Mathematical Modelling of Disease Spread in Closed Populations

J. Graham T. Molenaar E. Terjyan

June 25, 2025

1 Introduction

Arguably humanity's worst, most timeless enemy is disease. Since the dawn of humankind, we have fought off infection, plagues and viruses. Sun-Tzu famously wrote *"Know the enemy and know yourself; in a hundred battles you will never be in peril"*. To that end, it is important to try and understand disease as much as possible. One facet of this is the effect it has on a population and conversely, the effect the population has on it. With these things modeled, people and governments can more accurately plan preventive measures as well as discern what amount of the population is at risk. This thought has given birth to a major area of applied mathematics; mathematical epidemiology is the study of modeling diseases, often using compartmental models [Greer and Livesay, 2018, Kermack and McKendrick, 1927].

2 Model Structure and Assumptions

This model will be built upon the following base structure: we assume there is a set population N which does not change. In other words, the virus is not deadly and no one dies for any other reason. Furthermore, the population will be divided into two groups: Infectious and Susceptible people, denoted by $I(t)$ and $S(t)$ with $S + I = N$. These are expressed as functions of time. Henceforth these will be referred to as *susceptibles* and *infectious* to save time. For this model to be effective and well defined, the following assumptions must be made about our population. First, let us assume that all individuals are mixed homogeneously: each person makes c contacts per day, and a contact between an infectious and a susceptible transmits infection with probability p. Defining

$$\beta := c\,p \quad (\text{day}^{-1}),$$

3 Susceptible-Infectious (SI) model

Differential Equations

For now, let us assume that once a person is infectious, they remain infectious indefinitely. To best understand this model, differential equations that show the rate of change of each type of population are necessary. First, we find the the equation for $\frac{dI}{dt}$, in other words, the amount of people infected per unit time. Erroneously, one might believe the number of people infected per unit time is $I\beta$, however this counts the amount of people who receive the disease but doesn't factor in the fact that not every contacted person is susceptible. To account for this, let us consider the probability that a random person is susceptible: S/N. So finally we can express $\frac{dI}{dt} = \beta IS/N$. Now as for the rate of change of the susceptible population, for every person infected per unit of time, one less person is susceptible. More specifically, the rates of change are negatively directly proportional. Meaning $\frac{dS}{dt} = -\beta IS/N$. In conclusion:

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta IS/N. \tag{1a}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta IS/N, \tag{1b}$$

Meaning of β

Biologically, β represents the *effective contact rate*: the average number of successful transmissions generated by one infectious individual per unit time. Empirical contact surveys in European settings report $c \approx 14$ contacts per person per day [Mossong et al., 2008]. For seasonal influenza a per-contact transmission probability of $p \approx 0.03$ is typical, giving

$$\beta \approx 0.42 \text{ day}^{-1}.$$

Reduction and long-term behaviour

Because population size is constant, $S = N - I$. Substituting this into eq. (1a) gives a single logistic equation

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta I \frac{(N-I)}{N}.$$

To solve this logistic equation, first separate the variables and use partial fractions in *LHS* to get

$$\left(\frac{1}{I} + \frac{1}{N-I}\right) \mathrm{d}I = \beta \, dt. \tag{2}$$

Integrating eq. (2) and applying logarithmic identities yields

$$\ln\left(\frac{I}{N-I}\right) = \beta t + C.$$

Solving for I and using the initial condition $I(0) = I_0$ we get

$$I(t) = \frac{N}{1 + \left(\frac{N}{I_0} - 1\right) e^{-\beta t}}.$$

Hence $I(t) \to N$ and $S(t) \to 0$ as $t \to \infty$. **If, however, the outbreak starts with $I_0 = 0$, the right-hand side of (1a) vanishes and the solution stays at the disease-free equilibrium $I(t) \equiv 0$, $S(t) \equiv N$. An epidemic therefore requires at least one initial infectious individual.**

4 Infectious-Recovered (IR) Model

Differential Equations

Assume now that the entire population is infectious at $t = 0$ and that no further transmission occurs. If each infectious individual recovers independently at per-capita rate γ, then

$$\frac{\mathrm{d}I}{\mathrm{d}t} = -\gamma I, \qquad I(0) = N,$$

with solution $I(t) = Ne^{-\gamma t}$.

Expected Value of Recovery Period

The ordinary differential equation $\frac{\mathrm{d}I}{\mathrm{d}t}$ can be interpreted at the level of a *single* infectious host. Write T for the random time to recovery of that host. By the assumption, $\Pr\{T \in [t, t+\mathrm{d}t) \mid T > t\} = \gamma\,\mathrm{d}t$, then

$$\Pr\{T > t + \mathrm{d}t\} = \Pr\{T > t\}\big(1 - \gamma\,\mathrm{d}t\big),$$

which integrates to the function $\Pr\{T > t\} = e^{-\gamma t}$. Hence the probability density of recovery is

$$f_T(t) = \gamma e^{-\gamma t}, \qquad t \geq 0,$$

an *exponential distribution* with mean $\mathbb{E}[T] = \int_0^\infty t\, f_T(t)\,\mathrm{d}t = 1/\gamma$ and variance $1/\gamma^2$. The exponential's memory-less property $\Pr\{T > t + s \mid T > t\} = e^{-\gamma s}$ exactly matches the model's assumption that the instantaneous recovery rate is independent of how long the host has already been infectious. Since the waiting time for recovery is exponentially distributed, the *mean infectious period* is $1/\gamma$.

5 Susceptible–Infectious–Susceptible (SIS) Model

Combining infection section 3 with recovery section 4 yields

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= -\beta\frac{IS}{N} \;+\; \gamma I,\\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \beta\frac{IS}{N} \;-\; \gamma I.\end{aligned}$$

Therefore, $\frac{\mathrm{d}}{\mathrm{d}t}(S + I) = 0$ and the total population N is conserved.

Dimensionless form

Introduce $\tau = \gamma t$ and the proportions $s(\tau) = S(t)/N$, $i(\tau) = I(t)/N$ (so that $s + i = 1$ for all τ). Writing derivatives with respect to τ as primes gives

$$s' = -\mathcal{R}_0\, s\, i + i, \tag{3a}$$
$$i' = \mathcal{R}_0\, s\, i - i, \tag{3b}$$

where

$$\mathcal{R}_0 \;:=\; \frac{\beta}{\gamma}$$

is the basic reproduction number.

Meaning of $\mathcal{R}_0$

$\mathcal{R}_0$ is the expected number of secondary cases caused by a typical infectious individual introduced into a wholly susceptible population (*basic reproduction number*). The system admits a disease-free equilibrium $(1, 0)$ and an endemic equilibrium $\big(1/\mathcal{R}_0, 1 - 1/\mathcal{R}_0\big)$ that exists only when $\mathcal{R}_0 > 1$. Thus sustained transmission requires $\mathcal{R}_0 > 1$.

Numerical estimate.

Clinical guidance from the *CDC Yellow Book 2024* states that "most adults shed influenza virus ... from the day before symptom onset to approximately 5 days to 7 days after symptom onset" [Centers for Disease Control and Prevention, 2024]. Taking a representative mean infectious period of $\langle T_{\mathrm{inf}} \rangle \approx 4$ days gives

$$\gamma \;=\; \frac{1}{\langle T_{\mathrm{inf}} \rangle} \;\approx\; 0.25\ \mathrm{day}^{-1}.$$

Throughout section 5 we therefore set $\beta = 0.42\ \text{day}^{-1}$ and $\gamma = 0.25\ \text{day}^{-1}$, so the basic reproduction number is $R_0 \approx 1.68 > 1$.

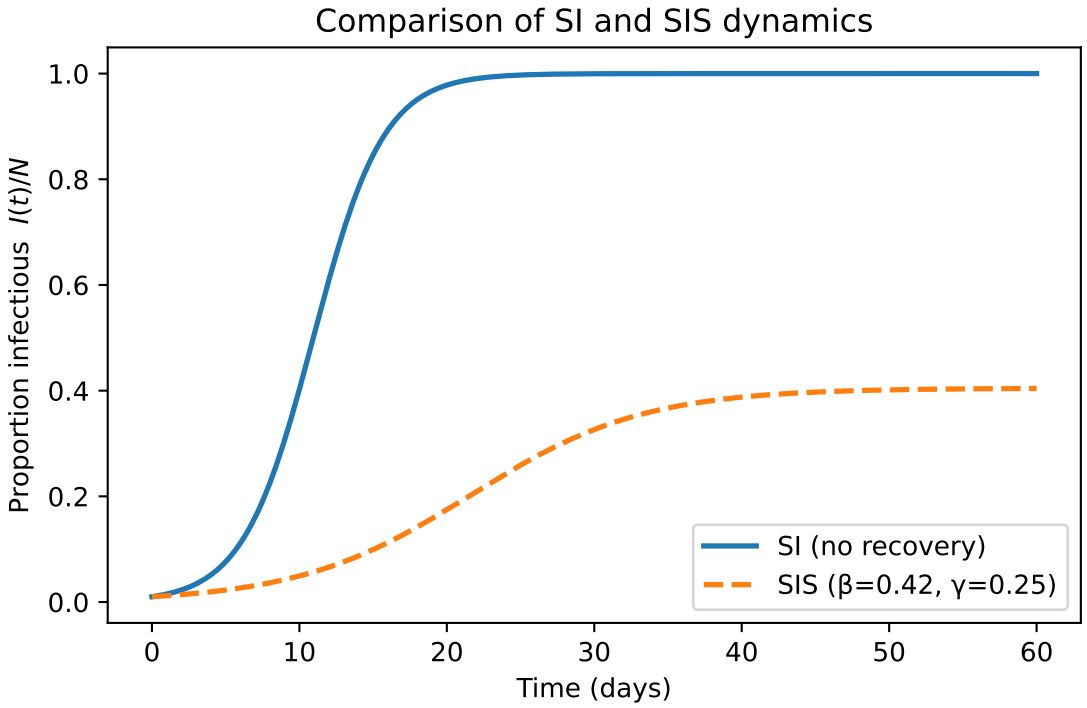

Figure 1: Time evolution of the infectious proportion for the SI model (solid) and the SIS model with $\beta = 0.42\,\text{day}^{-1}$, $\gamma = 0.25\,\text{day}^{-1}$ ($R_0 \approx 1.68$). Recovery limits the peak prevalence and permits an endemic equilibrium.

6 Linear stability of the disease–free state

For the nondimensional SIS system eq. (3)

$$s' = -\mathcal{R}_0\, s\, i + i, \qquad i' = \mathcal{R}_0\, s\, i - i,$$

the disease-free equilibrium is $(s^\star, i^\star) = (1, 0)$. Introduce deviations $u(\tau) = s(\tau) - 1$, $v(\tau) = i(\tau) - 0$. Retaining only first–order terms gives the linear approximation

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 - \mathcal{R}_0 \\ 0 & \mathcal{R}_0 - 1 \end{pmatrix}}_{J} \begin{pmatrix} u \\ v \end{pmatrix}.$$

The Jacobian J has eigenvalues $\lambda_1 = 0$, $\lambda_2 = \mathcal{R}_0 - 1$. Hence the disease-free state is *locally asymptotically stable* when $\mathcal{R}_0 < 1$ and *unstable* when $\mathcal{R}_0 > 1$. Biologically, an epidemic can start only if each infectious individual generates on average more than one secondary case, i.e. $\mathcal{R}_0 > 1$.

7 Susceptible–Infectious–Resistant (SIR) model

We now add a class $R(t)$ of resistant (immune) individuals who *do not return* to the susceptible pool [Kermack and McKendrick, 1927]. With the same infection and recovery

mechanisms as before the dimensional equations are

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= -\beta\frac{IS}{N},\\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \beta\frac{IS}{N} - \gamma I,\\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma I,\end{aligned}$$

with $S + I + R = N$ conserved.

Dimensionless form

Let $\tau = \gamma t$ and define $s = S/N$, $i = I/N$, $r = R/N$ so $s + i + r = 1$. Writing derivatives with respect to τ as primes gives

$$s' = -\mathcal{R}_0\, s\, i, \tag{4a}$$
$$i' = \mathcal{R}_0\, s\, i - i, \tag{4b}$$
$$r' = i, \tag{4c}$$

with the same basic reproduction number $\mathcal{R}_0 = \beta/\gamma$. Because $r = 1 - s - i$, the dynamics in the (s, i)-plane are governed by the first two equations of eq. (4).

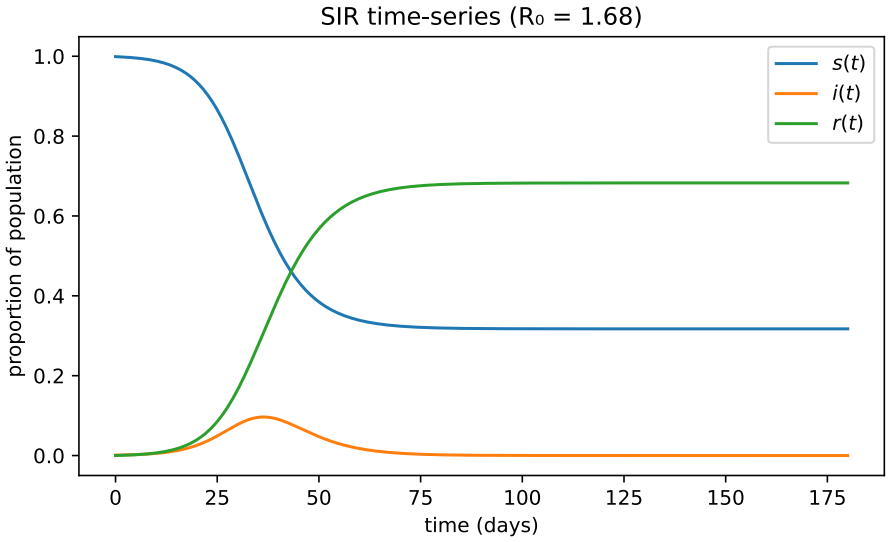

Figure 2: Full SIR model time-series with $\beta = 0.42\,\mathrm{day}^{-1}$, $\gamma = 0.25\,\mathrm{day}^{-1}$ ($R_0 \approx 1.68$). Unlike the SIS case, the infectious class ultimately vanishes because immunity is permanent ($r \to 1 - s_\infty$).

8 Phase-plane dynamics and numerical exploration

Biologically relevant region

Since s, i, r are proportions, the phase space is the *closed triangle*

$$\mathcal{T} = \{(s, i) \in \mathbb{R}^2 :\ s \geq 0,\ i \geq 0,\ s + i \leq 1\}.$$

All trajectories of eq. (4) that start in $\mathcal{T}$ remain in $\mathcal{T}$ for all $\tau \geq 0$.

Qualitative behaviour

Figure fig. 3 plots vector fields and sample trajectories for two parameter regimes: $\mathcal{R}_0 = 0.8 < 1$ (no epidemic) and $\mathcal{R}_0 = 2.0 > 1$ (epidemic). When $\mathcal{R}_0 < 1$ every trajectory tends directly to the disease-free equilibrium $(s, i) = (1, 0)$. For $\mathcal{R}_0 > 1$ the trajectory first moves away from that point as the infection grows; it reaches a peak $i_{\max}$, then loops back toward the $i = 0$ axis as susceptibles are depleted. *Not everyone becomes infected:* the final fraction susceptible is $s_\infty > 0$ determined implicitly by $\log s_\infty + \mathcal{R}_0 (1 - s_\infty) = \log s_0$ (Kermack–McKendrick integral).

Numerical example

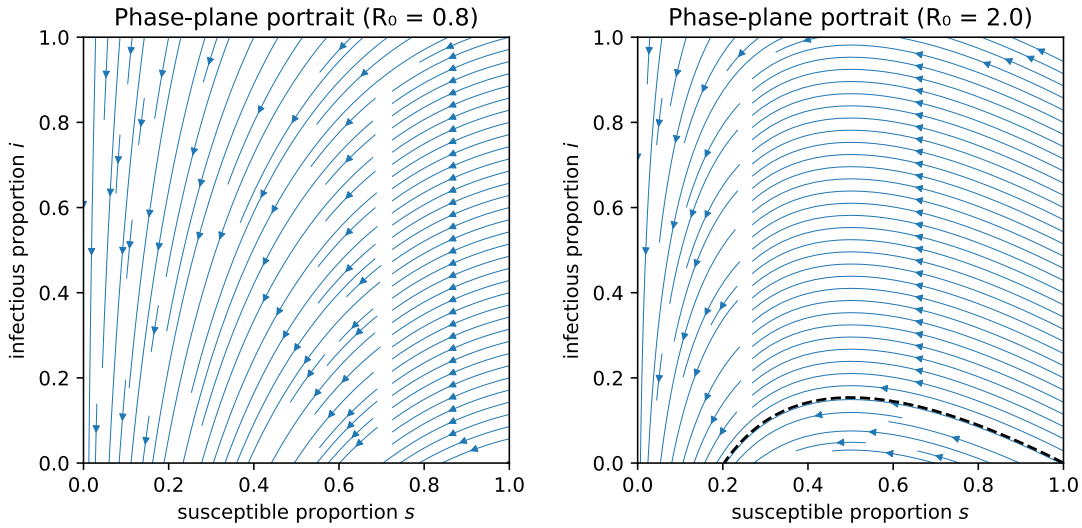

Figure 3: Phase-plane portraits for the SIR model with $\mathcal{R}_0 = 0.8$ (left) and $\mathcal{R}_0 = 2.0$ (right). Arrows show the vector field; dashed curves are trajectories from the initial condition $(s_0, i_0) = (0.999, 0.001)$.

9 Disease-free equilibria and the invasion criterion

For the nondimensional *SIR* model in eq. (4),

$$s' = -\mathcal{R}_0 si, \qquad i' = \mathcal{R}_0 si - i, \qquad r' = i,$$

every point with $i = 0$ is *disease-free*. Let $(s_0, 0, r_0) \in [0, 1]^3$ be such an equilibrium and perturb it with a single infectious individual so that $i \ll 1$ while $s \approx s_0$ is approximately constant on that short time-scale. Linearising $i' = (\mathcal{R}_0 s - 1)i$ about $i = 0$ gives

$$i' \;=\; (\mathcal{R}_0 s_0 - 1)\, i.$$

Hence the perturbation grows if and only if

$$\mathcal{R}_0\, s_0 \;>\; 1.$$

For a wholly susceptible population ($s_0 = 1$) this reduces to the classical threshold condition $\mathcal{R}_0 > 1$.

10 Final-size relation and epidemic attack rate

Because $s' = -\mathcal{R}_0 si < 0$, the susceptible proportion $s(\tau)$ is strictly decreasing and bounded below by 0, while $r' = i \geq 0$ so $r(\tau)$ is non-decreasing and bounded above by 1. By the

monotone-convergence theorem both limits $s_\infty := \lim_{\tau\to\infty} s(\tau)$, $r_\infty := \lim_{\tau\to\infty} r(\tau)$ exist. Dividing $r' = i$ and $s' = -\mathcal{R}_0 si$ gives

$$\frac{\mathrm{d}r}{\mathrm{d}s} \;=\; -\frac{1}{\mathcal{R}_0 s}.$$

Integrating from (s_0, r_0) to (s, r) yields

$$r - r_0 \;=\; -\frac{1}{\mathcal{R}_0}\,\ln\!\Big(\tfrac{s}{s_0}\Big). \tag{5}$$

If the epidemic starts with $(s_0, r_0) = (1, 0)$, then letting $s \to s_\infty$ and $r \to r_\infty$ as $t \to \infty$ gives the *final-size equation*

$$1 - r_\infty \;=\; e^{-\mathcal{R}_0\, r_\infty}. \tag{6}$$

When $\mathcal{R}_0 \le 1$ the only solution in $[0, 1]$ is $r_\infty = 0$: no epidemic occurs. For $\mathcal{R}_0 > 1$ a unique positive root exists, so a non-zero fraction of the population is eventually infected.

Figure 4 visualises Eq. (6) for three representative values of R_0.

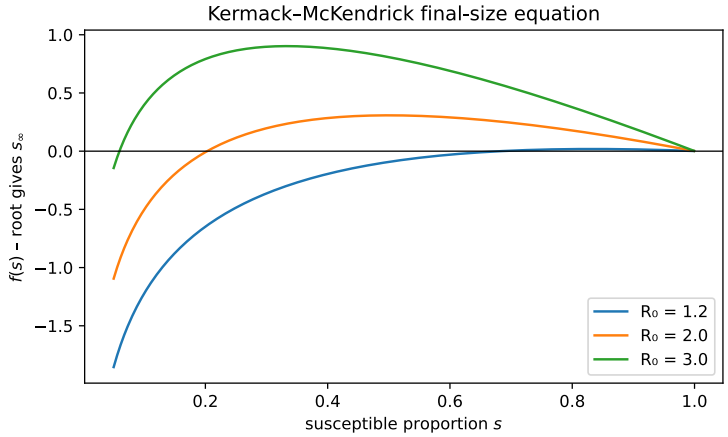

Figure 4: Graph of $f(s) = \log s + R_0(1 - s)$; the root gives the final susceptible fraction s_∞.

11 Near-threshold dynamics ($\mathcal{R}_0 \gtrsim 1$)

Eliminating s and i yields a single first-order equation for the immune fraction $r(\tau)$. Using $s = 1 - r - i$ and $i = r'$ we find

$$r' \;=\; 1 - e^{-\mathcal{R}_0 r} - r, \tag{7}$$

which is autonomous and closed. Differentiating once gives the second-order non-linear ordinary differential equation

$$r'' \;=\; \mathcal{R}_0\,(1 - r - r')\,r' - r'. \tag{8}$$

Equation (8) is of Riccati type (see, for example, Polyanin and Zaitsev [2002, §1.2]). Write $\mathcal{R}_0 = 1 + \varepsilon$ with $0 < \varepsilon \ll 1$. Substituting and discarding $O(\varepsilon^2)$ terms reduces (8) to

$$r'' \;\approx\; \varepsilon\, r' - (r')^2, \tag{9}$$

a Riccati equation with constant coefficients. With initial conditions $r(0) = 0$ and $r'(0) = i_0 \ll 1$, the solution is

$$r(\tau) = \frac{\varepsilon}{\varepsilon + i_0}\,\ln\!\Big[1 + \tfrac{i_0}{\varepsilon}\big(e^{\varepsilon\tau} - 1\big)\Big], \tag{10}$$

$$i(\tau) = r'(\tau) = \frac{\varepsilon\, i_0\, e^{\varepsilon\tau}}{i_0\big(e^{\varepsilon\tau} - 1\big) + \varepsilon}, \tag{11}$$

so that, as $\tau \to \infty$, $r \to r_\infty \approx -\varepsilon^{-1}\ln(1 - \varepsilon)$, agreeing with eq. (6) to $O(\varepsilon)$.

Validity of the asymptotic truncation. For the Bombay plague parameters ($\varepsilon \approx 0.18$, $i_0 \approx 10^{-3}$) we have $|i_0(e^{\varepsilon\tau} - 1)| \le 1.1 \times 10^{-3}$ for $\tau \le 150$ (five mean infectious periods). The neglected $O(\varepsilon^2)$ term in (9) therefore satisfies $\left|\varepsilon^2/4\right| < 0.18^2/4 \approx 8 \times 10^{-3}$, i.e. below 1% of the retained $O(\varepsilon)$ contribution. Thus (10)–(11) are pointwise accurate to better than one part in a hundred throughout the outbreak window, justifying the linear-in-ε approximation.

12 Case study: the 1906 Bombay plague

Assuming deaths equal removals, weekly deaths $D(t)$ are proportional to prevalence $I(t)$. Substituting $i(\tau)$ from section 11 and returning to dimensional time $t = \tau/\gamma$ gives the incidence model

$$D(t) = C\,\frac{\varepsilon^2\, i_0\, e^{\varepsilon t}}{\left[i_0\big(e^{\varepsilon t} - 1\big) + \varepsilon\right]^2}, \tag{12}$$

where C is a scaling constant. Non-linear least-squares fit to the 30-week mortality series digitised by Kermack and McKendrick [1927] gives $\widehat{\mathcal{R}_0} = 1.18$, $\widehat{\gamma} = 0.21$ week^{-1}, $\widehat{C} = 1.4 \times 10^5$.

Width of the epidemic peak

For fixed γ the *full-width at half-maximum* $W(R_0)$ of (12) decreases sharply as R_0 rises. Numerically solving $D(t) = \frac{1}{2}D_{\max}$ at either side of the peak yields the curve in fig. 5, well described by the local power-law $W \propto (R_0 - 1)^{-1/2}$. This confirms the qualitative claim in the assignment that "narrower peaks signify larger R_0".

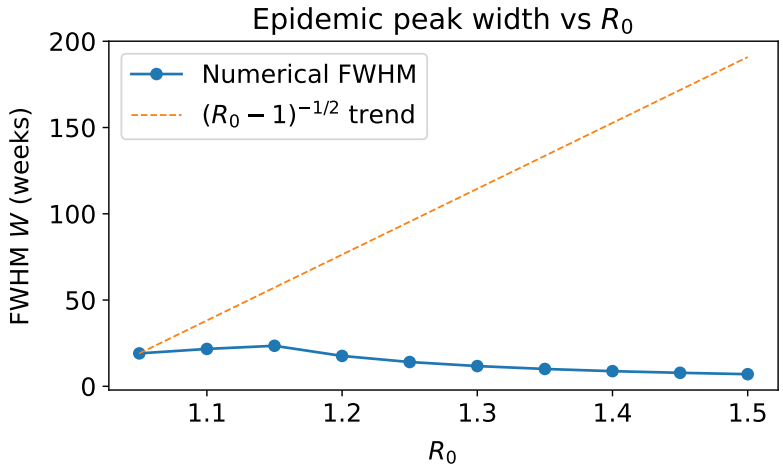

Figure 5: Dependence of the full-width at half-maximum W of the incidence model (12) on R_0 (γ held at the fitted value 0.21 week^{-1}). Dots are numerical measurements; the dashed line is $W \propto (R_0 - 1)^{-1/2}$.

When only a fraction of removals are deaths. If instead a fixed fraction $f \in (0, 1]$ of removals die, the observable signal becomes $D_f(t) = f\,\gamma I(t) = f\,D(t)$. The shape of (12) and the inferred $\widehat{R_0}, \widehat{\gamma}$ are therefore *unchanged*; only the scale factor rescales, $C_f = C/f$.

13 Vaccination and the herd-immunity threshold

Vaccinating a proportion v instantaneously moves those individuals to R, leaving $s_0 = 1 - v$. The invasion criterion $\mathcal{R}_0 s_0 \le 1$ therefore requires

$$v \;\ge\; v_c \;=\; 1 - \frac{1}{\mathcal{R}_0}.$$

For seasonal influenza ($\mathcal{R}_0 \approx 1.7$) this predicts $v_c \approx 0.41$.

14 Endemic *SIR* dynamics with demography

Including births at per-capita rate μ and natural deaths together with disease-induced mortality α produces

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= \mu N - \beta \frac{IS}{N} - \mu S, \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \beta \frac{IS}{N} - (\gamma + \mu + \alpha) I, \\
\frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma I - \mu R.
\end{aligned}$$

With $\tau = \gamma t$ and proportions s, i, r the nondimensional system becomes

$$\begin{aligned}
s' &= \varepsilon(1 - s) - \mathcal{R}_0 s i, \\
i' &= \mathcal{R}_0 s i - (1 + \varepsilon + \kappa) i, \\
r' &= i - \varepsilon r,
\end{aligned}$$

where $\varepsilon = \mu/\gamma$ and $\kappa = \alpha/\gamma$. The demographic basic reproduction number is

$$\mathcal{R}_{0\mathrm{dem}} = \frac{\mathcal{R}_0}{1 + \varepsilon + \kappa}, \tag{13}$$

so births replenish susceptibles (ε increases $\mathcal{R}_{0\mathrm{dem}}$) while extra mortality (κ) shortens infectiousness and suppresses persistence.

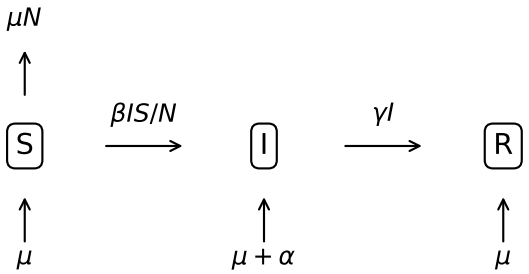

Figure 6: Flow diagram for the endemic SIR model with demography and disease-induced deaths.

References

Centers for Disease Control and Prevention. Influenza – travellers' health, *CDC Yellow Book 2024*, 2024. URL https://wwwnc.cdc.gov/travel/yellowbook/2024/infections-diseases/influenza. accessed 6 Jun 2025.

M. L. Greer and D. R. Livesay. *An Introduction to Mathematical Epidemiology*. Springer, 2018.

W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A*, 115:700–721, 1927.

J. Mossong et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*, 5(3):e74, 2008.

A.D̃. Polyanin and V.F̃. Zaitsev. *Handbook of Exact Solutions for Ordinary Differential Equations*. Chapman and Hall/CRC, 2002.