Transmission Waiting Time: A Unifying Metric for Outbreak Controllability

Christopher Y. Shin[1,2,3,4], Sang Woo Park[5], Cécile Viboud[4], Pardis C. Sabeti[2, 6, 7, 8, 9], Christophe Fraser[10], and Kaiyuan Sun[11, 12*]

[1]Harvard College, Harvard University, Cambridge, MA, USA
[2]Broad Institute of MIT and Harvard, Cambridge, MA, USA
[3]Harvard Global Health Institute, Harvard University, Cambridge, MA, USA
[4]Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA
[5]School of Biological Sciences, Seoul National University, Seoul, South Korea
[6]Howard Hughes Medical Institute, Chevy Chase, MD, USA
[7]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA
[8]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA
[9]Department of Medicine, Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA, USA
[10]Pandemic Sciences Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[11]Center for Infectious Disease Research, School of Medicine, Westlake University, Hangzhou, Zhejiang Province, China.
[12]Center for Interdisciplinary Studies , School of Science, Westlake University, Hangzhou, Zhejiang Province, China.
[*]Corresponding author: sunkaiyuan@westlake.edu.cn

Abstract

The feasibility of containing infectious disease outbreaks depends on whether interventions can act faster than pathogen transmission. Classical controllability theory links outbreak control to the timing of symptom onset relative to infectiousness, reflecting an era in which isolation and contact tracing were necessarily triggered by clinical symptoms. This framework therefore implies intrinsic limits for pathogens with substantial pre-symptomatic or asymptomatic transmission. However, the widespread availability of molecular diagnostics now enables infection detection and isolation independent of symptoms, fundamentally altering the operational basis of outbreak control. Here we introduce the transmission waiting time—the interval between infection and the first onward transmission event—as a symptom-agnostic timescale that defines the intrinsic speed limit for effective intervention. We derive this quantity analytically from two fundamental epidemiological parameters, the basic reproduction number and the generation-interval distribution, yielding closed-form benchmarks for the minimum isolation speed and coverage required for control. Applying this framework across diverse pathogens reveals substantial differences in intrinsic controllability driven by variation in both overall transmissibility and the timing of infectiousness. We show that transmission heterogeneity reshapes controllability primarily through changes in the temporal distribution of infectiousness, and that homogeneous predictions often provide conservative bounds. By mapping complex test–trace–isolate operations onto a low-dimensional delay–coverage intervention landscape, this framework clarifies the conditions under which such containment efforts are feasible.

Main

The controllability of infectious disease outbreaks is determined by a race between pathogen transmission and the speed of public health intervention.[1,2] For decades, this race was governed by a symptom-based paradigm, in which isolation and contact tracing were initiated only after the onset of clinical signs. Formalized following the 2003 SARS-CoV-1 epidemic, classical theory framed outbreak controllability in terms of a pathogen's basic reproduction number, R_0, and the fraction of transmission occurring before symptom onset.[1] This framework explains the successful containment of pathogens such as SARS-CoV-1 and smallpox—where symptoms reliably precede infectiousness—and the intrinsic difficulty of controlling pathogens such as influenza, where substantial transmission occurs before symptoms.[3–5] Reliance on symptom-triggered response imposes a structural delay on the order of the incubation period, rendering containment infeasible when transmission is largely pre-symptomatic or when infections are frequently asymptomatic.

SARS-CoV-2 emerged in a fundamentally different intervention landscape: its extensive pre-symptomatic transmission rendered symptom-based control too slow, while the widespread availability of molecular diagnostics enabled infection detection independent of symptom presentation.[6–11] This shift moved the bottleneck for outbreak control from clinical recognition to operational constraints—such as testing frequency, diagnostic turnaround time, notification delay, and the timeliness of isolation—that jointly determine intervention coverage and delay. Concurrent technological advances, including smartphone-based digital contact tracing and rapid, high-sensitivity point-of-care diagnostics, further demonstrated the potential to shorten detection-to-notification intervals and accelerate response.[12–19] Yet, despite these advances, a unifying theory of outbreak controllability that links operational intervention delays to the intrinsic temporal constraints imposed by pathogen biology remains lacking.

Transmission waiting time: the critical timescale for outbreak control

To address this gap, we introduce a theoretical framework centered on the transmission waiting time, $w(\tau)$: the interval between infection and the first onward transmission event. In contrast to the classical concept of generation interval distribution $g(\tau)$—which describes the timing of all secondary transmissions—$w(\tau)$ isolates the earliest transmission event, the moment an intervention must pre-empt to block onward spread.

Failure to pre-empt this earliest transmission event leaves a substantial opportunity for generating secondary infections and sustaining transmission chains, rendering outbreak control infeasible. The transmission waiting time therefore defines an intrinsic speed limit for epidemic containment. This intuition constitutes the central principle of the manuscript and motivates the development of a rigorous theoretical framework for pathogen controllability in the sections that follow.

Starting from a homogeneous population, we derive $w(\tau)$ analytically from two fundamental epidemiological quantities of a pathogen: the basic reproduction number R_0, which encodes a pathogen's overall transmissibility, and the generation interval distribution $g(\tau)$, which characterizes the average temporal profile of infectiousness over the course of an infection episode. Under this framework, secondary transmissions arise according to a non-stationary Poisson process with transmission hazard $h(\tau) = R_0\, g(\tau)$. Using standard survival-analysis techniques (Methods), the transmission waiting time distribution, conditional on at least one secondary transmission occur-

ring, is given by

$$w(\tau) = \frac{R_0\, g(\tau) \exp\!\left(-R_0 \int_0^\tau g(s)\, ds\right)}{1 - \exp(-R_0)}. \tag{1}$$

This expression shows that the timing of the first transmission event is fully determined by R_0 and $g(\tau)$.

We next examine how the transmission waiting time constrains isolation-based interventions. Although such interventions vary widely in implementation—from symptom-driven isolation and manual contact tracing to routine rapid-test screening and digital exposure notification[20–22]—their epidemiological effects can be summarized by two operational quantities: the isolation coverage, P_{iso}, defined as the fraction of infections that are ultimately identified and isolated, and the isolation-delay distribution, $\delta_{\text{iso}}(\tau)$, which characterizes the timing of isolation among those individuals. The probability that an infected individual remains unisolated at time τ is therefore

$$S_{\text{iso}}(\tau) = 1 - P_{\text{iso}} \int_0^\tau \delta_{\text{iso}}(\tau')\, d\tau'. \tag{2}$$

Subsequently, the corresponding effective reproduction number under intervention is given by

$$R_e = \int_0^\infty R_0 g(\tau)\, S_{\text{iso}}(\tau)\, d\tau. \tag{3}$$

Under perfect coverage ($P_{\text{iso}} = 1$), if isolation delays exactly match the transmission waiting time ($\delta^*_{\text{iso}}(\tau) = w(\tau)$), the effective reproduction number reduces to

$$R_e^* = 1 - \frac{R_0 e^{-R_0}}{1 - e^{-R_0}}, \tag{4}$$

which is strictly less than one for all R_0 and approaches unity only asymptotically as R_0 becomes large (Methods). Thus, any intervention that acts at least as rapidly as the intrinsic transmission waiting time is, in principle, sufficient to guarantee outbreak control.

Because $w(\tau)$ is defined conditional on at least one secondary transmission, it integrates to unity even when the probability that an infection generates no onward transmission—given by e^{-R_0}—is substantial for relatively small values of R_0. In this low R_0 regime, requiring isolation delays to exactly match $w(\tau)$ under perfect isolation coverage is therefore unnecessarily stringent. Relaxing this constraint, we instead impose $R_e = 1$ while retaining $\delta^*_{\text{iso}}(\tau) = w(\tau)$, thereby yielding the critical isolation coverage (Methods).

$$P^*_{\text{iso}} = \frac{(R_0 - 1)(1 - e^{-R_0})}{R_0 - 1 + e^{-R_0}}, \qquad \begin{cases} R_e = 1, \\ \delta^*_{\text{iso}}(\tau) = w(\tau). \end{cases} \tag{5}$$

This quantity depends solely on R_0 and increases monotonically toward unity as transmissibility rises, indicating that pathogens with large R_0 require isolation of nearly all infectious individuals—even when isolation operates at the speed defined by the transmission waiting time. Together, $w(\tau)$ and P^*_{iso} delineate the minimal speed and reach of isolation required for containment, directly linking the feasibility of outbreak control to a pathogen's intrinsic transmissibility and temporal infectiousness profile.

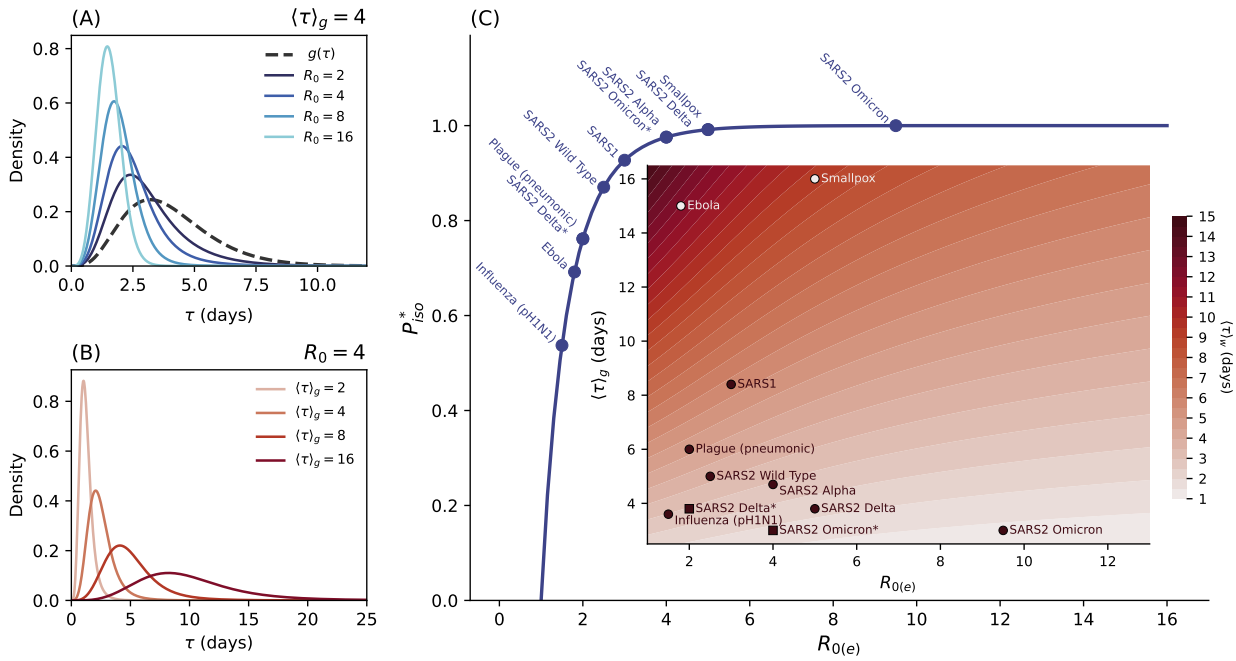


Figure 1: **Epidemiological determinants of transmission waiting time and isolation requirements.** (A) Effect of increasing the basic reproduction number R_0 on the transmission waiting time distribution $w(\tau)$, holding the generation-interval distribution $g(\tau)$ fixed (gamma-distributed with mean 4 days and shape parameter $\beta = 5$). The corresponding generation-interval distribution $g(\tau)$ is shown as a dashed curve. Larger R_0 shifts $w(\tau)$ toward earlier times, shortening the effective window for intervention. (B) Effect of decreasing the mean generation interval $\langle\tau\rangle_g$ on $w(\tau)$, holding $R_0 = 4$ and $\beta = 5$ fixed. Shorter generation intervals concentrate transmission earlier during infection, leading to similarly reduced waiting times. (C) Main panel: Relationship between the basic reproduction number R_0 and the minimum isolation coverage P^*_{iso} required for containment ($R_e = 1$). Points denote empirical estimates for selected pathogens, while the solid curve shows the theoretical prediction. Circle markers indicate basic reproduction numbers, and square markers indicate effective reproduction numbers. Inset: Heatmap showing how the mean transmission waiting time $\langle\tau\rangle_w$ varies jointly with R_0 and the mean generation interval, highlighting multiple epidemiological pathways to rapid transmission. SARS1: SARS-CoV-1; SARS2: SARS-CoV-2; Delta*: effective reproduction number of the SARS-CoV-2 Delta variant reduced to 2 under non-pharmaceutical interventions; Omicron*: effective reproduction number of the SARS-CoV-2 Omicron variant reduced to 4 due to prior immunity.

Epidemiological determinants of pathogen controllability

To assess the practical utility of the transmission waiting time framework, we first examined how waiting times depend on fundamental epidemiological parameters. Stochastic ordering analysis (Methods) shows that transmission waiting times become systematically shorter when transmission is more intense (higher R_0; Fig. 1A) and when infectiousness is concentrated earlier in infection (shorter generation intervals; Fig. 1B). Together, these results formalize the intuitive notion that both the overall magnitude and the temporal placement of infectiousness determine the intrinsic speed of transmission.

We next applied the transmission waiting time framework to a representative set of pathogens

that spread predominantly through direct person-to-person contact. For each pathogen, we compiled published estimates of the basic reproduction number R_0 and the generation-interval distribution $g(\tau)$. In addition, for selected SARS-CoV-2 variants, we also considered scenarios in which non-pharmaceutical interventions or prior immunity reduce transmission, and therefore report the effective reproduction number R_e instead of R_0. These inputs were used to derive the corresponding mean transmission waiting time, $\langle\tau\rangle_w$, together with the minimum isolation coverage required for containment, P^*_{iso}. Estimates of R_0 (or R_e, where applicable), the mean generation interval $\langle\tau\rangle_g$, the mean transmission waiting time $\langle\tau\rangle_w$, and the critical isolation threshold P^*_{iso} for all pathogens considered are summarized in Table 1.

Table 1: Epidemiological parameters and critical isolation requirements for selected pathogens. *indicates effective reproduction number R_e instead of basic reproduction number R_0 is reported. NPI: Non-pharmaceutical intervention

Pathogen	$\boldsymbol{R}_{0(\text{e})}$	$\langle\tau\rangle_g$	$\langle\tau\rangle_w$	$\boldsymbol{P}_{\text{iso}}$	**Reference(s)**
Ebola	1.8	15.0	11.96	0.69	[23,24]
Smallpox	5.0	16.0	9.24	0.99	[3,25,26]
SARS-CoV-1	3.0	8.4	5.83	0.93	[1,27]
Plague (pneumonic)	2.0	6.0	4.61	0.76	[28,29]
SARS-CoV-2 (Wild Type)	2.5	5.0	3.67	0.87	[30,31]
2009 H1N1 pandemic	1.5	3.6	2.98	0.54	[32,33]
SARS-CoV-2 (Delta*)	2.0	3.8	2.96	0.76	*$R_e = 2.0$ (under NPI)
SARS-CoV-2 (Alpha)	4.0	4.7	2.95	0.98	[34]
SARS-CoV-2 (Delta)	5.0	3.8	2.19	0.99	[35,36]
SARS-CoV-2 (Omicron*)	4.0	3.0	1.88	0.98	*$R_e = 4.0$ (prior immunity)
SARS-CoV-2 (Omicron)	9.5	3.0	1.37	1.00	[35,37–39]
Unit	–	days	days	–	–

Figure 1C visualizes these dependencies across the joint parameter space of R_0 and the mean generation interval, highlighting that rapid transmission can arise through multiple, non-mutually exclusive mechanisms: high transmissibility, rapid inter-generational turnover, or both. This comparative analysis reveals substantial variation in intrinsic controllability across pathogens.

Pathogens such as Ebola virus, which combine long generation intervals with moderate transmissibility, exhibit long transmission waiting times (approximately 12 days), providing relatively wide temporal windows for effective intervention.[23,24] Smallpox, despite having a generation interval of comparable length, is characterized by substantially higher transmissibility and therefore shorter transmission waiting times, requiring near-perfect isolation coverage and rendering containment operationally more challenging.[3,25,26] In contrast, the 2009 pandemic H1N1 influenza strain has a very short generation interval and consequently exhibits transmission waiting times of only 2–3 days even at modest values of R_0, illustrating that rapid inter-generational turnover alone can significantly constrain the feasibility of outbreak control.[32,33]

The evolutionary trajectory of SARS-CoV-2 further underscores the central role of the transmission waiting time in shaping pathogen controllability. The ancestral strain (wild type) exhibited intermediate waiting times (approximately 3–4 days), consistent with the transient containment achieved in several regions through early test–trace–isolate strategies.[30,31] However, successive variants of concern evolved toward both higher transmissibility and shorter generation intervals, resulting in progressively shorter transmission waiting times $\langle\tau\rangle_w$—a trend most clearly illustrated by the emergence of the Omicron variant.[34–39] With a transmission waiting time of roughly 1.37

days and a critical isolation coverage approaching unity, Omicron effectively placed isolation-based containment beyond operational reach in most settings.[35,37–39]

Importantly, both population-level non-pharmaceutical interventions (e.g., mobility restrictions) and accumulated population immunity can reduce the effective reproduction number and extend transmission waiting time. As shown in Table 1 and Fig. 1C, scenarios corresponding to the Delta variant under intervention and the Omicron variant in populations with substantial prior immunity exhibit longer transmission waiting times compared with their uninhibited spread in immunologically naïve populations. These comparisons highlight how reductions in effective transmissibility can meaningfully shift the transmission timescale, even when intrinsic viral properties favor rapid spread.

Controllability in the presence of transmission heterogeneity

The preceding analysis established outbreak controllability under the homogeneous assumption, in which all infected individuals share a common reproduction number R_0 and generation-interval distribution $g(\tau)$. Real-world populations, however, exhibit substantial heterogeneity in pathogen transmission. Infected individuals vary both in the number of secondary infections they generate and in the timing of infectiousness, and these two dimensions of heterogeneity are likely to be correlated.[40–42] A central question, therefore, is whether such heterogeneity qualitatively alters the conclusions derived under homogeneous assumptions, or whether it primarily manifests as a structured perturbation around that benchmark.

To address this question, we consider a population in which each infected individual is characterized by two latent attributes: an individual transmissibility parameter R, defined as the individual reproduction number, and an individual timing of infectiousness parameter μ_g, defined as the mean of the individual generation-interval distribution $g(\tau \mid \mu_g)$. Under this formulation, the individual-level transmission hazard is given by

$$h(\tau \mid R, \mu_g) = R\, g(\tau \mid \mu_g), \tag{6}$$

where $\mu_g = \int_0^\infty \tau\, g(\tau \mid \mu_g)\, d\tau$.

For a given pathogen, it is biologically plausible that R increases with μ_g, such that individuals who has longer mean generation interval tend, on average, to generate more secondary infections. To capture this dependence, we model the individual reproduction number as

$$R = R_0 \frac{\mu_g^\gamma}{\langle \mu_g^\gamma \rangle}\, \eta, \tag{7}$$

where μ_g follows a population-level distribution $f_g(\mu_g)$, reflecting heterogeneity in the timing of infectiousness. The parameter $\gamma \in [0, 2]$ governs the strength and form of the coupling between R and μ_g: $\gamma = 0$ corresponds to no dependence; $\gamma = 1$ implies linear scaling of R with μ_g; and $\gamma = 2$ represents linear scaling with total infectiousness over time, corresponding to the area under an individual infectiousness profile when both shedding intensity and duration increase with μ_g. The random variable η captures residual, μ_g-independent heterogeneity in transmissibility and its distribution $f_\eta(\eta)$ is assumed to be non-negative with unit mean $\langle \eta \rangle = \int_0^\infty \eta f_\eta(\eta)\, d\eta = 1$. Together with the normalization factor $\langle \mu_g^\gamma \rangle = \int_0^\infty \mu_g^\gamma f_g(\mu_g)\, d\mu_g$ this formulation ensures that the population average of individual reproduction number$\langle R \rangle = R_0$ for all choices of $\{f_g, \gamma, f_\eta\}$.

Because individual-level transmission attributes are generally unobservable in real time, interventions targeted to individual risk are impractical in most settings. We therefore consider uniform isolation strategies, characterized by a single isolation coverage P_{iso} and a shared isolation-delay

distribution $\delta_{\text{iso}}(\tau)$, applied identically across the population. Under this assumption, the probability that an infected individual remains unisolated at infection age τ is the same as in the homogeneous-mixing case: $S_{\text{iso}}(\tau) = 1 - P_{\text{iso}} \int_0^\tau \delta_{\text{iso}}(\tau')\, d\tau'$. The population-average effective reproduction number is then given by the expected number of secondary infections generated before isolation, averaged over the joint distribution of individual attributes μ_g and η:

$$R_e^{\text{pop}} = \mathbb{E}_{\mu_g,\eta}\left[\int_0^\infty h(\tau \mid R, \mu_g)\, S_{\text{iso}}(\tau)\, d\tau\right] \tag{8}$$

Notably, the residual heterogeneity term η, which captures variation in transmissibility independent of infectiousness timing, integrates out and does not affect the population-level effective reproduction number:

$$R_e^{\text{pop}} = R_0\, \mathbb{E}_{\mu_g}\left[\int_0^\infty \frac{\mu_g^\gamma}{\langle \mu_g^\gamma \rangle}\, g(\tau \mid \mu_g)\, S_{\text{iso}}(\tau)\, d\tau\right], \tag{9}$$

which implies R_e^{pop} depends only on heterogeneity in infectiousness timing and its correlation with transmissibility. Consequently, heterogeneity in transmissibility alone does not alter outbreak controllability at the population level; only heterogeneity that redistributes infectiousness in time—and its coupling to individual transmissibility R—can shift the isolation requirements relative to those predicted under homogeneous mixing.

The impact of heterogeneity in R on controllability under homogeneous μ_g

To elaborate on the point above, we first consider a limiting case in which transmission heterogeneity arises solely from variation in individual transmissibility R, while infectiousness timing is homogeneous across individuals. Specifically, the residual heterogeneity term η follows a gamma distribution with shape parameter k, and all individuals share an identical generation-interval distribution $g(\tau)$, implying no variation in μ_g.

Under these assumptions and in the absence of intervention, the offspring distribution is negative binomial with mean R_0 and dispersion parameter k, producing substantial variation in individual secondary infection counts when k is small. Nevertheless, because the population-level effective reproduction number R_e does not depend on η (Eq. (9)), and infectiousness timing is identical across individuals, the expression for R_e under intervention reduces exactly to the homogeneous-mixing result (Eq. (3)). Consequently, under the assumption of a homogeneous generation-interval distribution, variation in individual reproduction numbers—although it induces pronounced overdispersion in the distribution of secondary infections—does not modify the outbreak controllability when compared with the corresponding homogeneous reference scenario.

The impact of heterogeneity in infectiousness timing on controllability

To quantify how heterogeneity in infectiousness timing reshapes outbreak controllability, we consider a stylized pathogen with basic reproduction number $R_0 = 3$. Individual infectiousness timing parameter μ_g is drawn from a gamma distribution with mean $\langle \mu_g \rangle = 6$ days and shape parameter α, such that smaller α corresponds to greater heterogeneity, while $\alpha \to \infty$ recovers the homogeneous-mixing limit. Conditional on μ_g, the individual generation interval distribution $g(\tau \mid \mu_g)$ is gamma-distributed with fixed shape parameter and individual-specific mean (μ_g).

We allow individual transmissibility to scale with infectiousness timing according to $R \propto \mu_g^\gamma$, where γ controls the strength of coupling. All results assume uniform isolation with perfect coverage ($P_{\text{iso}} = 100\%$) and vary only the mean isolation delay $\langle \tau \rangle_{\text{iso}}$ (detailed in Methods).

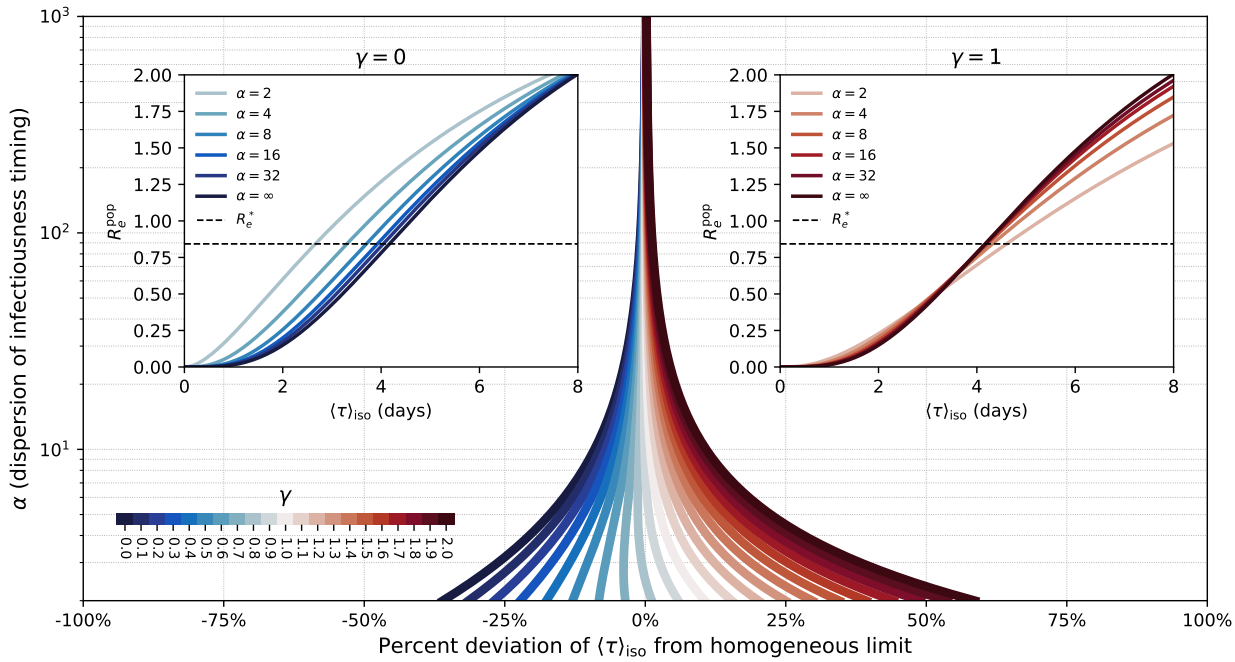


Figure 2: **Heterogeneity in infectiousness timing modifies outbreak controllability.** Inset panels show the population-average effective reproduction number R_e^{pop} as a function of the mean isolation delay $\langle\tau\rangle_{\text{iso}}$ for varying heterogeneity in infectiousness timing (α) under two correlation strengths between transmissibility and duration ($\gamma = 0$ and $\gamma = 1$). The horizontal dashed line denotes the critical threshold R_e^*, defined by the homogeneous transmission waiting time prediction under perfect isolation coverage (Eq. (4)). The main panel shows the percentage deviation in the isolation delay required to achieve $R_e^{\text{pop}} = R_e^*$ relative to the homogeneous-mixing benchmark.

Figure 2 illustrates how heterogeneity in μ_g, and its correlation with transmissibility, alters the isolation speed required for control. The inset panels show the population-average effective reproduction number R_e^{pop} as a function of $\langle\tau\rangle_{\text{iso}}$ for two limiting cases. When transmissibility is independent of infectiousness timing ($\gamma = 0$), increasing heterogeneity raises R_e^{pop} for a given isolation delay, requiring faster isolation than predicted under homogeneous mixing. This tightening reflects the increased contribution of early transmission from individuals with short generation intervals, which are difficult to pre-empt using a uniform intervention.

By contrast, when transmissibility increases with mean generation interval ($\gamma = 1$), these effects are strongly attenuated. Individuals with shorter μ_g also generate fewer secondary infections, partially offsetting their earlier transmission. As a result, isolation requirements remain close to the homogeneous-mixing prediction across a wide range of heterogeneity.

The main panel summarizes these effects by plotting, across values of α and γ, the percentage deviation in the mean isolation delay required to achieve $R_e^{\text{pop}} = R_e^*$ relative to the homogeneous benchmark. Deviations vanish as $\alpha \to \infty$, but become increasingly pronounced at low α. Increasing correlation strength progressively shifts these deviations toward zero and, under strong coupling, can even reverse their sign, indicating that heterogeneity may relax isolation-speed requirements when transmission is concentrated among individuals with longer, more interceptable infectiousness profiles.

Taken together, these results show that introducing heterogeneity does not overturn the qualitative conclusions of the transmission waiting time framework; instead, it acts as a structured

perturbation around the homogeneous limit. Given empirical evidence that infectiousness timing and overall shedding intensity are often positively correlated, the homogeneous theory therefore provides a conservative bound on outbreak controllability. Whether heterogeneity ultimately tightens or relaxes the speed requirements for containment is governed by how transmission potential is redistributed in time relative to intervention, rather than by variation in the individual reproduction number alone.

Operational determinants of pathogen controllability

The transmission waiting time framework links outbreak controllability to two operational quantities: the fraction of infections that are successfully isolated (isolation coverage) and the timing of isolation relative to infectiousness profile (isolation delay). A broad range of implementation-specific factors—including symptom recognition, testing frequency, diagnostic sensitivity, turnaround time, notification speed, contact tracing efficiency, and adherence—primarily influence pathogen controllability through their effects on these two quantities. We therefore evaluate intervention performance through the joint lens of isolation coverage and isolation delay, first using a stylized comparison between symptom-based and testing-based control to illustrate how advances in diagnostic technologies reshape intervention performance, and then embedding these insights within a more realistic intervention landscape using agent-based simulations.

Under the symptom-based intervention paradigm, isolation is initiated upon the onset of clinical symptoms. Consequently, the delay between infection and intervention is determined by the distribution of the incubation period—the time interval between infection and symptom onset—among individuals who develop symptoms. For pathogens with substantial pre-symptomatic transmission, this delay is structurally constrained to occur after a large fraction of transmission has already taken place.[7] Using SARS-CoV-2 as a case study, Fig. 3(A) compares incubation-period distributions with variant-specific transmission waiting times. From the ancestral lineage through Omicron, the mean incubation period consistently exceeds the corresponding transmission waiting time, indicating that symptom-triggered isolation systematically lags behind the earliest transmission events. The presence of a substantial fraction of asymptomatic infections further erodes effectiveness, rendering symptom-based containment infeasible for SARS-CoV-2. This limitation contrasts sharply with SARS-CoV-1. For SARS-CoV-1, the incubation period is shorter than the transmission waiting time (Fig. 3(A), inset), reflecting limited pre-symptomatic transmission. Combined with its high symptomatic rate, this temporal ordering—where symptom onset reliably precedes significant onward transmission—enabled symptom-based isolation and contact tracing to effectively interrupt transmission chains and achieve outbreak control, in alignment with the classical controllability theory established after the 2003 SARS outbreak.[1]

Testing-based interventions overcome the structural delay imposed by symptom onset through decoupling case detection from clinical presentation, enabling identification of infections regardless of symptom status. To illustrate this, we first consider a stylized test-and-isolate intervention: individuals are tested at regular intervals regardless of symptoms and immediately isolate upon receiving a positive result. The resulting intervention delay is governed by three components: the earliest time at which infection becomes detectable, determined by the assay's limit of detection relative to viral kinetics; the testing frequency; and the diagnostic turnaround time from sample collection to outcome notification.[10] Conventional diagnostic technologies involve a trade-off between sensitivity and speed: reverse transcription–polymerase chain reaction assays (RT–PCR, referred to hereafter as PCR) can detect low viral loads early in infection but often entail significant delays due to centralized laboratory processing, whereas rapid antigen lateral flow test (LFT) pro-

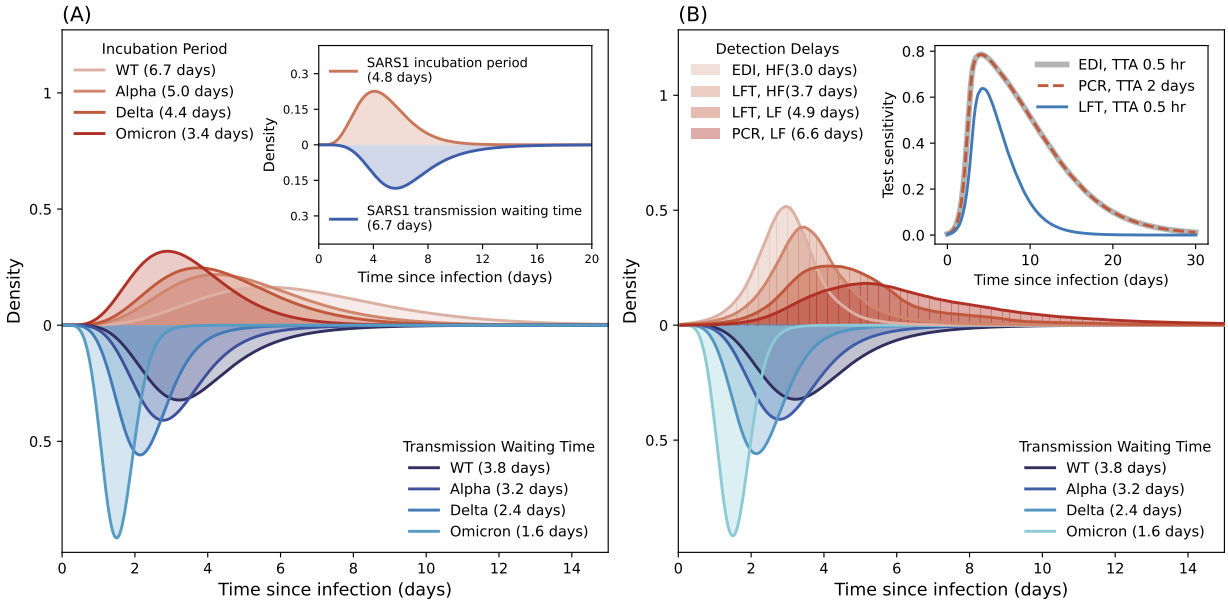


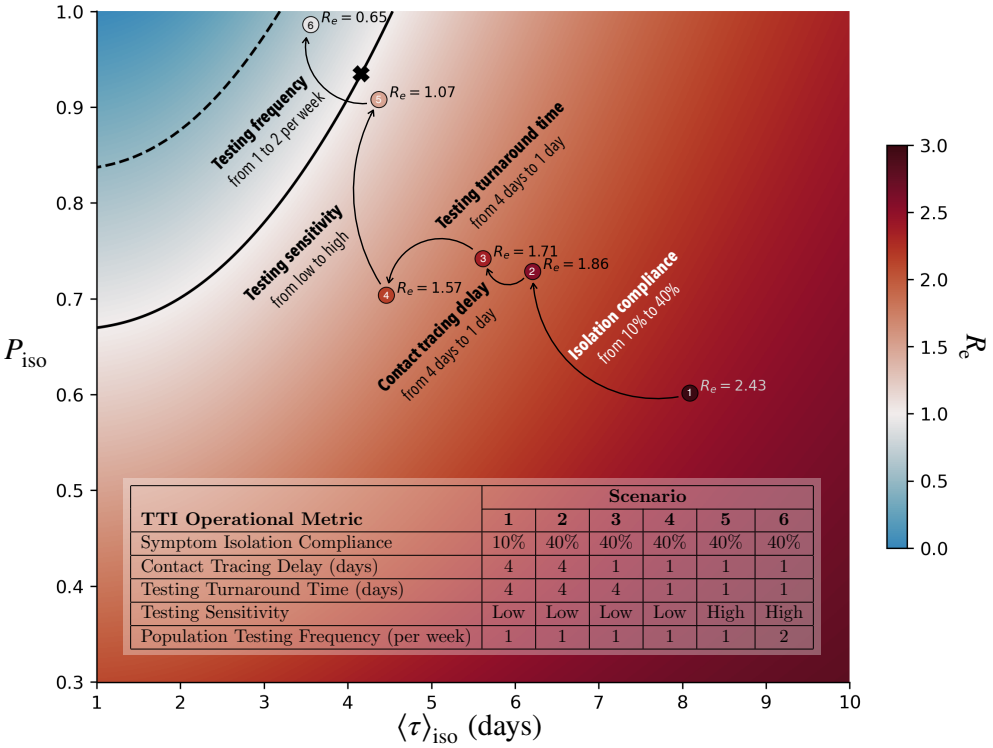
Figure 3: **Pathogen controllability through symptom-driven versus testing-driven isolation-based interventions. (A)** Distributions of incubation periods (symptom-based isolation delays) for SARS-CoV-2 variants (wild type (WT), Alpha, Delta, Omicron) compared with their corresponding transmission waiting times (means indicated in parentheses). Inset panel: distribution of incubation period and transmission waiting time for SARS-CoV-1 (SARS1, means indicated in parentheses). **(B)** Distributions of infection detection delays under population-wide testing strategies, compared with variant-specific transmission waiting times (means indicated in parentheses). We consider three diagnostic modalities: (1) PCR targeting viral RNA; (2) LFT targeting viral antigen; and (3) EDI that combine PCR-like sensitivity with LFT-like point-of-care speed. Test sensitivity as a function of time since infection as will as the testing turnaround (TTA) are reported in the inset panel. the Detection delays are shown for PCR and LFT under low-frequency (LF) testing, and for LFT and EDI under high-frequency (HF) testing.

vide near-immediate, point-of-care results yet typically identify infections only at later stages and higher viral loads, owing to their lower analytical sensitivity. Emerging molecular diagnostic innovation (EDI)—including CRISPR-based assays, isothermal amplification platforms, and portable microfluidic PCR systems—aim to combine high sensitivity with rapid turnaround.[43–49]

Figure 3(B) compares the distributions of isolation delays under population-wide testing strategies against variant-specific transmission waiting times. We assume a testing turnaround time of 2 days for PCR and 0.5 hours for both LFT and EDI. Two testing frequency scenarios are considered: a high-frequency (HF) scenario, in which the entire population is tested daily, and a low-frequency (LF) scenario, with population tested every three days. Test sensitivity as a function of time since infection is drawn from published literature for both PCR and LFT; for EDI, sensitivity is assumed to match that of PCR.[50]

Under the LF scenario, mean isolation delays exceed the transmission waiting time of the ancestral SARS-CoV-2 lineage (mean = 3.8 days)—6.6 days for PCR and 4.9 days for LFTs. Despite its higher sensitivity, PCR incurs longer delays due to its extended turnaround time (2 days vs. 0.5 hours). In contrast, the HF scenario—especially when combined with rapid-turnaround assays such as LFT and EDI—substantially compresses isolation delays to levels comparable to or shorter

than the transmission waiting times of early SARS-CoV-2 variants. High-frequency LFT testing reduces the mean detection delay to 3.7 days, while EDI further shorten it to approximately 3.0 days—sufficient to control early variants like Alpha without additional contact tracing or population level travel restrictions. Across all strategies evaluated, fewer than 10% of infections were missed, reflecting high intervention coverage due to effective infection ascertainment.



TTI Operational Metric	Scenario 1	2	3	4	5	6
Symptom Isolation Compliance	10%	40%	40%	40%	40%	40%
Contact Tracing Delay (days)	4	4	1	1	1	1
Testing Turnaround Time (days)	4	4	4	1	1	1
Testing Sensitivity	Low	Low	Low	Low	High	High
Population Testing Frequency (per week)	1	1	1	1	1	2

Figure 4: **Intervention landscape linking isolation timeliness and coverage to outbreak controllability under test–trace–isolate (TTI) strategies.** The heatmap shows the effective reproduction number R_e as a function of the mean isolation delay $\langle\tau\rangle_{\text{iso}}$ and isolation coverage P_{iso}, assuming moderate background population-level mitigation. Contour lines indicate the theoretical containment boundary ($R_e = 1$) in the absence of both population-level and case-based interventions ($R_0 = 6$, dashed line) and under population-level control alone ($R_e^{\text{pop}} = 3$, solid line). Conditional on background population-level measures, colored markers denote simulated TTI scenarios with incremental improvements in operational performance, while the black cross marks the theoretical control condition at which the intervention timescale coincides with the transmission waiting time, $\langle\tau\rangle_{\text{iso}} = \langle\tau\rangle_w$. **Inset table:** summary of the six incremental TTI scenarios.

Population-wide access to routine molecular testing can shift the distribution of isolation delays toward—and in some cases ahead of—the intrinsic transmission timescale that defines pathogen controllability within our framework. In practice, however, outbreak control rarely relies on testing alone. Instead, control is typically achieved through layered intervention systems that combine symptom-based surveillance, population-level mobility restrictions, routine diagnostic screening, and contact tracing, each operating under logistical constraints that introduce delays and limit coverage. As a result, real-world intervention systems are substantially more complex than the idealized scenarios considered above. Their effectiveness is constrained by operational limitations, such as finite testing capacity, diagnostic turnaround delays, incomplete contact tracing, and imper-

fect adherence. These factors jointly shape the performance of non-pharmaceutical interventions (NPIs), particularly test–trace–isolate (TTI) strategies.[20–22] To bridge the gap between the theoretical controllability thresholds defined by the transmission waiting time framework and realistic outbreak-control implementation, we therefore conducted a simulation-based intervention analysis using a branching process model.

Specifically, we simulated a hypothetical respiratory pathogen with a basic reproduction number $R_0 = 6$ and a mean generation interval of 6 days. We assumed the presence of background population-level interventions—such as travel restrictions—that reduced the effective reproduction number to $R_e^{\text{pop}} = 3$ prior to the implementation of TTI interventions. Against this population-wide intervention backdrop, we evaluated six TTI scenarios, each representing incremental improvements in symptom-based isolation compliance, testing frequency and sensitivity, diagnostic turnaround time, and contact-tracing delay (Fig. 4, inset table). To connect these detailed operational components to our theoretical framework, we summarize intervention performance using two aggregate quantities: the mean isolation delay, $\langle\tau\rangle_{\text{iso}}$, which captures the timeliness of isolation relative to infection, and the isolation coverage, P_{iso}, which represents the fraction of infections successfully isolated. Collectively, these scenarios span a realistic range of operational performance observed in practice. Full model details are provided in the Methods.

Figure 4 situates these scenarios within an intervention landscape defined by isolation delay and isolation coverage. The heatmap shows the effective reproduction number R_e as a function of $(\langle\tau\rangle_{\text{iso}}, P_{\text{iso}})$. Contour lines denote the theoretical containment boundary ($R_e = 1$) under moderate population-level mitigation ($R_e^{\text{pop}} = 3$, solid line) and in the absence of population-wide mitigation ($R_0 = 6$, dashed line). Conditional on the assumed background population-level control, colored markers indicate the simulated test–trace–isolate (TTI) scenarios, while the black cross marks the theoretical control condition at which the mean isolation delay coincides with the transmission waiting time, $\langle\tau\rangle_{\text{iso}} = \langle\tau\rangle_w$.

As operational performance improves—through reductions in isolation delay and/or increases in isolation coverage—scenarios move systematically across the landscape toward the containment boundary, accompanied by progressively lower effective reproduction numbers. The most effective configuration (Scenario 6), which combines high testing frequency, rapid diagnostic turnaround, efficient contact tracing, and background population-level mitigation, crosses this boundary and achieves $R_e < 1$.

Together, these results underscore the essential and complementary roles of intervention speed and coverage in outbreak control. By projecting complex, multi-component intervention systems onto a low-dimensional delay–coverage plane, the transmission waiting time framework provides a quantitative bridge between operational constraints and intrinsic controllability timescales. This representation offers a principled basis for comparing, evaluating, and optimizing TTI strategies under realistic implementation conditions.

Discussion

Over the past two decades, the operational determinants of outbreak control have shifted fundamentally. During the 2003 SARS-CoV-1 epidemic, containment relied primarily on symptom-based case detection and contact tracing. In contrast, the extensive pre-symptomatic and asymptomatic transmission of SARS-CoV-2 rendered symptom-triggered interventions structurally too slow, even when implemented at scale.[1,6–11] At the same time, widespread molecular diagnostics, routine screening, and digital exposure notification enabled symptom-independent identification and isolation, shifting the effective bottleneck from clinical recognition to operational performance—including testing fre-

quency, turnaround time, notification delays, tracing efficiency, and adherence.[12–19,51] Motivated by this transition, we developed a symptom-independent theory of outbreak controllability that compares the timing of interventions directly against the earliest opportunities for transmission.

Our framework generalizes the influential paradigm introduced by Fraser et al., which linked controllability to the basic reproduction number R_0 and the timing of symptom onset relative to infectiousness.[1] Whereas classical theory emphasizes the structural delay imposed by symptom-based surveillance, we anchor containment to a more fundamental temporal constraint: the first onward transmission event. We formalize this constraint through the transmission waiting time, $w(\tau)$, defined as the interval between infection and the first secondary transmission, conditional on at least one transmission occurring. Because failure to pre-empt this first event leaves substantial remaining opportunity for onward spread, $w(\tau)$ defines an intrinsic speed limit for isolation-based containment. Under homogeneous transmission, we show that $w(\tau)$ is fully determined by R_0 and the generation-interval distribution $g(\tau)$, yielding closed-form benchmarks for both the critical isolation-delay distribution and the minimum isolation coverage required to achieve $R_e \leq 1$. Because R_0 and $g(\tau)$ can often be estimated at the early phase of an outbreak, this formulation provides a practical basis for rapid assessment of whether isolation-based containment is biologically plausible, independent of whether case detection is symptom-driven or test-driven.[52]

A central implication of the transmission waiting time framework is that intrinsic controllability cannot be inferred from R_0 alone. Applying the framework across diverse directly transmitted pathogens reveals that pathogens with similar R_0 may nonetheless differ markedly in containment feasibility because of differences in transmission timing. Pathogens characterized by long generation intervals, such as Ebola and smallpox, admit relatively wide temporal windows for intervention, consistent with historical containment successes. In contrast, pathogens such as influenza can be intrinsically difficult to contain even at modest R_0, because rapid inter-generational turnover compresses the intervention window to only a few days.[3,4,32,33,53] The evolutionary trajectory of SARS-CoV-2 further illustrates how these timing constraints can tighten over short timescales: successive variants combined increasing transmissibility with progressively shorter transmission waiting times, pushing the critical isolation coverage toward unity and eroding the feasibility of isolation-based containment in most settings.[34–39]

The framework also provides a parsimonious operational interpretation of complex intervention systems. We show that heterogeneous components of test–trace–isolate (TTI) strategies can be summarized by two quantities: isolation coverage P_{iso} and the isolation-delay distribution $\delta_{\text{iso}}(\tau)$.[20–22] Testing frequency, diagnostic sensitivity, and turnaround time primarily shift $\delta_{\text{iso}}(\tau)$ earlier, whereas tracing efficiency and adherence primarily increase P_{iso}. Our simulation analysis illustrates how incremental improvements along these axes move realistic TTI programs through a low-dimensional delay–coverage landscape toward the containment boundary, clarifying why isolated improvements are often insufficient: for fast-transmitting pathogens, containment typically requires coordinated gains that simultaneously compress delays and increase coverage. This representation also accommodates additional interventions—such as masking, ventilation, distancing, or temporary closures—through their effects on contact rates and interaction structure, which effectively modify the realized infectiousness profile and shift the containment frontier.[54]

We further show that transmission heterogeneity modifies controllability primarily through temporal redistribution of infectiousness, rather than through variation in individual reproduction numbers alone. Under uniform isolation, heterogeneity in individual transmissibility that is independent of infectiousness timing integrates out of the population-average R_e and therefore does not alter the containment condition relative to the homogeneous benchmark. By contrast, heterogeneity in infectiousness timing can either tighten or relax isolation-speed requirements depending on its coupling to transmissibility.[42] When transmissibility is weakly coupled to mean generation interval,

greater dispersion increases the contribution of early transmission from short-duration infections, tightening speed requirements. When transmissibility increases with duration, these effects are attenuated—and can even reverse—because transmission is concentrated among longer-lasting infections that are more interceptable by interventions. Taken together, these results suggest that homogeneous predictions provide a robust benchmark and, under empirically plausible positive correlations between infectiousness timing and overall shedding intensity, often a conservative bound on the speed required for containment.

Several limitations highlight priorities for future work. First, our analytical results rely on simplified mixing assumptions and uniform intervention effects, whereas real outbreaks occurs on structured contact networks with setting-specific risks and correlated behaviors that can shape both transmission timing and intervention performance.[12,51,52,54] Extending the framework to incorporate contact structure and heterogeneity in intervention uptake—while retaining a tractable mapping to delay and coverage—would improve realism and interpretability. Second, we treat R_0 and $g(\tau)$ as fixed pathogen characteristics, although both vary across contexts due to behavior, immunity, and non-pharmaceutical interventions. Time-varying natural history and immune landscapes may therefore shift $w(\tau)$ and the controllability over the course of an epidemic. Third, we do not explicitly model feedbacks between control and pathogen evolution; early emergence can involve rapid adaptation that alters transmissibility, immune escape, or within-host kinetics, thereby changing the relevant transmission timescale.[9] Integrating evolutionary dynamics with operational constraints remains an important direction for prospective risk assessment. Finally, our intervention model is restricted to single-step contact tracing. Iterative tracing, particularly when accelerated by digital technologies and paired with rapid molecular diagnostics, warrants further exploration as a means to substantially strengthen outbreak control.

In conclusion, the transmission waiting time provides a symptom-independent timescale for outbreak controllability in the modern era of molecular diagnostics. By anchoring containment to the earliest transmission event, the framework yields interpretable benchmarks for the minimum speed and reach of isolation required for control, enables principled comparisons of intrinsic controllability across pathogens, and maps complex intervention systems onto a delay–coverage landscape that directly reflects operational levers. These features make the framework well suited to guide early-stage decision-making, prioritize investments in testing and tracing performance, and clarify when isolation-based containment is feasible versus when mitigation becomes the only realistic option.

Methods

1 Analytical expression for the transmission waiting time under homogeneous mixing

To quantify the timing of the earliest secondary transmission event generated by an infected individual, we adopt a survival-analysis framework and define the transmission waiting time as the interval between the time of primary infection and the occurrence of the first onward transmission event.[55] Under homogeneous mixing, individual infectiousness over infection age τ is fully characterized by the basic reproduction number R_0 and the generation-interval distribution $g(\tau)$.

We assume secondary transmission events arise from a non-homogeneous Poisson process with hazard

$$h(\tau) = R_0 g(\tau), \tag{M1}$$

where $g(\tau)$ is a probability density function satisfying

$$\int_0^\infty g(\tau)\, d\tau = 1. \tag{M2}$$

Let $G(\tau) = \int_0^\tau g(s)\, ds$ denote the corresponding cumulative distribution function. The survival function—the probability that no secondary transmission has occurred by time τ—is

$$S(\tau) = \exp\left(-\int_0^\tau h(s)\, ds\right) = \exp(-R_0 G(\tau))\,. \tag{M3}$$

Therefore, the (defective) cumulative distribution function for the time of the first secondary transmission is

$$F(\tau) = 1 - S(\tau) = 1 - \exp(-R_0 G(\tau))\,, \tag{M4}$$

and the corresponding (defective) density is

$$f(\tau) = \frac{dF(\tau)}{d\tau} = h(\tau)S(\tau) = R_0 g(\tau)\exp(-R_0 G(\tau))\,. \tag{M5}$$

Since $S(\infty) = \exp(-R_0)$, the probability that an infection generates at least one secondary transmission is

$$F(\infty) = 1 - \exp(-R_0). \tag{M6}$$

Conditioned on at least one secondary transmission occurring, the normalized transmission waiting-time density, denoted $w(\tau)$, is

$$w(\tau) = \frac{f(\tau)}{F(\infty)} = \frac{R_0 g(\tau)\exp(-R_0 G(\tau))}{1 - \exp(-R_0)}. \tag{M7}$$

Together, R_0 controls the overall intensity of transmission while $g(\tau)$ governs the temporal profile of infectiousness; both jointly determine the distribution of the earliest transmission event.

2 How pathogen characteristics shape waiting time: stochastic ordering analysis

This section shows how $w(\tau)$ shifts when (i) transmissibility increases (larger R_0) or (ii) infectiousness occurs earlier (a stochastically smaller generation interval).

The conditional CDF of the transmission waiting time is obtained by integrating (M7):

$$W(\tau \mid R_0, g) = \int_0^\tau w(s)\, ds = \frac{1 - \exp(-R_0 G(\tau))}{1 - \exp(-R_0)}. \tag{M8}$$

2.1 Dependence on the reproduction number R_0

Fix g (hence $G(\tau)$). For any fixed τ with $G(\tau) \in (0,1)$, define

$$\phi(x; G) = \frac{1 - e^{-xG}}{1 - e^{-x}}, \qquad x > 0.$$

Then $W(\tau \mid R_0, g) = \phi(R_0; G(\tau))$. One can verify that $\phi(x; G)$ is strictly increasing in x for each fixed $G \in (0,1)$ (e.g., by differentiating and showing $\phi'(x; G) > 0$). Therefore, if $R_0' > R_0$, then for all τ,

$$W(\tau \mid R_0', g) \;>\; W(\tau \mid R_0, g), \tag{M9}$$

implying first-order stochastic dominance: the waiting time under R_0' is stochastically smaller (earlier) than under R_0.

2.2 Dependence on the generation interval distribution

Fix R_0. Suppose g is stochastically smaller than g' in the sense that

$$G(\tau) \geq G'(\tau) \quad \text{for all } \tau, \tag{M10}$$

i.e., g concentrates more mass earlier in time. Define

$$\psi(y; R_0) = \frac{1 - e^{-R_0 y}}{1 - e^{-R_0}}, \qquad y \in [0, 1],$$

which is strictly increasing in y. Since $W(\tau \mid R_0, g) = \psi(G(\tau); R_0)$, condition (M10) implies

$$W(\tau \mid R_0, g) \;\geq\; W(\tau \mid R_0, g') \quad \text{for all } \tau, \tag{M11}$$

so the transmission waiting time associated with the "faster" generation interval g is stochastically smaller (earlier) than that associated with g'.

3 Derivation of critical intervention thresholds (homogeneous mixing case)

3.1 Effective reproduction number R_e^* under the critical isolation timing benchmark and perfect coverage

Consider an isolation-based intervention characterized by coverage P_{iso} and an isolation-delay density $\delta_{\text{iso}}(\tau)$ among those isolated. The probability that an infected individual remains unisolated at infection age τ is

$$S_{\text{iso}}(\tau) = 1 - P_{\text{iso}} \int_0^\tau \delta_{\text{iso}}(s)\, ds. \tag{M12}$$

Under perfect coverage ($P_{\text{iso}} = 1$) and the critical timing benchmark $\delta_{\text{iso}}^*(\tau) = w(\tau)$, we have

$$S_{\text{iso}}^*(\tau) = 1 - \int_0^\tau w(s)\, ds = 1 - W(\tau \mid R_0, g).$$

Using (M8), this becomes

$$S_{\text{iso}}^*(\tau) = \frac{\exp(-R_0 G(\tau)) - \exp(-R_0)}{1 - \exp(-R_0)}. \tag{M13}$$

The effective reproduction number under intervention is

$$R_e = \int_0^\infty R_0 g(\tau)\, S_{\text{iso}}(\tau)\, d\tau. \tag{M14}$$

Plugging (M13) into (M14) yields

$$\begin{aligned} R_e^* &= \int_0^\infty R_0 g(\tau)\, \frac{\exp(-R_0 G(\tau)) - \exp(-R_0)}{1 - \exp(-R_0)}\, d\tau \\ &= \frac{R_0}{1 - \exp(-R_0)} \left[\int_0^\infty g(\tau) \exp(-R_0 G(\tau))\, d\tau - \exp(-R_0) \int_0^\infty g(\tau)\, d\tau \right]. \end{aligned} \tag{M15}$$

Using the substitution $u = G(\tau)$ (so $du = g(\tau)\, d\tau$), the first integral becomes $\int_0^1 e^{-R_0 u}\, du = (1 - e^{-R_0})/R_0$, and the second integral equals 1. Therefore,

$$R_e^* = 1 - \frac{R_0 e^{-R_0}}{1 - e^{-R_0}}. \tag{M16}$$

3.2 Minimum required isolation coverage P^*_{iso} under the critical timing benchmark

Now impose the same critical timing benchmark $\delta^*_{\text{iso}}(\tau) = w(\tau)$ but allow imperfect coverage $P_{\text{iso}} < 1$. From (M12),

$$S_{\text{iso}}(\tau) = 1 - P_{\text{iso}}\, W(\tau \mid R_0, g). \tag{M17}$$

Setting the containment threshold $R_e = 1$ in (M14) gives

$$\begin{aligned} 1 &= \int_0^\infty R_0 g(\tau)\,[1 - P_{\text{iso}}\, W(\tau \mid R_0, g)]\, d\tau \\ &= R_0 - R_0 P_{\text{iso}} \int_0^\infty g(\tau)\, W(\tau \mid R_0, g)\, d\tau. \end{aligned} \tag{M18}$$

Using (M8), and substituting $u = G(\tau)$,

$$\int_0^\infty g(\tau)\, W(\tau \mid R_0, g)\, d\tau = \int_0^1 \frac{1 - e^{-R_0 u}}{1 - e^{-R_0}}\, du = \frac{1}{1 - e^{-R_0}} \left(1 - \frac{1 - e^{-R_0}}{R_0}\right). \tag{M19}$$

Plugging (M19) into (M18) and solving for P_{iso} yields

$$P^*_{\text{iso}} = \frac{(R_0 - 1)(1 - e^{-R_0})}{R_0 - 1 + e^{-R_0}}. \tag{M20}$$

4 Computational framework for heterogeneity analysis

This section specifies the model and numerical procedures used to generate Figure 2. We quantify how heterogeneity in infectiousness timing reshapes controllability under a uniform isolation intervention with perfect coverage ($P_{\text{iso}} = 100\%$). Notation follows the main text.

4.1 Heterogeneity model for infectiousness timing and transmissibility

Each infection is characterized by an individual generation-interval distribution parameterized by its mean μ_g. Conditional on μ_g, we assume

$$g(\tau \mid \mu_g) \sim \text{Gamma}(\text{shape} = \beta,\ \text{scale} = \theta_g), \qquad \beta = 5, \qquad \theta_g = \frac{\mu_g}{\beta}, \tag{M21}$$

so that $\mathbb{E}[\tau \mid \mu_g] = \mu_g$.

Across individuals, the mean timing parameter varies as

$$\mu_g \sim \text{Gamma}(\text{shape} = \alpha,\ \text{scale} = \theta_\mu), \qquad \theta_\mu = \frac{\langle \mu_g \rangle}{\alpha}, \qquad \langle \mu_g \rangle = 6 \text{ days}. \tag{M22}$$

Smaller α corresponds to greater dispersion in infectiousness timing, while $\alpha \to \infty$ recovers the homogeneous limit $\mu_g = \langle \mu_g \rangle$.

Individual transmissibility is allowed to scale with infectiousness timing via

$$R(\mu_g) = R_0\, \frac{(\mu_g / \langle \mu_g \rangle)^\gamma}{\mathbb{E}[(\mu_g / \langle \mu_g \rangle)^\gamma]}, \qquad R_0 = 3, \qquad \gamma \in [0, 2], \tag{M23}$$

which ensures $\mathbb{E}[R(\mu_g)] = R_0$ for all (α, γ).

4.2 Isolation-based intervention model and homogeneous limit

We consider a uniform isolation-based intervention with perfect coverage ($P_{\text{iso}} = 100\%$), parameterized by the mean isolation delay $\langle\tau\rangle_{\text{iso}}$. The key design feature of the intervention model is that $\langle\tau\rangle_{\text{iso}}$ is defined in the same waiting-time framework as pathogen transmission, ensuring a well-defined and interpretable homogeneous limit as $\alpha \to \infty$. Details of the numerical implementation is as follows:

1. We first introduce an auxiliary generation-interval distribution with the same shape parameter of the individual generation interval distribution

$$g_{\text{iso}}(\tau) \sim \text{Gamma}(\text{shape} = \beta,\ \text{scale} = \theta_{\text{iso}}), \qquad \beta = 5, \qquad \theta_{\text{iso}} = \frac{\mu_{\text{iso}}}{\beta}, \tag{M24}$$

   where $\mu_{\text{iso}} = \mathbb{E}_{g_{\text{iso}}}[\tau]$ serves as a numerical control parameter that sets the characteristic timescale of isolation delay.

2. We then define the isolation-delay distribution $\delta_{\text{iso}}(\tau)$ as the transmission waiting time distribution induced by g_{iso} under a homogeneous mixing assumption with reproduction number $R_0 = 3$:

$$\delta_{\text{iso}}(\tau) \equiv w(\tau; R_0, g_{\text{iso}}), \qquad R_0 = 3. \tag{M25}$$

   The corresponding survival function entering the transmission kernel is

$$S_{\text{iso}}(\tau) = 1 - \int_0^\tau \delta_{\text{iso}}(s)\, ds, \tag{M26}$$

   and the mean isolation delay reported in the figure is the mean of the constructed isolation delay distribution,

$$\langle\tau\rangle_{\text{iso}} = \int_0^\infty \tau\, \delta_{\text{iso}}(\tau)\, d\tau. \tag{M27}$$

This construction ensures that the intervention is expressed in the same waiting-time language as the homogeneous mixing scenario. In particular, when infectiousness timing becomes homogeneous ($\alpha \to \infty$, $\mu_g \to \langle\mu_g\rangle$), the population transmission process is fully characterized by the homogeneous generation-interval distribution $g(\tau \mid \langle\mu_g\rangle)$, and the corresponding transmission waiting time reduces to $w(\tau; R_0, g(\tau \mid \langle\mu_g\rangle))$. Choosing δ_{iso} to match this distribution therefore recovers the homogeneous benchmark for controllability. Deviations in the critical isolation delay for finite α can thus be interpreted directly as the effect of heterogeneity relative to this homogeneous limit.

4.3 Population-level R_e^{pop}, benchmark, and deviation metric

For an individual with parameters $(\mu_g, R(\mu_g))$, the expected number of secondary infections generated prior to isolation is

$$R(\mu_g) \int_0^\infty g(\tau \mid \mu_g)\, S_{\text{iso}}(\tau)\, d\tau. \tag{M28}$$

Averaging over the population distribution of μ_g yields

$$R_e^{\text{pop}}(\langle\tau\rangle_{\text{iso}}) = \mathbb{E}_{\mu_g}\left[R(\mu_g) \int_0^\infty g(\tau \mid \mu_g)\, S_{\text{iso}}(\tau)\, d\tau\right]. \tag{M29}$$

Numerically, we precompute the kernel

$$K(\mu_g, \langle\tau\rangle_{\text{iso}}) = \int_0^\infty g(\tau \mid \mu_g)\, S_{\text{iso}}(\tau)\, d\tau, \tag{M30}$$

on grids $\mu_g \in [0.1, 100]$ days (1000 points) and $\tau \in [0, 100]$ days with time step $\Delta\tau = 0.01$ days.

The dashed horizontal reference in the inset panels of Figure 2 is

$$R_e^* = 1 - \frac{R_0 e^{-R_0}}{1 - e^{-R_0}}, \qquad R_0 = 3, \tag{M31}$$

which corresponds to the homogeneous waiting-time benchmark under perfect coverage (i.e., when δ_{iso} matches the homogeneous transmission waiting time distribution).

For each (α, γ), we evaluate R_e^{pop} on the $\langle\tau\rangle_{\text{iso}}$ grid induced by the construction above and numerically identify $\langle\tau\rangle^*_{\text{iso}}(\alpha, \gamma)$ as the value satisfying

$$R_e^{\text{pop}}(\langle\tau\rangle^*_{\text{iso}}) = R_e^*. \tag{M32}$$

The main panel reports the deviation of this critical isolation delay relative to the homogeneous mean transmission waiting time $\langle\tau\rangle_w$ (computed by applying the same waiting-time transformation to the homogeneous $g(\tau \mid \langle\mu_g\rangle)$):

$$\frac{\langle\tau\rangle^*_{\text{iso}} - \langle\tau\rangle_w}{\langle\tau\rangle_w}, \tag{M33}$$

with axis tick labels displayed as percentages.

5 Branching process model for TTI simulations

We developed a stochastic branching-process model to evaluate the effectiveness of test–trace–isolate (TTI) strategies under realistic operational constraints. Transmission unfolds over discrete generations, with each infected individual generating a random number of secondary infections. In the absence of intervention, the basic reproduction number is fixed at $R_0 = 6$, and generation intervals are drawn from a gamma distribution with shape parameter $k = 5$ and mean 6 days.

5.1 Triggering of interventions and population-level control

The outbreak is allowed to evolve without intervention for four generations. A public-health response is then initiated at

$$t_{\text{intervention}} = 4 \times \langle\tau\rangle_g,$$

where $\langle\tau\rangle_g$ denotes the mean generation interval. At this time, population-level control measures—such as mobility restrictions and behavioral changes—are introduced, reducing transmissibility to a post-intervention baseline value $R_e^{\text{pop}} = 3$ in the absence of case-based interventions (i.e., without testing, tracing, or isolation).

5.2 Case-based isolation mechanisms

Following the onset of interventions, infected individuals may be isolated through three pathways:

1. **Symptom-based isolation.** Symptomatic individuals self-isolate upon symptom onset with probability $p^{\text{iso}}_{\text{sym}}$. Incubation periods are drawn from a gamma distribution with shape parameter 5 and mean 4 days. If symptom onset occurs after $t_{\text{intervention}}$, isolation is assumed to occur immediately at symptom onset.

2. **Routine diagnostic testing.** Individuals are screened at rate λ, independent of symptom status. Test sensitivity varies over the course of infection and is modeled as proportional to a proxy for infectiousness, $g(\tau)^{\alpha_{\text{test}}}$, scaled by a factor β_{test}. Following a positive test result, individuals isolate after a diagnostic turnaround delay drawn from a gamma distribution with shape parameter 5 and mean $\overline{\Delta\tau}_{\text{testing}}$.

3. **One-step contact tracing.** Each confirmed case triggers tracing of its direct transmission offspring. The contact tracing delay is drawn from a gamma distribution with shape parameter 5 and mean $\overline{\Delta\tau}_{\text{tracing}}$. Traced contacts isolate immediately upon notification, regardless of symptom status or test results.

5.3 Computational implementation of intervention effects

To quantify intervention impact, we reconstruct the full transmission tree and remove any transmission events that occur after isolation of the infector, together with all downstream infections arising from those prevented events. The effective reproduction number under intervention, R_e, is calculated as the mean number of secondary infections generated by individuals infected after $t_{\text{intervention}}$.

From the simulated transmission trees, we extract three summary metrics: (i) the effective reproduction number R_e among post-intervention infections; (ii) the isolation coverage P_{iso}, defined as the probability that an infected individual is isolated before generating any onward transmission; and (iii) the mean isolation delay $\langle\tau\rangle_{\text{iso}}$ among isolated individuals.

5.4 Incremental intervention scenarios

We simulate six increasingly stringent TTI scenarios, each representing incremental improvements in operational performance. All scenarios assume $R_e^{\text{pop}} = 3$, a mean generation interval of 6 days, and a mean incubation period of 4 days. Parameters are modified stepwise as follows:

- **Scenario 1 (baseline):** $p_{\text{sym}}^{\text{iso}} = 0.1$, weekly testing ($\lambda = 1/7$), mean testing turnaround time of 4 days, and mean contact tracing delay of 4 days.
- **Scenario 2:** Increase symptom-based isolation compliance to $p_{\text{sym}}^{\text{iso}} = 0.4$.
- **Scenario 3:** Reduce the mean contact tracing delay to 1 day.
- **Scenario 4:** Reduce the mean testing turnaround time to 1 day.
- **Scenario 5:** Increase test sensitivity by setting $\alpha_{\text{test}} = 0.5$ and $\beta_{\text{test}} = 0.9$.
- **Scenario 6:** Increase testing frequency to twice per week ($\lambda = 2/7$).

These scenarios illustrate the cumulative benefits of improving individual TTI components. As intervention performance improves through faster isolation, higher coverage, and more frequent and sensitive testing, the effective reproduction number declines. Under the most stringent configuration, the combined system crosses the containment threshold and achieves $R_e < 1$.

References

[1] Fraser, C.; Riley, S.; Anderson, R. M.; Ferguson, N. M. *Proceedings of the National Academy of Sciences* **2004**, *101*, 6146–6151.

[2] Anderson, R. M.; May, R. M. *Infectious Diseases of Humans: Dynamics and Control*; Oxford University PressOxford, 1991.

[3] Gani, R.; Leach, S. *Nature* **2001**, *414*, 748–751.

[4] Swanson, K. C.; Altare, C.; Wesseh, C. S.; Nyenswah, T.; Ahmed, T.; Eyal, N.; Hamblion, E. L.; Lessler, J.; Peters, D. H.; Altmann, M. *PLOS Neglected Tropical Diseases* **2018**, *12*, e0006762, Publisher: Public Library of Science (PLoS).

[5] Lau, J. T.; Tsui, H.; Lau, M.; Yang, X. *Emerging Infectious Diseases* **2004**, *10*, 587–592, Publisher: Centers for Disease Control and Prevention (CDC).

[6] Ferretti, L.; Wymant, C.; Kendall, M.; Zhao, L.; Nurtay, A.; Abeler-Dörner, L.; Parker, M.; Bonsall, D.; Fraser, C. *Science* **2020**, *368*, eabb6936.

[7] He, X. et al. *Nature Medicine* **2020**, *26*, 672–675, Publisher: Springer Science and Business Media LLC.

[8] Li, R.; Pei, S.; Chen, B.; Song, Y.; Zhang, T.; Yang, W.; Shaman, J. *Science* **2020**, *368*, 489–493, Publisher: American Association for the Advancement of Science (AAAS).

[9] Kissler, S. M. et al. *PLOS Biology* **2021**, *19*, e3001333.

[10] Larremore, D. B.; Wilder, B.; Lester, E.; Shehata, S.; Burke, J. M.; Hay, J. A.; Tambe, M.; Mina, M. J.; Parker, R. *Science Advances* **2021**, *7*, Publisher: American Association for the Advancement of Science (AAAS).

[11] Hellewell, J. et al. *The Lancet Global Health* **2020**, *8*, e488–e496, Publisher: Elsevier BV.

[12] Ferretti, L.; Wymant, C.; Petrie, J.; Tsallis, D.; Kendall, M.; Ledda, A.; Di Lauro, F.; Fowler, A.; Di Francia, A.; Panovska-Griffiths, J.; Abeler-Dörner, L.; Charalambides, M.; Briers, M.; Fraser, C. *Nature* **2024**, *626*, 145–150.

[13] Mina, M. J.; Parker, R.; Larremore, D. B. *New England Journal of Medicine* **2020**, *383*, Publisher: Massachusetts Medical Society.

[14] Kojaku, S.; Hébert-Dufresne, L.; Mones, E.; Lehmann, S.; Ahn, Y.-Y. *Nature Physics* **2021**, *17*, 652–658, Publisher: Springer Science and Business Media LLC.

[15] Lavezzo, E. et al. *Nature* **2020**, *584*, 425–429.

[16] Ferguson, N. et al. *Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand*; 2020.

[17] Park, Y. J. et al. *Emerging Infectious Diseases* **2020**, *26*, 2465–2468, Publisher: Centers for Disease Control and Prevention (CDC).

[18] Pung, R.; Cook, A. R.; Chiew, C. J.; Clapham, H. E.; Sun, Y.; Li, Z.; Dickens, B. L.; Ma, S.; Mak, K.; Tan, C. C.; Heng, D.; Chen, M. I.-C.; Lee, V. J. *Epidemiology* **2021**, *32*, 79–86, Publisher: Ovid Technologies (Wolters Kluwer Health).

[19] Grasselli, G.; Pesenti, A.; Cecconi, M. *JAMA* **2020**, *323*, 1545, Publisher: American Medical Association (AMA).

[20] Quilty, B. J. et al. *The Lancet Public Health* **2021**, *6*, e175–e183, Publisher: Elsevier BV.

[21] Middleton, C.; Larremore, D. B. *Science Advances* **2024**, *10*, eadk5108.

[22] Aleta, A.; Martín-Corral, D.; Pastore Y Piontti, A.; Ajelli, M.; Litvinova, M.; Chinazzi, M.; Dean, N. E.; Halloran, M. E.; Longini Jr, I. M.; Merler, S.; Pentland, A.; Vespignani, A.; Moro, E.; Moreno, Y. *Nature Human Behaviour* **2020**, *4*, 964–971, Publisher: Springer Science and Business Media LLC.

[23] Chowell, G.; Hengartner, N.; Castillo-Chavez, C.; Fenimore, P.; Hyman, J. *Journal of Theoretical Biology* **2004**, *229*, 119–126.

[24] Chowell, G.; Nishiura, H. *BMC Medicine* **2014**, *12*, 196.

[25] Riley, S.; Ferguson, N. M. *Proceedings of the National Academy of Sciences* **2006**, *103*, 12637–12642.

[26] Eichner, M. *American Journal of Epidemiology* **2003**, *158*, 110–117.

[27] Lipsitch, M.; Cohen, T.; Cooper, B.; Robins, J. M.; Ma, S.; James, L.; Gopalakrishna, G.; Chew, S. K.; Tan, C. C.; Samore, M. H.; Fisman, D.; Murray, M. *Science* **2003**, *300*, 1966–1970.

[28] Nishiura, H. *Journal of Epidemiology & Community Health* **2006**, *60*, 640–645.

[29] Majumder, M. S.; Cohn, E. L.; Santillana, M.; Brownstein, J. S. *PLoS Currents* **2018**,

[30] Li, Q. et al. *New England Journal of Medicine* **2020**, *382*, 1199–1207, Publisher: Massachusetts Medical Society.

[31] Wu, Z.; McGoogan, J. M. *JAMA* **2020**, *323*, 1239, Publisher: American Medical Association (AMA).

[32] White, L. F.; Wallinga, J.; Finelli, L.; Reed, C.; Riley, S.; Lipsitch, M.; Pagano, M. *Influenza and Other Respiratory Viruses* **2009**, *3*, 267–276.

[33] Fraser, C. et al. *Science* **2009**, *324*, 1557–1561.

[34] Davies, N. G. et al. *Science* **2021**, *372*, eabg3055.

[35] Madewell, Z. J.; Yang, Y.; Longini, I. M.; Halloran, M. E.; Vespignani, A.; Dean, N. E. *BMC Infectious Diseases* **2023**, *23*, 429.

[36] Liu, Y.; Rocklöv, J. *Journal of Travel Medicine* **2021**, *28*, taab124.

[37] Liu, Y.; Rocklöv, J. *Journal of Travel Medicine* **2022**, *29*, taac037.

[38] Sun, K. et al. *Science Translational Medicine* **2022**, *14*, eabo7081.

[39] Pulliam, J. R. C.; Van Schalkwyk, C.; Govender, N.; Von Gottberg, A.; Cohen, C.; Groome, M. J.; Dushoff, J.; Mlisana, K.; Moultrie, H. *Science* **2022**, *376*, Publisher: American Association for the Advancement of Science (AAAS).

[40] Sun, K. et al. *Science* **2021**, *371*, eabe2424.

[41] Lloyd-Smith, J. O.; Schreiber, S. J.; Kopp, P. E.; Getz, W. M. *Nature* **2005**, *438*, 355–359, Publisher: Springer Science and Business Media LLC.

[42] Chen, P. Z.; Bobrovitz, N.; Premji, Z.; Koopmans, M.; Fisman, D. N.; Gu, F. X. *Elife* **2021**, *10*, e65774.

[43] Gootenberg, J. S.; Abudayyeh, O. O.; Lee, J. W.; Essletzbichler, P.; Dy, A. J.; Joung, J.; Verdine, V.; Donghia, N.; Daringer, N. M.; Freije, C. A.; others *Science* **2017**, *356*, 438–442.

[44] Welch, N. L.; Zhu, M.; Hua, C.; Weller, J.; Mirhashemi, M. E.; Nguyen, T. G.; Mantena, S.; Bauer, M. R.; Shaw, B. M.; Ackerman, C. M.; others *Nature medicine* **2022**, *28*, 1083–1094.

[45] Arizti-Sanz, J.; Bradley, A.; Zhang, Y. B.; Boehm, C. K.; Freije, C. A.; Grunberg, M. E.; Kosoko-Thoroddsen, T.-S. F.; Welch, N. L.; Pillai, P. P.; Mantena, S.; others *Nature biomedical engineering* **2022**, *6*, 932–943.

[46] Dao Thi, V. L.; Herbst, K.; Boerner, K.; Meurer, M.; Kremer, L. P.; Kirrmaier, D.; Freistaedter, A.; Papagiannidis, D.; Galmozzi, C.; Stanifer, M. L.; others *Science translational medicine* **2020**, *12*, eabc7075.

[47] Sun, Y.; Yu, L.; Liu, C.; Ye, S.; Chen, W.; Li, D.; Huang, W. *Journal of translational medicine* **2021**, *19*, 74.

[48] Peto, L.; Rodger, G.; Carter, D. P.; Osman, K. L.; Yavuz, M.; Johnson, K.; Raza, M.; Parker, M. D.; Wyles, M. D.; Andersson, M.; others *Journal of clinical microbiology* **2021**, *59*, 10–1128.

[49] Ganguli, A.; Mostafa, A.; Berger, J.; Aydin, M. Y.; Sun, F.; Ramirez, S. A. S. d.; Valera, E.; Cunningham, B. T.; King, W. P.; Bashir, R. *Proceedings of the National Academy of Sciences* **2020**, *117*, 22727–22735.

[50] Hellewell, J.; Russell, T. W.; SAFER Investigators and Field Study Team; Crick COVID-19 Consortium; CMMID COVID-19 working group; Beale, R.; Kelly, G.; Houlihan, C.; Nastouli, E.; Kucharski, A. J. *19*, 106.

[51] Kendall, M.; Ferretti, L.; Wymant, C.; Tsallis, D.; Petrie, J.; Di Francia, A.; Di Lauro, F.; Abeler-Dörner, L.; Manley, H.; Panovska-Griffiths, J.; Ledda, A.; Didelot, X.; Fraser, C. *Science* **2024**, *385*, Publisher: American Association for the Advancement of Science (AAAS).

[52] Fraser, C. *PLoS ONE* **2007**, *2*, e758.

[53] Biggerstaff, M.; Cauchemez, S.; Reed, C.; Gambhir, M.; Finelli, L. *BMC Infectious Diseases* **2014**, *14*, 480.

[54] Zhang, J.; Litvinova, M.; Liang, Y.; Wang, Y.; Wang, W.; Zhao, S.; Wu, Q.; Merler, S.; Viboud, C.; Vespignani, A.; Ajelli, M.; Yu, H. *Science* **2020**, *368*, 1481–1486, Publisher: American Association for the Advancement of Science (AAAS).

[55] Salminen, P. *Advances in Applied Probability* **1988**, *20*, 411–426, Publisher: Cambridge University Press (CUP).

Conflict of interest statement

P.C.S. hold several patents related to genomic and bioinformatic technologies, and is a co-founder and equity holder in Delve Biosciences and Lyra Labs, a board member and equity holder in Polaris Genomics, and an equity holder of NextGenJane. P.C.S was formerly a co-founder of Sherlock Biosciences and board member of Danaher Corporation, until December 2024. The findings and conclusions in this study are those of the authors and do not necessarily represent the official position of the funding agencies, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Funding statement

C.Y.S. was funded by the Harvard Global Health Institute Research & Internships Grant (#192206). S.W.P. was supported by the New Faculty Startup Fund from Seoul National University. Kaiyuan Sun was supported by Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM502).