Delayed introduction, contact variation, and susceptible dynamics explain spatial asynchrony during Korea's large pertussis outbreak

Sang Woo Park[1,*]

1 School of Biological Sciences, Seoul National University, Seoul, Korea

*Corresponding author: sangwoopark@snu.ac.kr

Abstract

Infectious disease outbreaks typically exhibit a large-scale spatial synchrony, reflecting coupling through shared environmental forcing and human mobility. A few exceptions to this theoretical expectation have been reported, but the conditions under which synchrony breaks down remain poorly understood. South Korea's 2024–2025 pertussis outbreak provides a striking example, during which epidemic trajectories diverged substantially throughout the country: notably, the decline in spatial synchrony with distance was more pronounced than that reported in many previous outbreaks. Integrating high-resolution surveillance data from 252 municipalities and a Bayesian transmission model, I show that heterogeneity in introduction timing and differences in local contact levels can drive such fine-scale asynchrony even when all locations share identical seasonal forcing for transmission rates. In particular, regional variation in baseline contact levels can drive differential levels of susceptible depletion, thus shaping the potential for multiple epidemic waves. This analysis offers a unique perspective on spatiotemporal epidemic dynamics that depart from classical epidemic theory.

Introduction

Spatiotemporal variation in epidemic dynamics offers a powerful means for understanding factors that determine the timing and magnitude of epidemic waves across populations [1, 2, 3, 4]. Ecological theory predicts that spatial coupling—via shared environmental forcing and human mobility—can synchronize epidemic dynamics between neighboring locations, where epidemic waves propagate through space [5, 6, 7]. Consistent with this theory, large-scale spatial synchrony has been observed across diverse pathogens, including measles [7], dengue [8], influenza [9], RSV [4], and SARS-CoV-2 [10], highlighting the near-ubiquity of spatial coupling in shaping epidemic dynamics. However, an exception to this pattern has been observed in pre-vaccination pertussis epidemics, which exhibited substantially weaker spatial correlation than measles epidemics reflecting a stronger sensitivity to stochasticity and transients [11].

Pertussis is a vaccine preventable respiratory disease, caused by the bacterium *Bordetella pertussis* [12]. Following the introduction of DTwP (diphtheria, tetanus, and whole cell pertussis) vaccine in 1954 and DTaP (diphtheria, tetanus, and acellular pertussis) vaccine in 1989 in Korea, the number of pertussis cases in Korea decreased dramatically, with less than 10 cases reported in 1995 compared to $> 10,000$ cases reported annually around 1955–1963 [13]. However, pertussis cases have resurged in Korea and many other countries despite high vaccine coverage, raising concerns about the possibility of vaccine waning [14, 15, 16] and pathogen evolution [17, 18, 19].

Between 2024 and 2025, an unprecedented large pertussis outbreak occurred in Korea, resulting in more than 50,000 cases by early 2025 [20]. This outbreak marked a dramatic increase in the disease compared to 292 cases reported in 2023 and 21 cases reported in 2022, despite high vaccine coverage (91.1% of identified 2024 cases had a known vaccination history) [21, 20]. A considerable spatiotemporal variation in disease burden has been observed during the outbreak, but we currently have limited understanding of the underlying spatiotemporal variation in epidemic dynamics [20]. Interestingly, this outbreak exhibited a pronounced spatial asynchrony, which deviates from theoretical expectations based on spatial coupling, raising questions about mechanisms driving this fine-scale spatiotemporal heterogeneity in pertussis epidemic (Figure 1).

To understand mechanisms driving the apparent spatial asynchrony during the large pertussis outbreak in Korea, I combine a detailed spatiotemporal surveillance data from 252 municipalities with mathematical modeling approaches. I begin by showing that asynchrony in pertussis epidemics, captured by heterogeneous epidemic timing, is associated with differences in introduction delays. In contrast, the analysis of effective reproduction number reveals synchrony in underlying transmission patterns, suggesting that spatial asynchrony can arise despite synchrony in transmission. To test this hypothesis, I develop a spatiotemporal transmission model that allows for shared transmission term and variation in initial conditions and fit this model to observed data using Bayesian framework. Overall, the paper offers insight into how

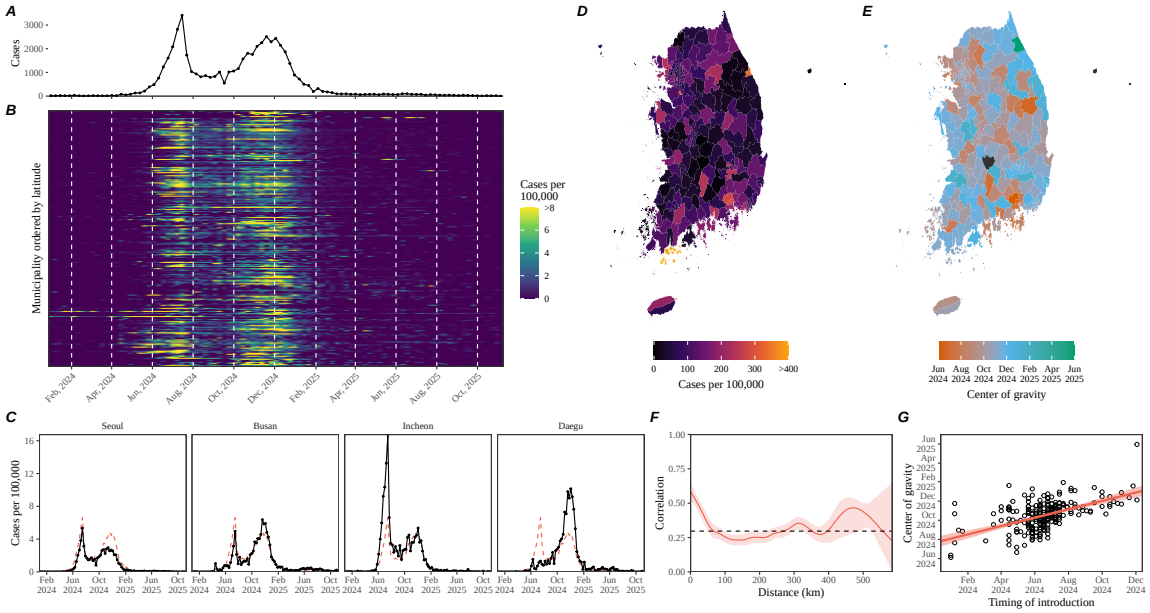

Figure 1: **Spatiotempordal dynamics of pertussis outbreak in Korea, 2024—2025.** (A) Total number of reported cases in Korea. (B) Number of reported cases per 100,000 individuals by municipality, ordered by latitude. (C) Number of reported cases per 100,000 individuals in four most populated cities in Korea. (D) Map of pertussis cases per 100,000 individuals between January 2024 and November 2025. (E) Map of center of gravity, which represents the mean timing of cases. (F) Spatial synchrony in logged cases as a function of distance. The red solid line and shaded regions represent the median estimate and the corresponding 95% confidence interval. The dashed horizontal line represents the mean correlation. (G) Relationship between center of gravity and the timing of introduction, defined as the first week when the number of reported cases is greater than 1 in 100,000. The red solid line and shaded regions represent the linear regression fit and the corresponding 95% confidence interval.

differences in initial conditions (i.e., introduction delays and initial susceptibility) can drive fine-scale spatiotemporal asynchrony in population dynamics.

Materials and Methods

Data

The weekly surveillance data on pertussis cases across 252 municipalities (1st week of 2024 to the 44th week of 2025) were obtained from a publicly available website, the Infectious Disease Portal, by the Korea Disease Control and Prevention Agency [22]. Population size data as of September 2025 were obtained from publicly available website by the Ministry of the Interior and Safety [23]. Vaccine coverage data

were obtained from a publicly available website by the Korea Disease Control and Prevention Agency [24]. Longitude and latitude data for each municipality were obtained from a publicly available GitHub repository [25]. Finally, shape files used for constructing maps were obtained from a publicly available website [26].

Center of gravity

I quantified the center of gravity to characterize variation in the mean timing of epidemic. Typically, the center of gravity is calculated based on annual incidence for recurrent epidemics. Since the paper focuses on analyzing a single outbreak, I used the entire time series of reported cases C_t to computer the center of gravity for each municipality:

$$\text{Center of gravity} = \frac{\sum_{t=1}^{t_{\max}} C_t \times t}{\sum_{t=1}^{t_{\max}} C_t}. \tag{1}$$

I then calculated the correlation coefficient between the center of gravity and the timing of introduction, defined as the first week when the number of reported cases is greater than 1 in 100,000.

Spatial synchrony of reported cases

I characterized the spatial synchrony of pertussis epidemic, which captures changes in pairwise correlation as a function of distance [7]. I used logged values of reported cases between May 2024 and March 2025, excluding municipalities that had no reported cases ($n = 250$). Spatial synchrony was calculated using the `ncf` package in R [27]. I used 1,000 bootstraps to generate the median estimate and the corresponding 95% confidence intervals.

Effective reproduction number

I calculated the effective reproduction number $\mathcal{R}(t)$ to characterize changes in pertussis transmission over time. First, I took the case time series C_t and fitted a generalized additive model assuming a Poisson error to smooth the time series [28]. Then, I estimated $\mathcal{R}(t)$ using the method of [29]:

$$\mathcal{R}(t) = \frac{i(t)}{\sum_{k=1}^{n} i(t-k)g(k)}, \tag{2}$$

where $i(t)$ is the smoothed incidence and n is the maximum length of generation interval in weeks (assumed to be 5 weeks). The generation-interval distribution $g(k)$ was assumed to follow a gamma distribution with a mean of 9.47 days and a standard deviation of 6.22 days based on the estimated serial-interval distribution during this outbreak [30]. The reproduction number estimates can be unrealistically high when the number of infections is close to zero. Therefore, I only used estimates between

May 2024 and March 2025 throughout the paper. I truncated all $\mathcal{R}(t)$ estimates exceeding 5 to a maximum value of 5.

First, I compared correlation coefficients between $\mathcal{R}(t)$ estimates and those between logged cases across all pairwise municipality combinations (Figure 2B in the main text). For this analysis, I only used data from municipalities with more than 400 total cases. Then, I compared the spatial synchrony for $\mathcal{R}(t)$ estimates and synchrony for logged cases (Figure 2C in the main text). For this analysis, I used $\mathcal{R}(t)$ estimates from all municipalities without any exclusion.

Finally, I fitted a generalized additive model to logged values of $\mathcal{R}(t)$ to quantify the effect of susceptible depletion and temporal variation in intrinsic transmission [31, 28, 32]:

$$\log\left(\mathcal{R}_{t,m}\right) = \alpha_m + \beta d_{t,m} + s(t), \tag{3}$$

where $\mathcal{R}_{t,m}$ represents $\mathcal{R}(t)$ estimate at time t in municipality m; α_m represents the municipality-specific intercept term; β represents the effect of susceptible depletion; $d_{t,m}$ represents the cumulative cases at time t in municipality m; and $s(t)$ represents the smooth term capturing temporal variation in transmission. For this analysis, I only used $\mathcal{R}(t)$ estimates from municipalities with more than 400 total cases.

Transmission model

Finally, I fitted a simple transmission model using Bayesian inference to test whether variation in introduction timing and contact variation alone can explain the observed epidemic patterns. Specifically, I extended the SEIR model, which is commonly used for pertussis transmission [33], to allow for a joint estimation of a single, time-varying transmission term that is shared across all regions and a separate estimation of initial infected fraction, basic reproduction number, and reporting rates. To do so, I first discretized the SEIR model following the approach of [34]:

$$\beta(t) = \mathcal{R}_0(1 - \exp(-\gamma\Delta t))\delta(t) \tag{4}$$

$$\mathrm{FOI}(t) = \frac{\beta(t) I_m(t - \Delta t)}{N_m} \tag{5}$$

$$\Delta S_m(t) = \left[1 - \exp(-\mathrm{FOI}(t)\Delta t)\right] S_m(t - \Delta t) \tag{6}$$

$$\Delta E_m(t) = \left[1 - \exp(-\sigma\Delta t)\right] E_m(t - \Delta t) \tag{7}$$

$$\Delta I_m(t) = \left[1 - \exp(-\gamma\Delta t)\right] I_m(t - \Delta t) \tag{8}$$

$$S_m(t) = S(t - \Delta t) - \Delta S_m(t) \tag{9}$$

$$E_m(t) = E(t - \Delta t) - \Delta E_m(t) + \Delta S_m(t) \tag{10}$$

$$I_m(t) = I(t - \Delta t) - \Delta I_m(t) + \Delta E_m(t) \tag{11}$$

where $S_m(t)$, $E_m(t)$, and $I_m(t)$ represent the number of susceptible, exposed, and infectious individuals in municipality m at time t, respectively; N_m represents the population size of municipality m; Δt represents the simulation time step, which assumed to be 1 week; $\beta(t)$ represents the shared time-varying transmission rate; $\mathcal{R}_0$

represents the shared basic reproduction number [35]; $\delta(t)$ represents the normalized time-varying transmission rate; σ represents the rate at which individuals develop infectiousness, which is assumed to be $\sigma = -\log(1 - 7/8)$ such that the mean latent period is 8 days; and γ represents the rate at which individuals recover, which is assumed to be $\sigma = -\log(1 - 7/15)$ such that the mean latent period is 15 days.

Here, I modeled changes in transmission using the $\delta(t)$ term, which is given a normal prior around 1:

$$\delta(t) \sim \text{Normal}(1, 0.2). \tag{12}$$

To allow for smooth variation in transmission, I also incorporated a random walk prior:

$$\delta(t) \sim \text{Normal}(\delta(t - \Delta t), \sigma_\delta), \tag{13}$$

$$\sigma_\delta \sim \text{Half-Normal}(0, 0.1), \tag{14}$$

where a half-normal prior on σ_δ constrains the smoothness of $\delta(t)$. I also estimated a municipality-specific basic reproduction number:

$$\mathcal{R}_{0,m} \sim \text{Normal}(17, 2). \tag{15}$$

The municipality-specific observation process is modeled based on a Poisson distribution:

$$\text{cases}_{t,m} \sim \text{Poisson}(\rho_m \Delta S_m(t)) \tag{16}$$

$$\rho_m \sim \text{Uniform}(0, 1) \tag{17}$$

where $\text{cases}_{t,m}$ represents the reported cases at time t in municipality m; ρ_m represents the municipality-specific reporting rate; and $\Delta S_m(t)$ represents the expected number of new infections between time $t - \Delta t$ and t.

Finally, I estimated a municipality-specific initial infected fraction $i_m(0)$ and a shared initial susceptible fraction $s(0)$ by imposing the following priors:

$$s(0) \sim \text{Uniform}(0, 1), \tag{18}$$

$$i_m(0) \sim \text{Half-Normal}(0, 0.001). \tag{19}$$

Then, the initial conditions for the model is specified as follows:

$$S(0) = N_m s(0), \tag{20}$$

$$E_m(0) = N_m i_m(0), \tag{21}$$

$$I_m(0) = N_m i_m(0). \tag{22}$$

I simultaneously fitted this model time series data between May 2024 and March 2025 from all municipalities with more than 400 total cases (a total of 42 municipalities) using a Bayesian inference software rstan using 4 chains, each with 2000 iterations [36]. Convergence was assess by ensuring low R-hat (< 1.01), high effective sample

size (> 800), no divergent transitions, and no iterations that exceeded the maximum tree depth.

To assess the model fit, I computed R squared values for each municipality. This is done by calculating the squared value of the correlation coefficient between logged values of the observed cases+1 and a posterior median of the logged predictions $\rho_m \Delta S_m(t)$.

Finally, I evaluated the impact of initial conditions on pertussis epidemic dynamics by varying $\mathcal{R}_0$ between 14.2 and 22.2 and $i(0)$ between 5×10^{-6} and 2.3×10^{-3}. For each simulation, I quantified the center of gravity.

As an alternative hypothesis, I also considered a model assuming identical $\mathcal{R}_0$ across municipality and estimated a municipality-specific initial susceptible fraction $s_m(0)$. Same prior distributions were used for this analysis.

Transmission model validation

I validated the model by fitting the model to remaining municipalities (i.e., those with less than 400 total cases). In doing so, I assumed that changes in transmission rate $\delta(t)$ is known and only estimate the initial conditions, $s(0)$ and $i_m(0)$, basic reproduction number $\mathcal{R}_{0,m}$, and the reporting rate ρ_m. For this validation, I used the optimization function in rstan [36] rather than performing a full Bayesian inference. Then, I computed R squared values for each municipality in the same way as before.

Relationship between estimated initial susceptible and vaccine coverage

I assessed the relationship between estimated initial susceptible from the alternative model and vaccine coverage using linear regression. Specifically, assuming that the waning of vaccinal immunity is the main cause for pertussis re-emergence, I compared DTaP (Diphtheria-Tetanus-Pertussis) vaccine coverage for 3 years old as of 2015 and the estimated initial susceptible in each municipality and performed a linear regression. I used vaccine coverage as of 2015 because this was the oldest data that were publicly available. I considered using estimates from municipalities with > 400 cases, since the parameter estimates are expected to be more reliable.

Results

Asynchrony in pertussis epidemics associated with introduction delays

First, I begin by characterizing the spatiotemporal dynamics of the pertussis epidemic in Korea. The outbreak consisted of two waves that peaked around late July and December 2024 (Figure 1A) and exhibited large spatial heterogeneity in timing

(Figure 1B,C) and magnitude (Figure 1D). For example, both Seoul, the capital city located in the northwest region, and Busan, the second most populated city located at the southeast end, experienced two waves, whereas Daegu, the fourth most populated city located $<$ 100km away from Busan, only experienced the second wave (Figure 1C; see Figure S1 for epidemic dynamics in other regions).

To further quantify spatial heterogeneity in pertussis waves across Korea, I calculate the center of gravity (Figure 1E), which represents the mean timing of cases [1]. Notably, I find a sharp contrast between early (orange) and late (light blue) waves between many neighboring municipalities (Figure 1E), suggesting spatial asynchrony in pertussis dynamics. To systematically evaluate this apparent asynchrony, I estimate the spatial synchrony function, which captures changes in correlation in epidemic dynamics as a function of distance [7]. The spatial synchrony function confirms a sharp reduction in cross-correlation with distance (Figure 1F). For example, within 100km of distance, the cross-correlation drops below 0.25. By comparison, pertussis epidemics in the US maintained cross-correlations above 0.2 even at 1,000 km [37], highlighting the fine-scale asynchrony in pertussis epidemics in Korea. Interestingly, the variation in the center of gravity significantly correlates with the introduction timing ($\rho = 0.53$; 95%CI: 0.44—0.62; Figure 1G), suggesting that differences in introduction timing may contribute to the apparent spatial asynchrony.

Synchrony in temporal changes in pertussis transmission

I then ask whether the observed epidemic asynchrony can be explained by the differences in transmission patterns. To do so, I quantify the effective reproduction number $\mathcal{R}(t)$, which measures the average number of new infections caused by a single infected individual [38, 29]. For this analysis, I focus on municipalities with more than 400 total cases reported because $\mathcal{R}(t)$ estimates can be unstable when the number of infections are low.

Notably, $\mathcal{R}(t)$ estimates appear to be largely homogeneous across different regions, despite the apparent heterogeneity in the observed epidemic dynamics (Figure 2A). Comparisons of correlation coefficients in $\mathcal{R}(t)$ versus correlation coefficients in reported cases between all pairwise municipality combinations confirm this observation (Figure 2B). In 830 out of 861 municipality combinations, correlation in $\mathcal{R}(t)$ is higher than correlation in cases. This is reflected in the spatial synchrony function for $\mathcal{R}(t)$ (Figure 2C), where the mean cross-correlation for $\mathcal{R}(t)$ ($\rho = 0.71$) is considerably higher than that for cases ($\rho = 0.30$).

To better understand factors driving changes in $\mathcal{R}(t)$, I perform a regression analysis using a generalized additive model with three covariates [31, 28, 32]: (1) a municipality-level fixed effect allowing each municipality to have its own intercept; (2) a cumulative-cases term capturing the effect of susceptible depletion; and (3) a single smooth term shared across municipalities, which allows for temporal changes in pertussis transmission over time, such as seasonal forcing. First, the smooth term suggests a strong transmission reduction around August and February,

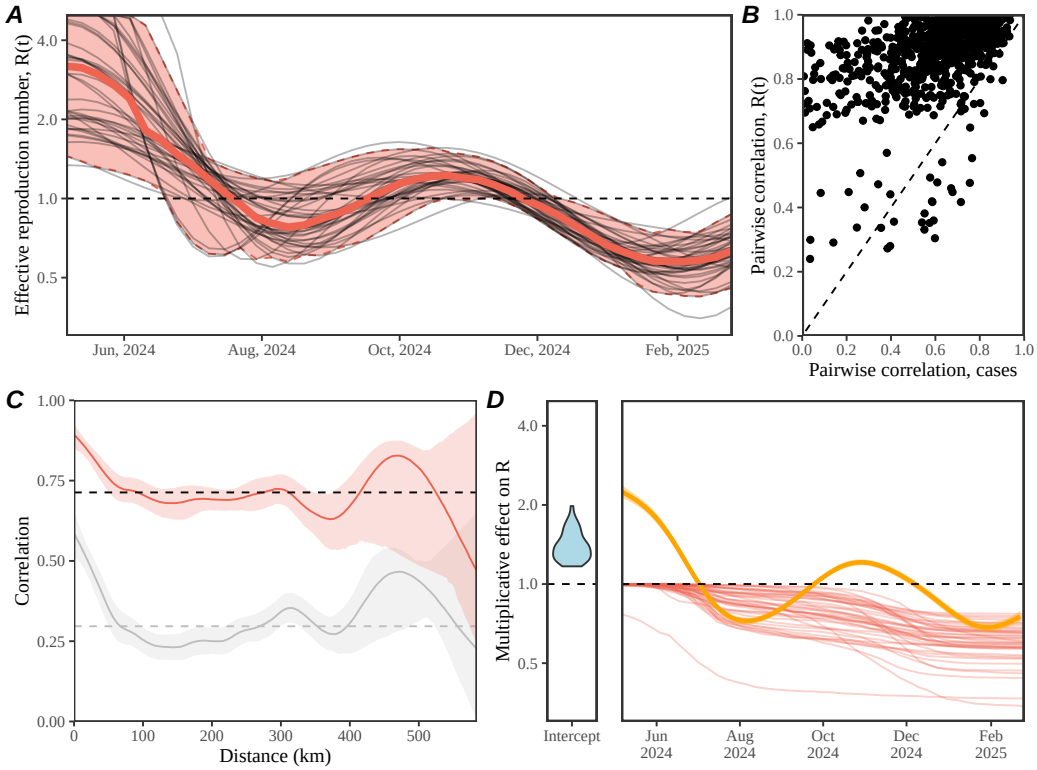

Figure 2: **Spatial variation in the estimates of effective reproduction number.** (A) Each black line represents the estimated effective reproduction number for a given municipality with more than 400 total cases. The red solid line and shaded regions represent the median and 95% quantiles of effective reproduction number estimates. (B) Relationship between correlation coefficient for $\mathcal{R}(t)$ and correlation coefficient for reported cases across all pairwise municipality combinations with more than 400 total cases. (C) Spatial synchrony in $\mathcal{R}(t)$ (red) versus logged cases (black) as a function of distance. The solid line and shaded regions represent the median estimate and the corresponding 95% confidence interval. The dashed horizontal line represents the mean correlation. (D) Estimated multiplicative effects on $\mathcal{R}(t)$ from the regression analysis. The blue shaded violin represents the distribution of intercept effect, where a separate intercept term is estimated for each municipality. The red lines represent the estimated susceptible depletion effect for each municipality. The orange line and shaded regions represent the estimated changes in transmission from a smooth term and the corresponding 95% confidence interval.

which coincide with the timing of school breaks (Figure 2D). Second, the model estimates a significant effect of effect of susceptible depletion (Figure 2D; $p < 0.001$). Notably, this simple regression model explains 84% of variance in $\mathcal{R}(t)$, suggesting that the observed temporal fluctuations in transmission can be largely explained by school-related seasonality and the gradual depletion of susceptible individuals. Taken together, the analysis of center of gravity and $\mathcal{R}(t)$ suggest that differences in

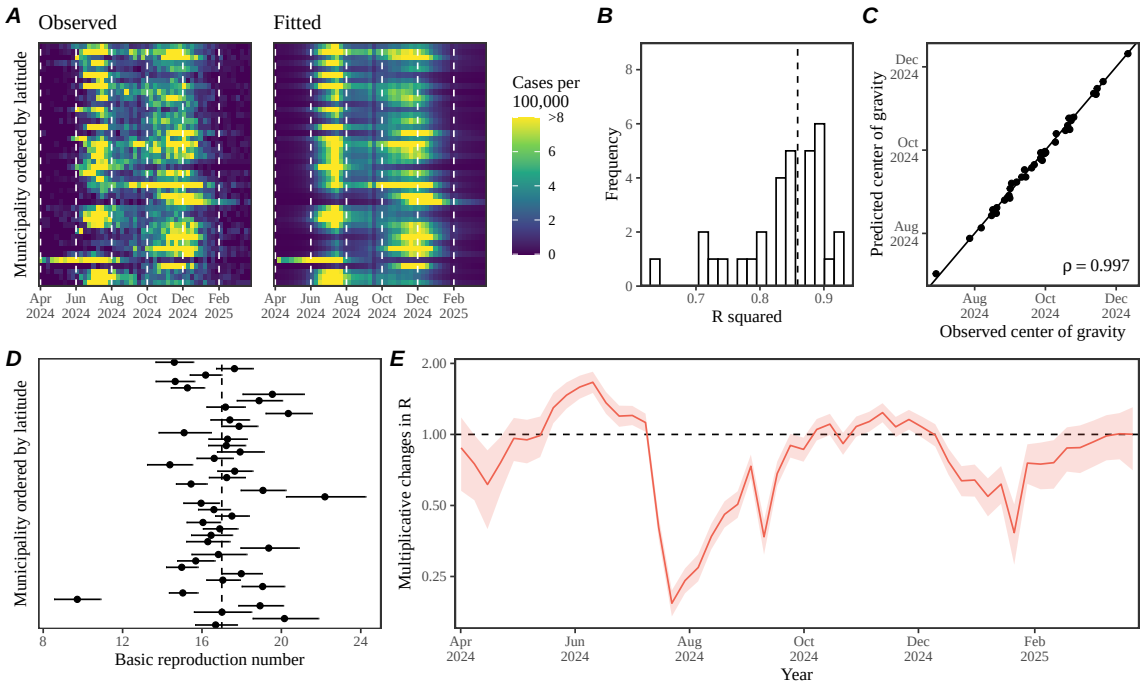

Figure 3: **Fine-scale spatiotemporal variation in pertussis epidemic can be explained with a simple SEIR model.** (A) Comparisons between the observed and predicted epidemic dynamics across municipalities with more than 400 total cases, ordered by latitude. (B) The bar plot represents the distribution of R squared values for model fits. The vertical dashed line represents the median. (C) Relationship between the estimated and predicted center of gravity. The solid line indicated the one-to-one relationship. (D) Estimated basic reproduction number $\mathcal{R}_0$ across municipalities, ordered by latitude. Points and error bars represent posterior medians and corresponding 95% credible intervals. (E) Estimated time-varying pertussis transmission multiplier. The red line and shaded regions represent the posterior median and the corresponding 95% credible interval.

introduction timing and susceptible host dynamics are likely main drivers for spatial asynchrony in pertussis epidemics in Korea, rather than variation in seasonal transmission patterns.

Mathematical modeling of fine-scale spatio-temporal pertussis epidemic

To test whether variation in introduction timing and susceptible host dynamics can explain the observed epidemic patterns, I extend a standard susceptible-exposed-infected-removed (SEIR) model [33] to jointly estimate a single, time-varying transmission term, which is shared across all regions. In particular, the transmission rate

$\beta_r(t)$ in region r is modeled as follows:

$$\beta_r(t) = \beta_r(0)\delta(t), \tag{23}$$

where $\beta_r(0)$ represents the baseline level transmission rates, which accounts for variation in baseline contact levels and is estimated separately for each region r. Then, $\delta(t)$ represents a time-varying transmission term, which are shared across all regions and captures a multiplicative change in transmission over time. I also allow the initial susceptible fraction $S(0)$ to be shared across all regions to capture the average level of susceptibility. Initial infected fraction $I_r(0)$ and reporting rate ρ_r are allowed to vary across regions, where variation in $I(0)$ effectively captures differences in introduction timing. For this analysis, I also focus on municipalities with more than 400 total cases. The model is fitted using a Bayesian inference software rstan [36].

This simple transmission model that assume shared transmission rate can capture fine-scale spatiotemporal variation of pertussis outbreak in Korea (Figure 3A) with a median R squared value of 0.86 (95% quantile: 0.71—0.92; Figure 3B). Furthermore, I find a near-perfect correlation between the predicted and observed center of gravity (Figure 3C). These results suggest that the model assuming shared time-varying transmission can sufficiently explain the fine-scale spatial asynchrony observed during the pertussis outbreak in Korea. So how do contact levels vary across regions and transmission change across time?

I estimate a considerable variation in $\mathcal{R}_0$ across regions (Figure 3D). The median estimates for region-level $\mathcal{R}_0$ vary from 9.7 to 22.1 with an average of 17, consistent with previous estimates [39, 16, 40]. The estimated initial susceptible fraction is 0.15 (95% CI: 0.14–0.16), which correspond to an initial effective reproduction number of 2.6 when combined with the average $\mathcal{R}_0$ estimate—this is consistent with the initial $\mathcal{R}(t)$ estimate. Finally, I also estimate a large reduction in transmission around late July 2024 and late January 2025 (Figure 3E). The estimated timing of transmission reduction is also consistent with results from the $\mathcal{R}(t)$ analysis and coincides with the timing of school holidays in Korea. Parameter estimates for the initial infected fraction $I(0)$ and reporting rates also reveal large variation across regions (Supplementary Figure S2). While estimates of $\mathcal{R}_0$ or $I(0)$ do not correlate with population sizes, reporting rate negatively correlates with population size (Supplementary Figure S3).

I validate the model by taking the estimated transmission term $\delta(t)$ and fitting the remaining parameters (i.e., initial infected, initial susceptible, basic reproduction number, and reporting probability) to data from the remaining 210 municipalities with less than 400 total cases (Figure S4). While model predictions appear consistent with the observed patterns (Figure S4A), I estimate a lower value for the median R squared: 0.59 (95% quantile: 0.01–0.85; Figure S4B). Nonetheless, this analysis suggests that the majority of asynchrony in pertussis epidemics can be explained via the differences in introduction timing and contact variation.

As an alternative hypothesis, I also considered the possibility that the spatiotemporal asynchrony may be driven by variation in susceptibility. To do so, I assumed a

fixed basic reproduction number, $\mathcal{R}_0 = 17$ and estimated a separate initial susceptible fraction $S_r(0)$ for each region and a shared time-varying transmission term $\delta(t)$. This alternative model provided nearly identical fits to data compared to the original model (Supplementary Figure S5–S6) and similar levels of out-of-sample validation (Supplementary Figure S7). However, I did not find any correlation between past vaccine coverage and the estimated levels of susceptibility (Supplementary Figure S7). Moreover, the estimated range for the initial susceptible fraction was unrealistically wide compared to the reported range for past vaccine coverage (Supplementary Figure S7). Thus, I tentatively conclude that a model assuming contact variation is likely more parsimonious and continue the analysis.

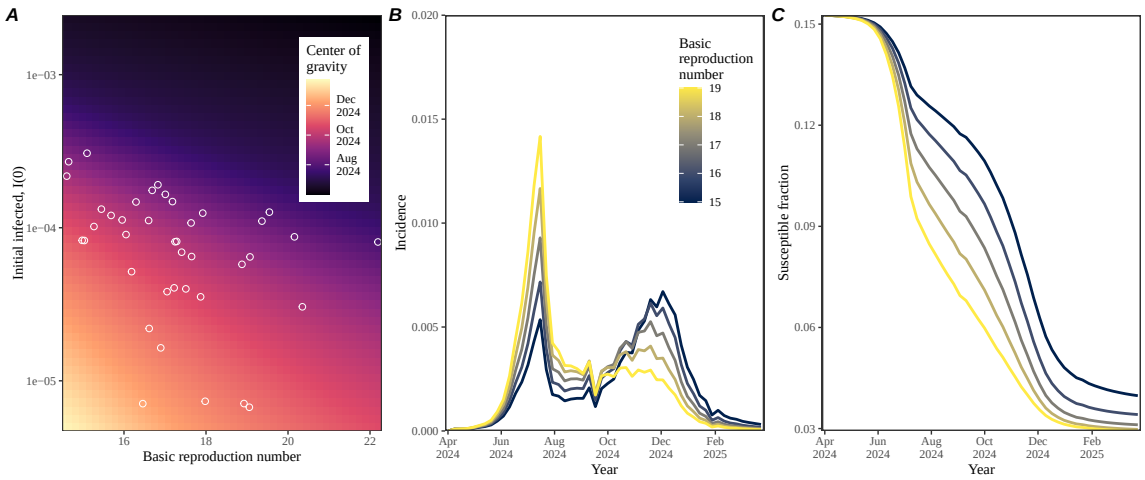

Figure 4: **Impact of delayed introduction, contact variation, and susceptible host dynamics on epidemic waves.** (A) Relationship between the basic reproduction number $\mathcal{R}_0$ and infected I(0) fraction with center of gravity. Points represent the estimates for each municipality. (B) Impact of basic reproduction number on pertussis epidemic waves. (C) Impact of basic reproduction number on pertussis susceptible host dynamics. Colors in panels B and C represent values of basic reproduction number.

So how does variation in introduction timing and contact rates generate heterogeneity in epidemic waves? I address this question by simulating the model between April 2024 and April 2025 across plausible ranges of $\mathcal{R}_0$ and $I(0)$. First, increasing $I(0)$ leads to an earlier epidemic because it permits earlier introduction of the infection to the population (Figure 4A). This is consistent with the earlier observations (Figure 1F). Similarly, increasing $\mathcal{R}_0$ also results in an earlier epidemic. This is because high levels of transmission allows for a large first wave (Figure 4B), which prevents the second wave through the overcompensatory susceptible depletion (Figure 4C). In contrast, when $\mathcal{R}_0$ is low, there will be enough susceptible hosts for the population to experience the second wave (Figure 4C). These results are consistent with an earlier finding, which showed that contact variation can explain a sudden

shift in RSV seasonality from a later wave to an earlier wave [41].

Discussion

Understanding mechanisms that drive spatiotemporal variation in epidemic dynamics is a major aim for infectious disease dynamics and control [1, 2, 3, 4, 7, 8, 9]. The large pertussis outbreaks in Korea that occurred between 2024 and 2025 presents a unique perspective on a fine-scale spatiotemporal asynchrony heterogeneity in epidemic dynamics: some municipalities experienced two epidemic waves while other places experienced only the first or second without clear patterns of synchrony. I propose that this variation can be explained by the differences in introduction delays, contact variation, and susceptible host dynamics, where high levels of transmission can cause a large first wave and prevent the second wave through susceptible depletion. Incorporating this idea into a simple transmission model by allowing for a joint estimation of a shared transmission term confirms this hypothesis. Together, these findings shed light into mechanisms that can generate spatial asycnrhony in epidemic dynamics, offering a general framework for understanding the spatiotemporal dynamics of re-emerging infections.

Even though the model presented in this paper explains the observed dynamics reasonably well, there are several limitations to this work. First, the model is purely deterministic and does not account for stochasticity, which is likely an important factor for characterizing the establishment of infections after the initial invasion and the subsequent dynamics, especially in small populations [42, 43, 44]. Second, I estimated that variation in contact rates can explain the observed epidemic patterns, but I was unable to identify the mechanism that explains this variation. As an alternative hypothesis, I also considered the possibility that spatial asycnrhony can be explained by susceptible variation given the role of waning immunity in driving pertussis re-emergence [16, 45, 46]. However, I still did not find a clear correlation between past vaccine coverage and estimated variation in susceptible fraction. Detailed serological data may allow us to test different model formulations and further quantify potential variation in population-level susceptibility [47, 48, 49].

Resulting parameter estimates must be interpreted with care as the model relies on several simplifying assumptions. For example, the main model assumed the same values of initial susceptible fraction across all municipalities and tried to estimate separate basic reproduction number $\mathcal{R}_0$. Therefore, any regional variation in $S(0)$, driven by differences in vaccine coverage and waning, would be captured in the estimate of the $\mathcal{R}_0$. I also did not consider an age structure in the model, but age-structured contact patterns play an important role in shaping pertussis transmission [33]. For example, [50] recently suggested transmission among elderly may contributed to the persistence of pertussis in Korea. Differences in age structure across regions may also contribute to variation in contact rates and population-level immunity [51], but detailed age structured data would be needed to address these

challenges. Despite these limitations, this work provides a quantitative foundation and empirical evidence for understanding epidemic asynchrony.

The continued emergence of new pathogens and re-emergence of vaccine preventable infections, such as pertussis, highlights the importance of identifying hotspots of future outbreaks and developing targeted intervention strategies [52, 53, 48]. The apparent asynchrony in pertussis epidemic in Korea, driven by the differences in introduction timing and contact rates, underlines the importance of detailed pathogen surveillance for early detection and strategic allocation of prevention and control resources. Immunological surveillance platforms will be particularly useful for evaluating potential risk of pathogen re-emergence in the future [47, 48, 49]. A detailed understanding of spatial variation in contact patterns and population-level immunity will be crucial for predicting spatiotemporal dynamics of future outbreaks.

Acknowledgements

S.W.P. thanks Bryan Grenfell for helpful discussion. S.W.P. was supported by the New Faculty Startup Fund from Seoul National University.

Author contributions

S.W.P. conceived of the study, conducted the analysis, and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Data availability

All data and code used in this paper are available on https://github.com/parklab-snu/asynchrony-pertussis.

Supplementary Figures

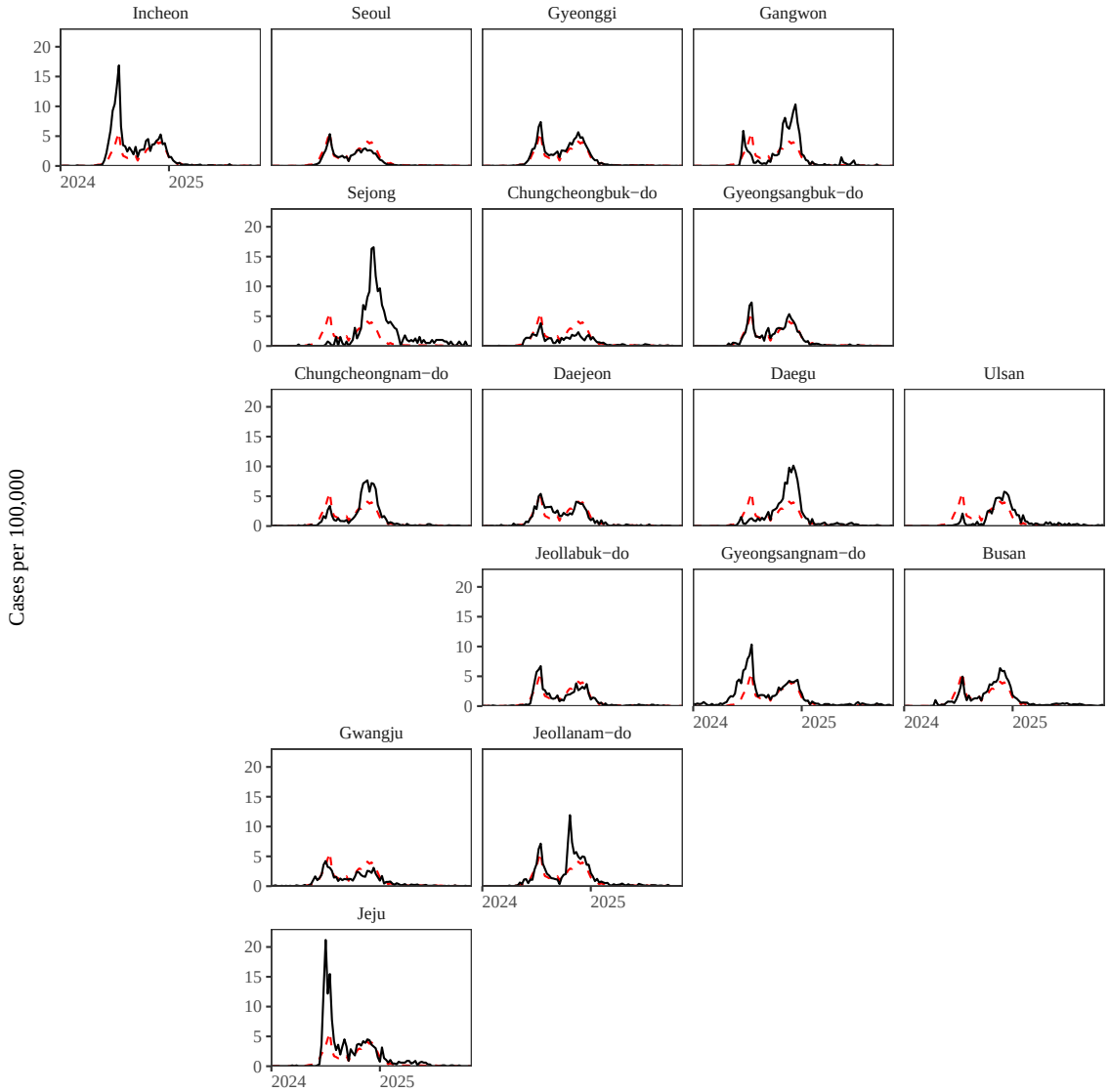

Figure S1: **Spatiotempordal dynamics of pertussis outbreak in Korea across first-level administrative divisions, 2024—2025.** Time series are arranged based on their approximate locations.

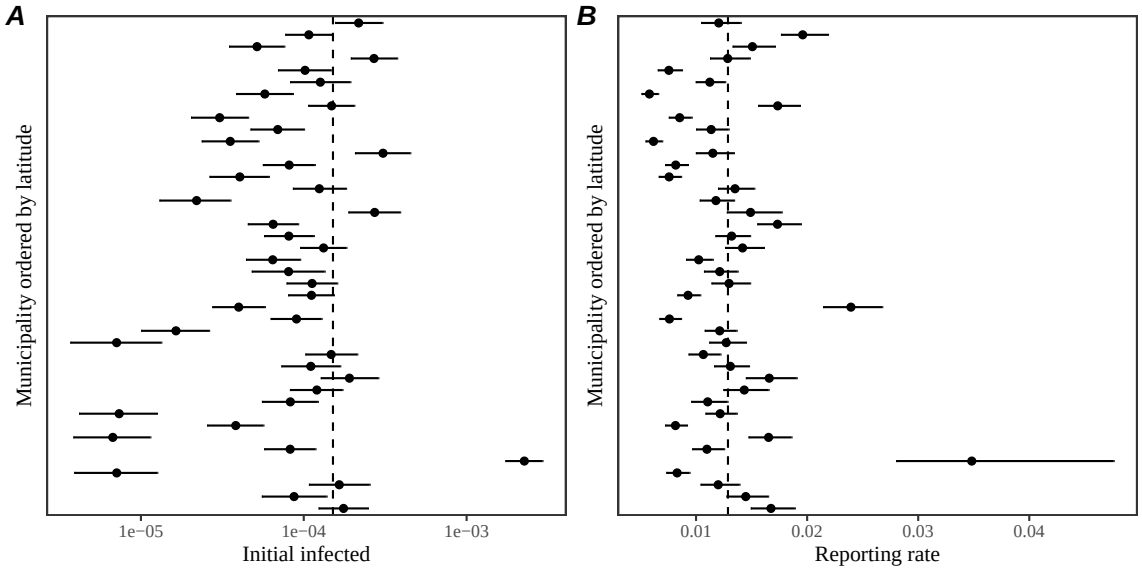

Figure S2: **Estimated initial infected fraction (A) and reporting rates (B) using time series data from municipalities with less than 400 total cases.** Points represent posterior medians. Error bars represent 95% credible intervals.

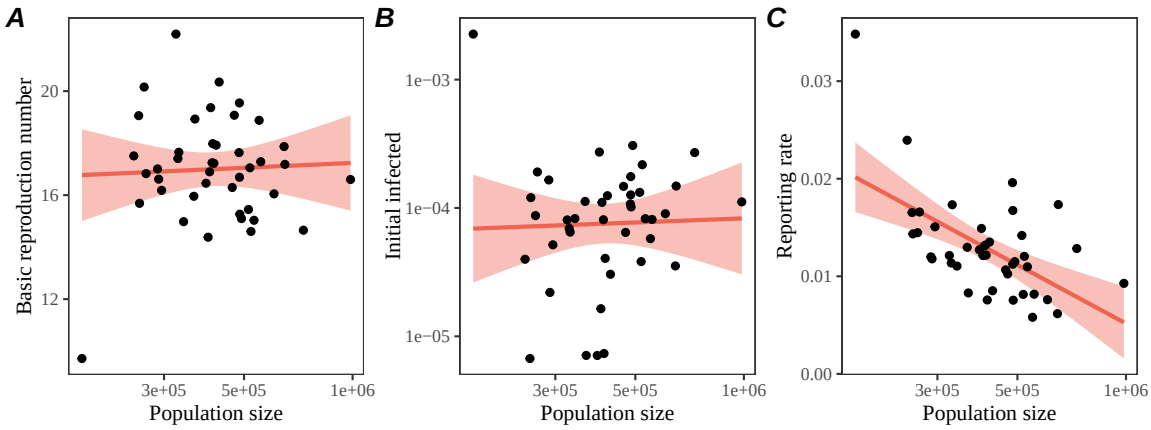

Figure S3: **Relationship between population size and parameter estimates for the basic reproduction number (A), initial infected fraction (B), and reporting rate (C).** Points represent posterior medians. Error bars represent 95% credible intervals. Red lines and shaded regions represent the linear regression fit and associated 95% confidence intervals.

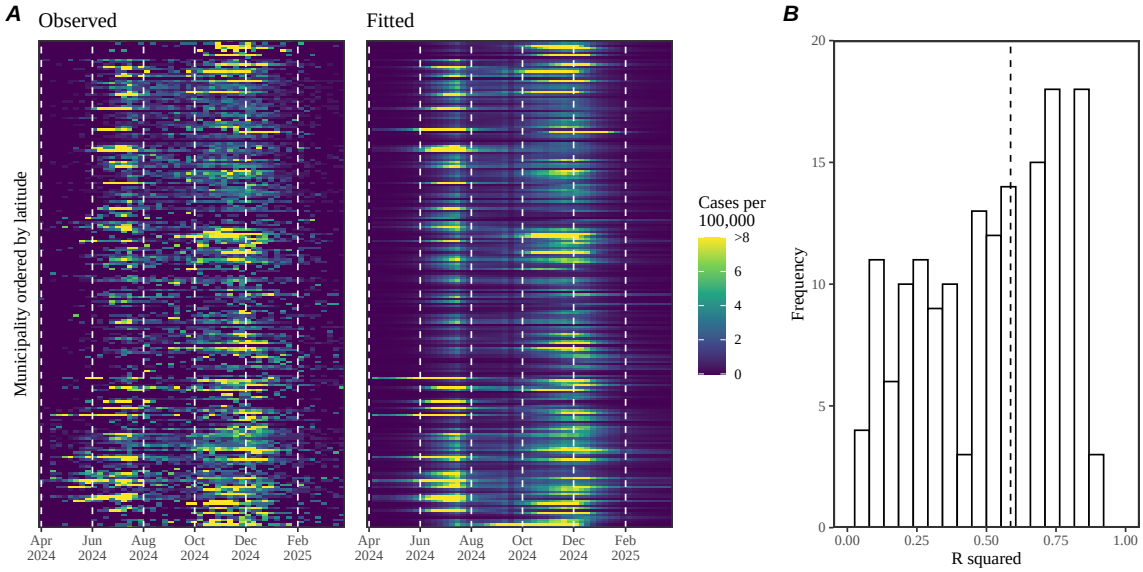

Figure S4: **Validation of the transmission model using time series data from municipalities with less than 400 total cases.** (A) Comparisons between the observed and predicted epidemic dynamics across municipalities with less than 400 total cases, ordered by latitude. (B) The bar plot represents the distribution of R squared values for model fits. The vertical dashed line represents the median. Initial infected, initial susceptible, basic reproduction number, and reporting probability were estimated for each region while assuming a known transmission variation term $\delta(t)$.

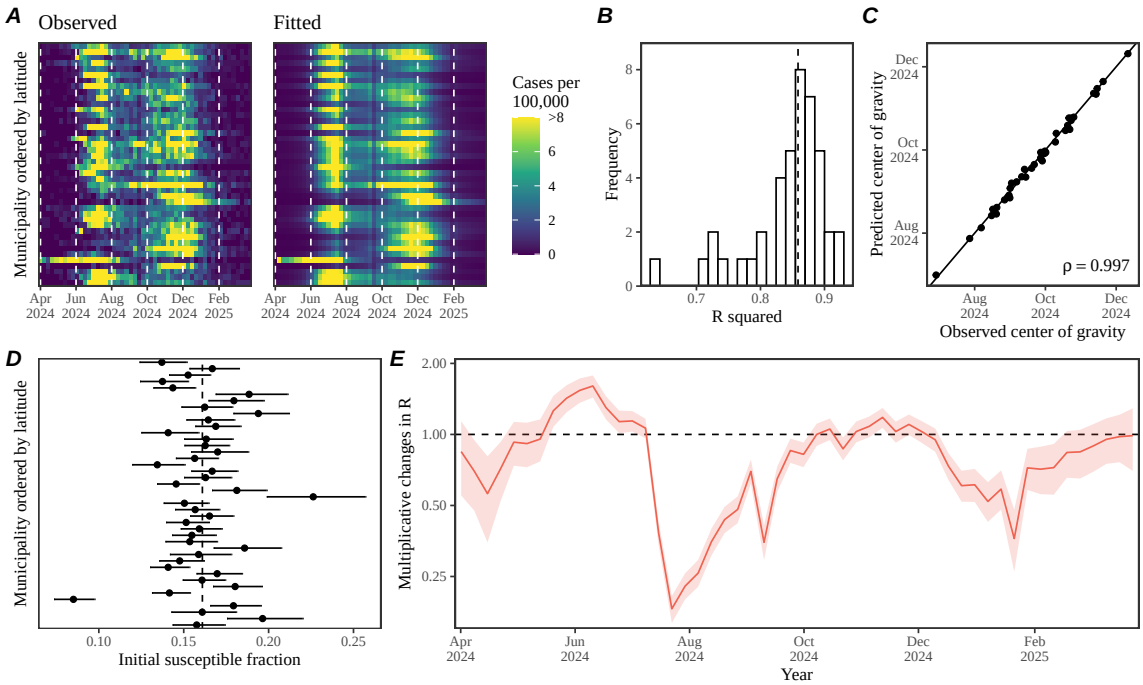

Figure S5: **Alternative SEIR model fits assuming variation in initial susceptible fraction.** (A) Comparisons between the observed and predicted epidemic dynamics across municipalities with more than 400 total cases, ordered by latitude. (B) The bar plot represents the distribution of R squared values for model fits. The vertical dashed line represents the median. (C) Relationship between the estimated and predicted center of gravity. The solid line indicated the one-to-one relationship. (D) Estimated initial susceptible fraction $S(0)$ across municipalities, ordered by latitude. Points and error bars represent posterior medians and corresponding 95% credible intervals. (E) Estimated time-varying pertussis transmission multiplier. The red line and shaded regions represent the posterior median and the corresponding 95% credible interval.

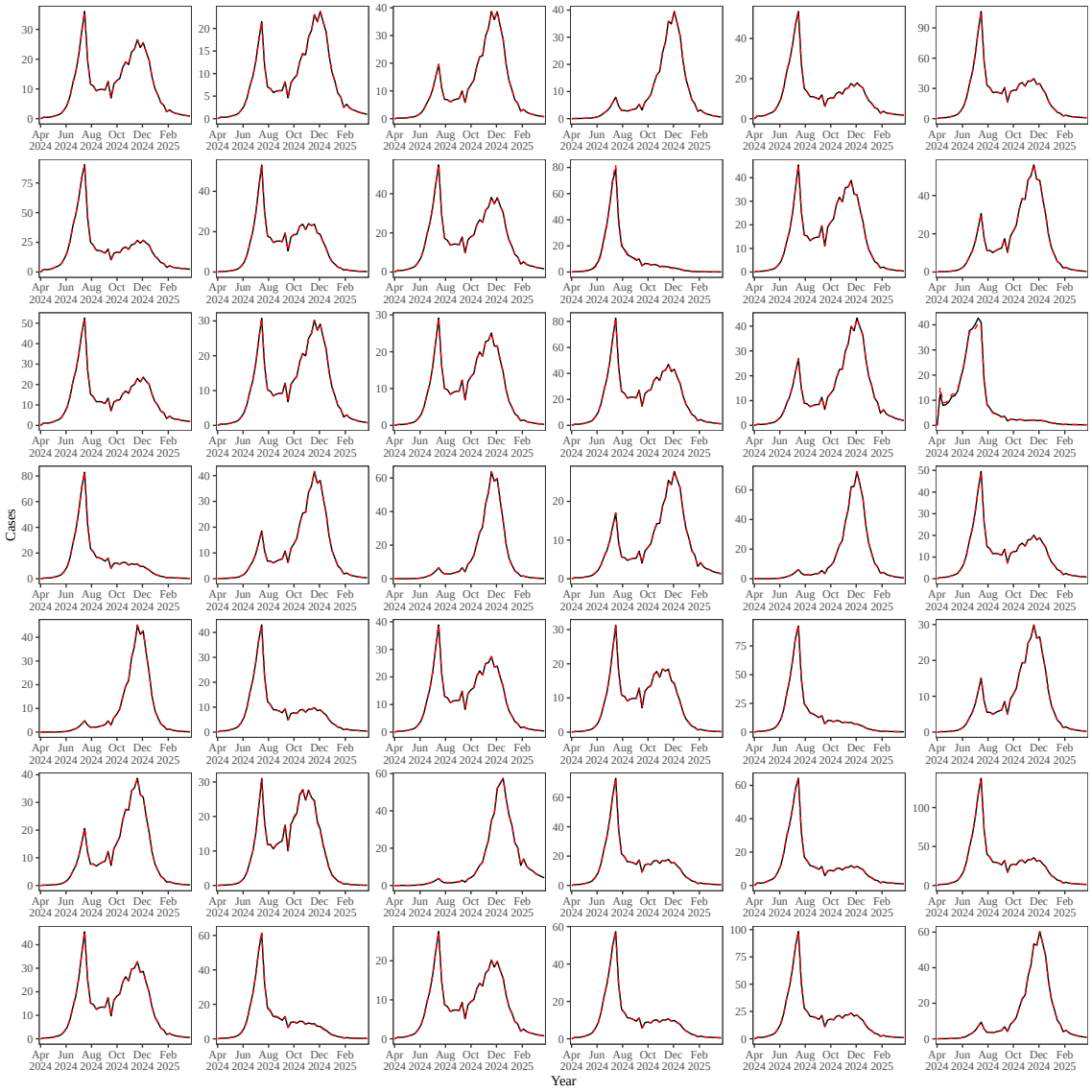

Figure S6: **Comparisons between model fits across two different models.** Solid lines represent posterior median predictions based on the model assuming separate $\mathcal{R}_0$ in each region. Red lines represent posterior median predictions based on the model assuming separate $S(0)$ in each region.

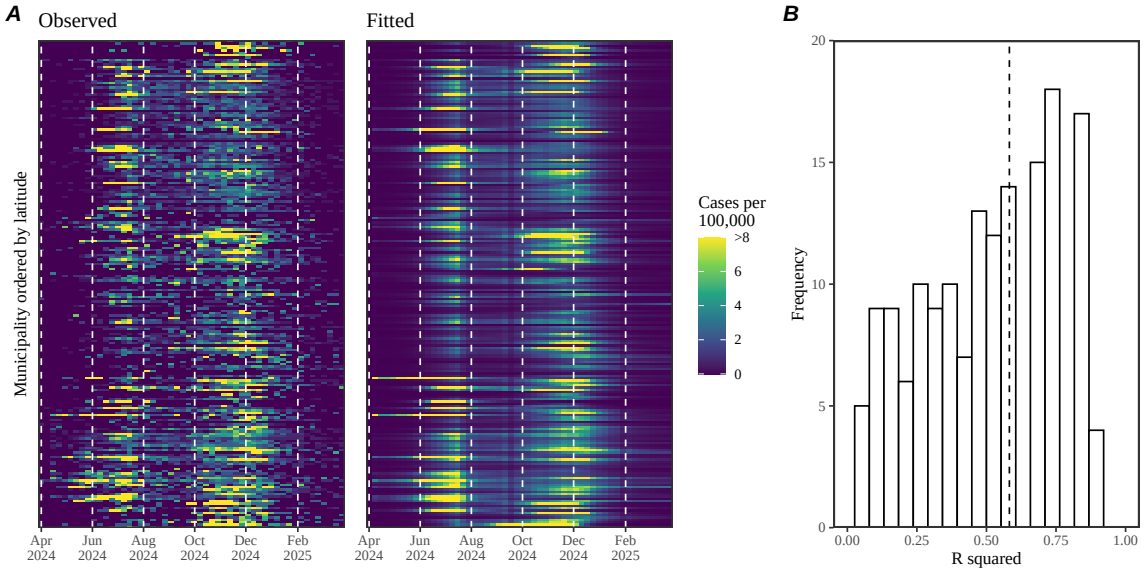

Figure S7: **Validation of the alternative transmission model using time series data from municipalities with less than 400 total cases.** (A) Comparisons between the observed and predicted epidemic dynamics across municipalities with less than 400 total cases, ordered by latitude. (B) The bar plot represents the distribution of R squared values for model fits. The vertical dashed line represents the median. Initial infected, initial susceptible, and reporting probability were estimated for each region while assuming a known transmission variation term $\delta(t)$ and basic reproduction number.

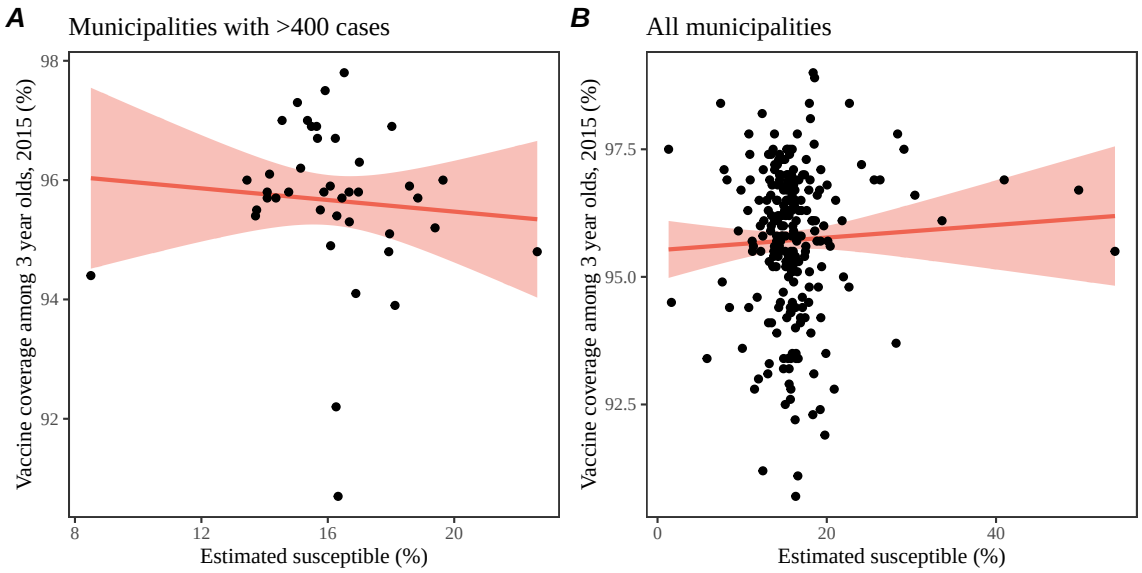

Figure S8: **Relationship between estimated initial susceptible and vaccine coverage for (A) municipalities with $>$ 400 cases and (B) for all municipalities.** Each point represents the estimate initial susceptible and vaccine coverage value for each municipality. Red lines and shaded regions represent the linear regression fit and corresponding 95% confidence intervals.

References

[1] Virginia E Pitzer, Cécile Viboud, Wladimir J Alonso, Tanya Wilcox, C Jessica Metcalf, Claudia A Steiner, Amber K Haynes, and Bryan T Grenfell. Environmental drivers of the spatiotemporal dynamics of respiratory syncytial virus in the United States. *PLoS pathogens*, 11(1):e1004591, 2015.

[2] Benjamin D Dalziel, Stephen Kissler, Julia R Gog, Cecile Viboud, Ottar N Bjørnstad, C Jessica E Metcalf, and Bryan T Grenfell. Urbanization and humidity shape the intensity of influenza epidemics in US cities. *Science*, 362(6410):75–79, 2018.

[3] Margarita Pons-Salort, M Steven Oberste, Mark A Pallansch, Glen R Abedi, Saki Takahashi, Bryan T Grenfell, and Nicholas C Grassly. The seasonality of nonpolio enteroviruses in the united states: Patterns and drivers. *Proceedings of the National Academy of Sciences*, 115(12):3078–3083, 2018.

[4] Rachel E Baker, Ayesha S Mahmud, Caroline E Wagner, Wenchang Yang, Virginia E Pitzer, Cecile Viboud, Gabriel A Vecchi, C Jessica E Metcalf, and Bryan T Grenfell. Epidemic dynamics of respiratory syncytial virus in current and future climates. *Nature communications*, 10(1):5512, 2019.

[5] Michael P Hassell, Hugh N Comins, and Robert M Mayt. Spatial structure and chaos in insect population dynamics. *Nature*, 353(6341):255–258, 1991.

[6] JA Sherratt. Periodic travelling waves in cyclic predator–prey systems. *Ecology Letters*, 4(1):30–37, 2001.

[7] Bryan T Grenfell, Ottar N Bjørnstad, and Jens Kappey. Travelling waves and spatial hierarchies in measles epidemics. *Nature*, 414(6865):716–723, 2001.

[8] Derek AT Cummings, Rafael A Irizarry, Norden E Huang, Timothy P Endy, Ananda Nisalak, Kumnuan Ungchusak, and Donald S Burke. Travelling waves in the occurrence of dengue haemorrhagic fever in thailand. *Nature*, 427(6972):344–347, 2004.

[9] Cécile Viboud, Ottar N Bjørnstad, David L Smith, Lone Simonsen, Mark A Miller, and Bryan T Grenfell. Synchrony, waves, and spatial hierarchies in the spread of influenza. *science*, 312(5772):447–451, 2006.

[10] Robert S Paton, Christopher E Overton, and Thomas Ward. The rapid replacement of the SARS-CoV-2 Delta variant by Omicron (B. 1.1. 529) in England. *Science Translational Medicine*, 14(652):eabo5395, 2022.

[11] Pejman Rohani, David JD Earn, and Bryan T Grenfell. Opposite patterns of synchrony in sympatric disease metapopulations. *Science*, 286(5441):968–971, 1999.

[12] Paul E Kilgore, Abdulbaset M Salim, Marcus J Zervos, and Heinz-Josef Schmitt. Pertussis: microbiology, disease, treatment, and prevention. *Clinical microbiology reviews*, 29(3):449–486, 2016.

[13] Young June Choe, Young-Joon Park, Chaewon Jung, Geun-Ryang Bae, and Duk-hyoung Lee. National pertussis surveillance in South Korea 1955–2011: epidemiological and clinical trends. *International Journal of Infectious Diseases*, 16(12):e850–e854, 2012.

[14] James D Cherry. The epidemiology of pertussis: a comparison of the epidemiology of the disease pertussis with the epidemiology of bordetella pertussis infection. *Pediatrics*, 115(5):1422–1427, 2005.

[15] Aaron M Wendelboe, Annelies Van Rie, Stefania Salmaso, and Janet A Englund. Duration of immunity against pertussis after natural infection or vaccination. *The Pediatric infectious disease journal*, 24(5):S58–S61, 2005.

[16] Helen J Wearing and Pejman Rohani. Estimating the duration of pertussis immunity using epidemiological signatures. *PLoS pathogens*, 5(10):e1000647, 2009.

[17] Frits R Mooi, IH Van Loo, and Audrey J King. Adaptation of bordetella pertussis to vaccination: a cause for its reemergence? *Emerging infectious diseases*, 7(3 Suppl):526, 2001.

[18] Alex-Mikael Barkoff, Jussi Mertsola, Denis Pierard, Tine Dalby, Silje Vermedal Hoegh, Sophie Guillot, Paola Stefanelli, Marjolein Van Gent, Guy Berbers, Didrik F Vestrheim, et al. Surveillance of circulating Bordetella pertussis strains in Europe during 1998 to 2015. *Journal of clinical microbiology*, 56(5):10–1128, 2018.

[19] Elodie Lesne, Breeze E Cavell, Irene Freire-Martin, Ruby Persaud, Frances Alexander, Stephen Taylor, Mary Matheson, Cécile ACM van Els, and Andrew Gorringe. Acellular pertussis vaccines induce anti-pertactin bactericidal antibodies which drives the emergence of pertactin-negative strains. *Frontiers in Microbiology*, 11:2108, 2020.

[20] Hye-Jin Kim, Young Joon Park, Dongkeun Kim, Jihee Lee, Yun Kyoung Kim, Soonryu Seo, Jin Seon Yang, Yeoeun Yun, Eunbyeol Wang, Subin Park, Seo Yeon Ko, Jin Lee, Jeeyeon Shin, Wookeon Lee, and Seonhee Ahn. Pertussis surveillance and occurrence report in the Republic of Korea: from 2024 to the first half of 2025. *Public Health Weekly Report*, 18(43):1631–1651, 11 2025.

[21] Hyun Mi Kang, Taek-Jin Lee, Su Eun Park, and Soo-Han Choi. Pertussis in the post-COVID-19 era: resurgence, diagnosis, and management. *Infection & Chemotherapy*, 57(1):13, 2024.

[22] Korea Disease Control and Prevention Agency. Notifiable Infectious Diseases. 2025. https://dportal.kdca.go.kr/pot/is/rginEDW.do. Accessed Nov 2, 2025.

[23] Ministry of the Interior and Safety. Resident Population and Household Statistics by Administrative District. 2025. https://jumin.mois.go.kr/. Accessed Nov 2, 2025.

[24] Korea Disease Control and Prevention Agency. 2015 National Immunization Coverage Survey. 2025. https://nip.kdca.go.kr/irhp/infm/goNatnVcntStatView.do. Accessed Nov 2, 2025.

[25] cubensys. Korea_district. 2018. https://github.com/cubensys/Korea_District. Accessed Nov 2, 2025.

[26] Hyung-jun Kim. Download the latest administrative divisions (SHP) of the Republic of Korea, GIS Developer. 2023. http://www.gisdeveloper.co.kr/?p=2332. Accessed Nov 2, 2025.

[27] Ottar N. Bjornstad. *ncf: Spatial Covariance Functions*, 2022. R package version 1.3-2.

[28] Simon N Wood. *Generalized additive models: an introduction with R*. Chapman and Hall/CRC, 2017.

[29] Anne Cori, Neil M Ferguson, Christophe Fraser, and Simon Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American journal of epidemiology*, 178(9):1505–1512, 2013.

[30] U Jin Cho, Seonghui Cho, Hyeokjin Lee, Seul-Ki Kang, Bryan Inho Kim, Yukyung Nam, Chiara Achangwa, Jun-Sik Lim, Dong Han Lee, and Sukhyun Ryu. Transmission Dynamics and Parameters for Pertussis during School-Based Outbreak, South Korea, 2024. *Emerging Infectious Diseases*, 31(7):1330, 2025.

[31] Dennis E te Beest, Michiel van Boven, Mariëtte Hooiveld, Carline van den Dool, and Jacco Wallinga. Driving factors of influenza transmission in the netherlands. *American journal of epidemiology*, 178(9):1469–1477, 2013.

[32] Stephen M Kissler, Christine Tedijanto, Edward Goldstein, Yonatan H Grad, and Marc Lipsitch. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*, 368(6493):860–868, 2020.

[33] Pejman Rohani, Xue Zhong, and Aaron A King. Contact network structure explains the changing epidemiology of pertussis. *Science*, 330(6006):982–985, 2010.

[34] Daihai He, Edward L Ionides, and Aaron A King. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of the Royal Society Interface*, 7(43):271–283, 2010.

[35] Roy M Anderson and Robert M May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1991.

[36] Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of statistical software*, 76:1–32, 2017.

[37] Marc Choisy and Pejman Rohani. Changing spatial epidemiology of pertussis in continental USA. *Proceedings of the Royal Society B: Biological Sciences*, 279(1747):4574–4581, 2012.

[38] Jacco Wallinga and Marc Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609):599–604, 2007.

[39] Roy M Anderson and Robert M May. Directly transmitted infections diseases: control by vaccination. *Science*, 215(4536):1053–1060, 1982.

[40] Jennie S Lavine and Pejman Rohani. Resolving pertussis immunity and vaccine effectiveness using incidence time series. *Expert review of vaccines*, 11(11):1319–1329, 2012.

[41] Sang Woo Park, Inga Holmdahl, Emily Howerton, Wenchang Yang, Rachel E Baker, Gabriel A Vecchi, Sarah Cobey, C Jessica E Metcalf, and Bryan T Grenfell. Interplay between climate and childhood mixing can explain a sudden shift in RSV seasonality in Japan. *Nature Communications*, 2025.

[42] Bärbel F Finkenstädt, Ottar N Bjørnstad, and Bryan T Grenfell. A stochastic model for extinction and recurrence of epidemics: estimation and inference for measles outbreaks. *Biostatistics*, 3(4):493–510, 2002.

[43] James O Lloyd-Smith, Sebastian J Schreiber, P Ekkehard Kopp, and Wayne M Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359, 2005.

[44] FMG Magpantay, M Domenech De Cellés, P Rohani, and AA King. Pertussis immunity and epidemiology: mode and duration of vaccine-induced immunity. *Parasitology*, 143(7):835–849, 2016.

[45] Tina Tan, Tine Dalby, Kevin Forsyth, Scott A Halperin, Ulrich Heininger, Daniela Hozbor, Stanley Plotkin, Rolando Ulloa-Gutierrez, and Carl Heinz Wirsing Von König. Pertussis across the globe: recent epidemiologic

trends from 2000 to 2013. *The Pediatric infectious disease journal*, 34(9):e222–e232, 2015.

[46] Matthieu Domenech de Cellès, Felicia MG Magpantay, Aaron A King, and Pejman Rohani. The impact of past vaccination coverage and immunity on pertussis resurgence. *Science translational medicine*, 10(434):eaaj1748, 2018.

[47] Michael J Mina, C Jessica E Metcalf, Adrian B McDermott, Daniel C Douek, Jeremy Farrar, and Bryan T Grenfell. A global lmmunological observatory to meet a time of pandemics. *Elife*, 9:e58989, 2020.

[48] Hai Nguyen-Tran, Sang Woo Park, Kevin Messacar, Samuel R Dominguez, Matthew R Vogt, Sallie Permar, Perdita Permaul, Michelle Hernandez, Daniel C Douek, Adrian B McDermott, et al. Enterovirus D68: a test case for the use of immunological surveillance to develop tools to mitigate the pandemic potential of emerging pathogens. *The Lancet Microbe*, 3(2):e83–e85, 2022.

[49] Hai Nguyen-Tran, Sang Woo Park, Matthew R Vogt, Perdita Permaul, Alicen B Spaulding, Michelle L Hernandez, Jennifer A Bohl, Sucheta Godbole, Tracy J Ruckwardt, Peter W Krug, et al. Dynamics of endemic virus re-emergence in children in the USA following the COVID-19 pandemic (2022–23): A prospective, multicentre, longitudinal, immunoepidemiological surveillance study. *The Lancet Infectious Diseases*, 26(1):22–33, 2026.

[50] Seonghui Cho, Dong Wook Kim, Chiara Achangwa, Junseo Oh, and Sukhyun Ryu. Pertussis in the elderly: Plausible amplifiers of persistent community transmission of pertussis in South Korea. *Journal of Infection*, 89(3), 2024.

[51] Tom Britton, Frank Ball, and Pieter Trapman. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *science*, 369(6505):846–849, 2020.

[52] David L Heymann and Guénaël R Rodier. Hot spots in a wired world: Who surveillance of emerging and re-emerging infectious diseases. *The Lancet infectious diseases*, 1(5):345–353, 2001.

[53] Mahmoud M Naguib, Patrik Ellström, Josef D Järhult, Åke Lundkvist, and Björn Olsen. Towards pandemic preparedness beyond COVID-19. *The Lancet Microbe*, 1(5):e185–e186, 2020.