The Impact of Neglecting Vaccine Unwillingness in Epidemiology Models

Glenn Ledder

Abstract

With significant population fractions in many societies who refuse vaccines, it is important to reconsider how vaccination is incorporated into compartmental epidemiology models. It is still most common to apply the vaccination rate to the entire class of susceptibles, rather than to use the more realistic assumption that the vaccination rate function should depend only on the population of susceptibles who are willing and able to receive a vaccination. This study uses a simple generic disease model to address two questions: (1) How much error is introduced in key model outcomes by neglecting vaccine unwillingness?, and (2) Can the error be reduced by incorporating vaccine unwillingness into the vaccination rate constant rather than the rate diagram? The answers depend greatly on the time scale of interest. For the endemic time scale, where longterm behavior is studied with equilibrium point analysis, the error in neglecting unwillingess is large and cannot be improved upon by decreasing the vaccination rate constant. For the epidemic time scale, where the first big epidemic wave is studied with numerical simulations, the error can still be significant, particularly for diseases that are relatively less infectious and vaccination programs that are relatively slow.

1   Introduction

Vaccination is one of many processes that can be incorporated into epidemiology models. Its effect on the rate diagram varies because the benefits of vaccination depends on the disease and the vaccine. In the best scenario, vaccination confers guaranteed and permanent immunity. Lesser benefits are also common:

1. Temporary or partial immunity;

2. Decreased probability of being infected;

3. Decreased severity of infection;

4. Faster recovery;

5. Decreased risk of infecting others.

Historically, it has been standard to incorporate vaccination as a linear process with rate proportional to the full susceptible population. In 2022, Ledder [8] presented the idea that vaccination should only be applied to a subclass of willing susceptibles, taking into account that some susceptibles are either unwilling or unable to be vaccinated. Numerous studies, both before and after the onset of the COVID-19 pandemic, have shown that the fraction of individuals who are unwilling or unable is higher than might previously have been thought based on public health policies, as there has always been a significant minority of people who do not follow those policies [2]. A study published by the Annenberg Foundation in 2023 showed that only 71% of Americans thought vaccines were safe (and this presumably included some who thought vaccines are not effective) and that only 63% of Americans at that time thought that the COVID vaccine was less dangerous than the COVID disease [1]. Given recent trends [3, 10, 11, 12], it is reasonable to guess that 30% is a conservative estimate of the fraction of Americans who will be unwilling to receive a vaccine for the next epidemic. While other societies may exhibit different attitudes toward vaccination, it is reasonable to expect that some level of vaccine unwillingness is present everywhere.

Restricting vaccination to a subclass of susceptibles makes a model more complicated. At minimum, we need an additional state variable, either for willing susceptibles or unwilling susceptibles, and, depending

on the biological assumptions about the disease, classes other than the susceptibles may also need to be subdivided. This raises the question of whether the extra complexity is justified by the improved realism. This is not an uncommon issue in epidemiology models; for example, it is standard practice to use a single-phase transition model component for processes such as incubation and recovery in spite of the fact that doing so tacitly assumes that times for such processes are exponentially distributed.

In this study, we examine the extent of differences in model outcomes caused by adopting the naive assumption that all susceptibles are in the vaccination queue (with or without deprecating the rate constant to account for unwillingness) rather than the more realistic assumption that only some susceptibles are in that queue. For a conservative approach, we allow for the possibility that benefits from vaccination are relatively minor. Compared to susceptible individuals, we assume three benefits of vaccination: lower rate of infection, lower infectiousness, and faster recovery.

The significance of the decision on how to incorporate vaccination into the model may depend on the time frame for which the model is intended; hence, we separately consider scenarios for the endemic and epidemic phases of the disease. In the latter case, the amount of time over which the model is tracked could affect the perceived importance of the vaccination modeling, so we examine how the effect of the difference progresses over time. The key questions are these:

1. How much error results from completely neglecting vaccine unwillingness?

2. To what extent can the error be corrected by incorporating unwillingness into the vaccination rate constant rather than into the rate diagram?

We examine how the answer depends on the basic reproduction number of the disease ($\mathcal{R}_0$), the size of the unwilling fraction (W), the vaccination rate constant scaled with respect to disease duration (θ), the relative susceptibility of vaccinated individuals ($\sigma < 1$), the relative infectiousness of previously-vaccinated infectious individuals ($\eta < 1$), and the relative recovery rate of previously-vaccinated individuals ($\alpha > 1$).

2 An SIR Model with Vaccination

Figure 1 shows an SIR model that includes vaccination and demographics. The total birth rate is taken to match the death rate in order to keep the population at a constant value, which we can take without loss of generality to be 1. In addition to classifying population members as susceptible, infectious, or removed, we classify individuals as either willing or unwilling/unable to accept vaccination, with fixed fractions W and U.[1] The model allows for three possible vaccination benefits, relative to unvaccinated susceptibles:

1. The infection rate is multiplied by a factor $\sigma \leq 1$;

2. Infectiousness is multiplied by a factor $\eta \leq 1$;

3. Recovery rate is increased by a factor $\alpha \geq 1$.

These three parameters combine into a single parameter that represents the relative disease burden of vaccinated individuals:
$$\xi = \frac{\sigma\eta}{\alpha}. \tag{1}$$

It is reasonable to assume that this imperfect vaccination has a roughly equal quantitative impact for the three modifications. Hence, when needed, we make the additional simplifying assumption that the effects of vaccination on the infection rate, infectiousness, and recovery time are equal, that is,

$$\alpha\sigma = 1, \qquad \eta = \sigma, \qquad \xi = \sigma^3. \tag{2}$$

To keep the models simple, we assume that disease recovery confers lifetime immunity. Relaxing this assumption would increase the importance of vaccine unwillingness because recovered individuals whose immunity lapses would again be susceptible, when additional vaccination would again require willingness.

To accommodate the differences between vaccinated and unvaccinated susceptibles, it is necessary to have three susceptible classes, one willing but pre-vaccinated (S_W), one vaccinated (V), and one unwilling

---

[1]See Jiang et al [5] for a more nuanced model in which attitudes change in response to disease levels.

(S_U) and to distinguish between infectious individuals who were (I_V) and were not (I) vaccinated. For convenience, we define the total population infectiousness as

$$\widehat{I} = I + \eta I_V \tag{3}$$

and use this quantity to simplify the formulas. The figure illustrates the more general endemic model with demographic processes of birth and natural death. In this case, willingness is incorporated into the birth rates.
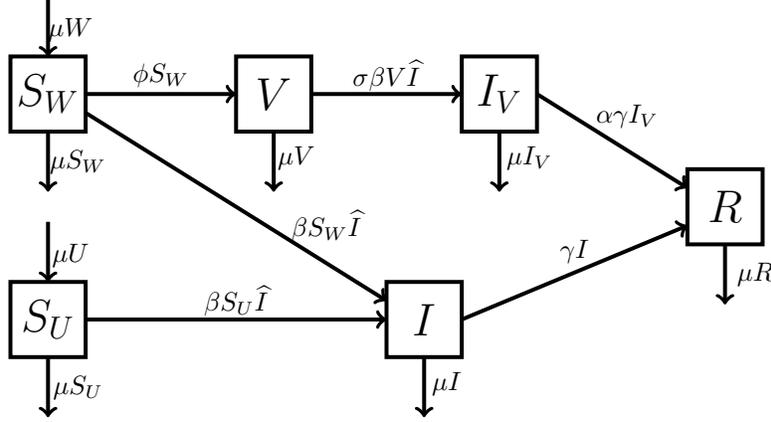

Figure 1: An SVIR model with vaccination applied only to a willing subclass of susceptibles.

From the diagram, we have the model

$$
\begin{aligned}
\dot{I} &= -(\gamma + \mu)I + \beta S\widehat{I}, \\
\dot{I}_V &= -(\alpha\gamma + \mu)I_V + \sigma\beta V\widehat{I}, \\
\dot{S} &= \mu - \mu S - \phi S_W - \beta S\widehat{I}, \\
\dot{V} &= \phi S_W - \mu V - \sigma\beta V\widehat{I}, \\
\dot{S_W} &= \mu W - (\phi + \mu)S_W - \beta S_W\widehat{I},
\end{aligned}
\tag{4}
$$

where we have used the full unvaccinated susceptible population $(S = S_W + S_U)$ as a variable in place of the unwilling subpopulation and use dots to indicate time derivatives.

2.1 Scaling on an Epidemic Time Scale

Full understanding of epidemiological models is facilitated by a proper scaling. Because the population in our model is constant, we can assume without loss of generality that the model populations represent population fractions. We still need to scale the time. There are two possibilities: the mean lifespan $1/\mu$ and the mean infectious duration $1/(\gamma + \mu)$ [7]. These scales yield slow and fast dimensionless time variables, given respectively as

$$T = \mu t. \qquad \tau = (\gamma + \mu)t. \tag{5}$$

For both scalings, we create dimensionless parameters using the fast time as a reference:

$$\epsilon = \frac{\mu}{\gamma + \mu} \ll 1, \qquad \mathcal{R}_0 = \frac{\beta}{\gamma + \mu}, \qquad \theta = \frac{\phi}{\gamma + \mu}, \qquad a = 1 - \alpha^{-1}, \tag{6}$$

3

with a added merely for algebraic convenience. For the epidemic scaling, we replace the time derivative using $d/dt = (\gamma + \mu)d/d\tau$ and then divide the resulting equations by the factor $\gamma + \mu$ to obtain the scaled model

$$
\begin{aligned}
\frac{dI}{d\tau} &= -I + \mathcal{R}_0 S\widehat{I}, & I(0) &= \epsilon I_0, \\
\frac{dI_V}{d\tau} &= -\alpha(1 - \epsilon a)I_V + \sigma\mathcal{R}_0 V\widehat{I}, & I_V(0) &= 0, \\
\frac{dS}{d\tau} &= \epsilon(1 - S) - \theta S_W - \mathcal{R}_0 S\widehat{I}, & S(0) &= 1 - \epsilon I_0, \\
\frac{dV}{d\tau} &= \theta S_W - \epsilon V - \sigma\mathcal{R}_0 V\widehat{I}, & V(0) &= 0, \\
\frac{dS_W}{d\tau} &= \epsilon(W - S_w) - \theta S_W - \mathcal{R}_0 S_W\widehat{I}, & S_W(0) &= WS(0),
\end{aligned}
\tag{7}
$$

where we have assumed an initial state in which nobody has been vaccinated or is removed from risk, so that the population consists only of a small number of infectives and the rest susceptibles. The factor ϵ in the initial conditions codifies the smallness of the initial infectious population without loss of generality.

Without doing any analysis, the scaled model offers some insights into the relative importance of the different epidemiological and demographic processes. In the epidemic scaling, all of the demographic terms (birth and natural death) are $O(\epsilon)$ relative to the epidemiological terms. Demographic processes are too slow to have much impact in the epidemic phase of a disease. Changes in the full susceptible class and the willing susceptible subclass are dominated by processes that decrease these populations. We can consider the epidemic phase to have ended when these populations have been so decreased that the birth process is sufficient to start recovery of the susceptible population. This suggests that the primary goal of epidemic analysis should be to identify the total population fraction that has escaped the epidemic at the time τ_∞, arbitrarily taken as the time when the total infection rate of susceptibles $(\mathcal{R}_0 S\widehat{I})$ is equal to the birth rate (ϵ); then we define the output of the epidemic model in terms of the willing population fraction as the function

$$
f(W; \mathcal{R}_0, \sigma, \eta, \alpha, \theta, \epsilon, I_0) = S(\tau_\infty) + V(\tau_\infty).
\tag{8}
$$

The function notation with the semicolon is used to identify f as a family of functions of the single variable W, distinguished by a vector of parameters. The function f needs to be determined from numerical simulations. An additional outcome of the epidemic analysis will be the graphs of $\widehat{I}$ and V, showing the total infectiousness and vaccinated population size over time.

For the modification of the model where we try to account for unwillingness without having two susceptibles classes, the S, V, and S_W equations in (7) are replaced by

$$
\begin{aligned}
\frac{dS}{d\tau} &= \epsilon(1 - S) - W\theta S - \mathcal{R}_0 S\widehat{I}, & S(0) &= 1 - \epsilon I_0, \\
\frac{dV}{d\tau} &= W\theta S - \epsilon V - \sigma\mathcal{R}_0 V\widehat{I}, & V(0) &= 0,
\end{aligned}
\tag{9}
$$

where we have used $W\theta$ as the vaccination rate constant so that the initial vaccination rate is $W\theta$, matching that of the more general model $[\theta S_W(0) = W\theta]$, and the S_W equation is replaced by $S_W = S$.

2.2 Scaling on an Endemic Time Scale

To scale the model for the endemic scenario, we apply the derivative transformation $d/dt = \mu\, d/dT$ to obtain the system

$$
\begin{aligned}
\epsilon\frac{dI}{dT} &= -I + \mathcal{R}_0 S\widehat{I}, \\
\epsilon\frac{dI_V}{dT} &= -\alpha(1 - \epsilon a)I_V + \sigma\mathcal{R}_0 V\widehat{I}, \\
\frac{dS}{dT} &= 1 - S - \epsilon^{-1}\theta S_W - \epsilon^{-1}\mathcal{R}_0 S\widehat{I}, \\
\frac{dV}{dT} &= \epsilon^{-1}\theta S_W - V - \epsilon^{-1}\sigma\mathcal{R}_0 V\widehat{I}, \\
\frac{dS_W}{dT} &= W - (\epsilon^{-1}\theta + 1)S_W - \epsilon^{-1}\mathcal{R}_0 S_W\widehat{I}.
\end{aligned}
\tag{10}
$$

While the model is now dimensionless, it is not yet properly scaled because some of the variables are $O(\epsilon)$ rather than $O(1)$ [7]. This is most easily seen in the S equation, where it appears that S is dominated by the processes that decrease its value, as is true in the epidemic phase. In actuality, the variables S_W and $\widehat{I}$ are $O(\epsilon)$ on the endemic time scale, so that all three negative terms in the S equation are comparable in importance. To correctly scale all the state variables and simplify the notation, we make substitutions using the following definitions:

$$
I = \epsilon Y, \qquad \eta I_V = \epsilon X, \qquad \mathcal{R}_0\widehat{I} = \epsilon Z, \qquad \theta S_W = \epsilon P.
\tag{11}
$$

With these changes, and with $\sigma\eta$ replaced by $\alpha\xi$ using (1), we obtain the scaled model

$$
\begin{aligned}
\epsilon Y' &= -Y + SZ, \\
\epsilon X' &= -\alpha X + \alpha\xi VZ + \epsilon\alpha aX, \\
S' &= 1 - S - P - SZ, \\
V' &= P - V - \sigma VZ, \\
\epsilon P' &= \theta(W - P) - \epsilon P - \epsilon PZ,
\end{aligned}
\tag{12}
$$

with

$$
Z = \mathcal{R}_0(X + Y).
\tag{13}
$$

There are no initial conditions because the purpose of the endemic scaling is to determine long-term behavior.

As in the epidemic version, there are insights to be gained from mere examination of the scaled model. The primary processes governing changes in the willing susceptible population, represented by P, are birth and vaccination, with natural death and infection occurring at a slower rate and so having little impact. On the endemic scale, any difference between P and W is resolved quickly. This biological observation corresponds to the mathematical observation that the P equation is decoupled from the rest (at leading order in the asymptotic regime $\epsilon \to 0$) when it comes to determining stability.

There are two primary outcomes of interest for the endemic model, depending on which of the two possible equilibria, disease-free (DFE) or endemic (EDE), is stable. If the benefit of vaccination when everyone is willing is sufficient to make the DFE stable, then the most important question to ask is how much willingness is required to stabilize the DFE; this minimum level of willingness will be seen to depend on the disease, represented by the basic reproduction number in the absence of vaccination ($\mathcal{R}_0$), and the overall relative disease burden of the vaccinated, represented by ξ. For cases where the EDE is stable, the basic reproduction number is not the primary endemic outcome of interest. Instead we are interested primarily in the equilibrium rate of new infections in the asymptotic limit $\epsilon \to 0$; this is given by a function

$$
F(W; \mathcal{R}_0, \sigma, \eta, \alpha, \theta) = SZ + \sigma VZ = 1 - S - V,
\tag{14}
$$

which we'll see can be determined analytically. Leading order approximation is quite accurate, as the parameter ϵ is very small. For example, if the expected disease duration is about 2 weeks and the expected lifespan of the population is about 80 years, then $\epsilon \approx 0.0005$.

2.3 Parameter Estimates

Given that the model is not for a specific disease, we want to consider a relatively wide range of parameter values. We consider basic reproduction values in the range $1 \leq \mathcal{R}_0 \leq 6$, corresponding to diseases of relatively slight to very significant infectiousness, although not including diseases as infectious as measles. Most sources estimate the basic reproduction number for influenza at roughly 2–3, while the original strain of the COVID-19 virus had a basic reproduction number closer to 6 [6], with later strains very likely higher. The vaccination rate parameter is the ratio of mean disease duration to mean time for willing susceptibles to be vaccinated. With a typical disease duration of 2 weeks and a range of 4 to 20 weeks for mean vaccination time, we have the range $0.1 \leq \theta \leq 0.5$. The relative infection rate for vaccinated individuals (σ) can vary considerably. It is 0 for a near-perfect vaccine, such as the measles vaccine, while considerably lower for the seasonal flu vaccine. CDC data suggests that the worst case for the flu vaccine is a 30% reduction, while a more typical case is 50% reduction [4]. We refer to the values $\sigma = 0.7$ and $\sigma = 0.5$, corresponding to these percentages, as "marginal" and "moderate" vaccine effectiveness. For simplicity, we assume that the impact of vaccination on the infection duration ($1/\alpha$) and transmission rate of infectives (η) is the same as the relative rate of infection (σ). As noted earlier, a reasonable estimate for ϵ is 0.0005. The value of this parameter plays little role in the outcomes, as its effect is primarily on the amount of time between successive jumps in the infection rate after the first epidemic.

3 Endemic Analysis

Both the disease-free and endemic equilibria require the Jacobian matrix for stability determination; the general form of this matrix is

$$
J = \begin{bmatrix}
-\Gamma(1 - \mathcal{R}_0 S) & \Gamma \mathcal{R}_0 S & \Gamma Z & 0 & 0 \\
\Gamma \alpha \xi \mathcal{R}_0 V & -\Gamma \alpha (1 - \xi \mathcal{R}_0 V) + \alpha a & 0 & \Gamma \alpha \xi Z & 0 \\
-\mathcal{R}_0 S & -\mathcal{R}_0 S & -(1 + Z) & 0 & -1 \\
-\sigma \mathcal{R}_0 V & -\sigma \mathcal{R}_0 V & 0 & -(1 + \sigma Z) & 1 \\
-\mathcal{R}_0 P & -\mathcal{R}_0 P & 0 & 0 & -\Gamma \theta - (1 + Z)
\end{bmatrix}, \tag{15}
$$

where for notational convenience, we have defined the additional quantity $\Gamma = \epsilon^{-1}$.

3.1 The Disease-Free Equilibrium (DFE) and Basic Reproduction Number

By inspection, we see that the equilibrium equations with all disease states set to 0 yields the DFE

$$
X = Y = 0, \qquad V = P = \frac{W}{1 - \epsilon \theta^{-1}}, \qquad S = 1 - P. \tag{16}
$$

The Jacobian for this solution is

$$
J_{DFE} = \left[\begin{array}{cc|ccc}
-\Gamma(1 - \mathcal{R}_0 S) & \Gamma \mathcal{R}_0 S & 0 & 0 & 0 \\
\Gamma \alpha \xi \mathcal{R}_0 V & -\Gamma \alpha (1 - \xi \mathcal{R}_0 V) + \alpha a & 0 & 0 & 0 \\
\hline
-\mathcal{R}_0 S & -\mathcal{R}_0 S & -1 & 0 & -1 \\
-\sigma \mathcal{R}_0 V & -\sigma \mathcal{R}_0 V & 0 & -1 & 1 \\
-\mathcal{R}_0 P & -\mathcal{R}_0 P & 0 & 0 & -\Gamma \theta - 1
\end{array} \right], \tag{17}
$$

The lower right block of this Jacobian is triangular; therefore, its diagonals contain three of the eigenvalues, all negative. Stability is determined by the upper left block.

The upper left block also contains the information needed to determine the basic reproduction number using the next generation method [13], as it constitutes $\hat{F} - \hat{V}$, where

$$
\hat{F} = \begin{bmatrix} \Gamma \mathcal{R}_0 S & \Gamma \mathcal{R}_0 S \\ \Gamma \alpha \xi \mathcal{R}_0 V & \Gamma \alpha \xi \mathcal{R}_0 V \end{bmatrix}, \qquad \hat{V} = \begin{bmatrix} \Gamma & 0 \\ 0 & \alpha(\Gamma - a) \end{bmatrix}. \tag{18}
$$

These combine to make the next generation matrix

$$
\hat{F} \hat{V}^{-1} = \begin{bmatrix} \mathcal{R}_0 S & \alpha^{-1}(1 - \epsilon a)^{-1} \mathcal{R}_0 S \\ \alpha \xi \mathcal{R}_0 V & (1 - \epsilon a)^{-1} \xi \mathcal{R}_0 V \end{bmatrix}. \tag{19}
$$

The largest eigenvalue of this matrix is the basic reproduction number:

$$\mathcal{R}_v = \mathcal{R}_0 \left[S + \frac{\xi V}{1 - \epsilon a} \right] = \mathcal{R}_0 \left[1 - \frac{W}{1 - \epsilon \theta^{-1}} + \frac{\xi W}{(1 - \epsilon \theta^{-1})(1 - \epsilon a)} \right], \tag{20}$$

using (16) and the fact that the determinant of the matrix is 0, which makes the eigenvalues 0 and the trace of the matrix. In the limit as $\epsilon \to 0$, this formula reduces to the asymptotic approximation

$$\mathcal{R}_v \sim \mathcal{R}_0[1 - (1 - \xi)W] = \mathcal{R}_0(U + \xi W), \qquad U \equiv 1 - W. \tag{21}$$

This result was to be expected, as ξ is the relative disease burden through universal vaccination and W is the fraction of individuals who are willing to be vaccinated. The small fraction of willing susceptibles who do not make it into class V either because of prior infection or unrelated death accounts for the minor deviation of the exact result from the leading order approximation. As a practical matter, deviations from leading order can be ignored, as they are far less significant than the smallest errors in estimation of parameter values.

Stability of the DFE can be checked using the Routh-Hurwitz conditions for the submatrix in the upper left corner, which we call J_{12}, identifying this as the submatrix constructed from rows 1 and 2 and columns 1 and 2. The relevant quantities are the negative trace and the determinant, both of which must be positive; to leading order these are

$$c_1 = -\operatorname{tr}(J_{12}) = \Gamma[1 - \mathcal{R}_0 S + \alpha - \alpha \xi \mathcal{R}_0 V], \tag{22}$$

$$c_2 = \det(J_{12}) = \Gamma^2 \alpha[(1 - \mathcal{R}_0 S) - \xi \mathcal{R}_0 V] = \Gamma^2 \alpha(1 - \mathcal{R}_0 U - \xi \mathcal{R}_0 W) = \Gamma^2 \alpha(1 - \mathcal{R}_v). \tag{23}$$

As expected, the condition $\mathcal{R}_v < 1$ is necessary for DFE stability. It remains to show that it is also sufficient. To that end, we rearrange the inequality $c_2 > 0$ to obtain

$$-\xi \mathcal{R}_0 V > \mathcal{R}_0 S - 1$$

and substitute this into the formula for c_1 to obtain

$$\epsilon c_1 > 1 - \mathcal{R}_0 S + \alpha + \alpha(\mathcal{R}_0 S - 1)$$
$$= 1 + (\alpha - 1)\mathcal{R}_0 S > 0.$$

Retaining the $O(\epsilon)$ terms makes for a minor change in the calculation without changing the final result that $\mathcal{R}_v < 1$ is the binding constraint for stability.

3.2   The Endemic Disease Equilibrium (EDE)

In epidemiology models that lack feedback, as is the case here, we should expect there to be a unique and asymptotically stable EDE that exists only if $\mathcal{R}_v > 1$. Here we verify these expectations to leading order.

As a first step to finding the EDE, we use the Y' and X' equations from (12) to obtain equilibrium results

$$Y = SZ, \qquad X = \xi V Z + \epsilon a X = (1 + \epsilon a)\xi V Z + O(\epsilon^2).$$

The need to retain $O(\epsilon)$ terms at this stage will be seen in the stability analysis, where we need the key result obtained by substituting these equilibrium results into (13):

$$\mathcal{R}_0 S + \xi \mathcal{R}_0 V \sim 1 - \epsilon a \xi \mathcal{R}_0 V. \tag{24}$$

From here, leading order approximations in the calculation of the EDE will be sufficient.

Using the leading order approximation $P \sim W$, the S' and V' equilibrium equations become

$$S(1 + Z) = U, \qquad V(1 + \sigma Z) = W. \tag{25}$$

Combining the second of these equations with (24) (to leading order) to eliminate V yields

$$(1 - \mathcal{R}_0 S)(1 + \sigma Z) = \xi \mathcal{R}_0 V(1 + \sigma Z) = \xi \mathcal{R}_0 W = \mathcal{R}_v - \mathcal{R}_0 U.$$

Rearranging this equation yields

$$\mathcal{R}_0(U - S) + \sigma(1 - \mathcal{R}_0 S)Z = \mathcal{R}_v - 1.$$

Substituting from the first equation of (25) puts this result into the form

$$[\sigma + (1 - \sigma)\mathcal{R}_0 S]Z = \mathcal{R}_v - 1. \tag{26}$$

This result serves two purposes: It provides an explicit formula for Z once S is known, and it shows that $\mathcal{R}_v > 1$ is a necessary condition for the existence of an EDE.

We can now eliminate Z from (26) by first multiplying by S and then substituting again from (25). This yields a quadratic equation for S:

$$(U - S)[\sigma + (1 - \sigma)\mathcal{R}_0 S] = (\mathcal{R}_v - 1)S. \tag{27}$$

Because $\mathcal{R}_0 S \leq 1$, the quantity in the square brackets is also less than or equal to 1; hence,

$$U - S \geq (\mathcal{R}_v - 1)S,$$

from which we obtain the bound $S \leq \mathcal{R}_v^{-1} U$. Thus, there is a unique solution in the interval $0 \leq S \leq \mathcal{R}_v^{-1} U$ and so $\mathcal{R}_v > 1$ is sufficient as well as necessary for existence of a unique EDE.

Proof of the asymptotic stability of the EDE follows the outlines of a simplified form of the Routh-Hurwitz conditions, presented here.

Theorem 1. *Let J be the Jacobian matrix for a four-component dynamical system. Let $c_1 = -\operatorname{tr}(J)$, c_2 be the sum of all 2×2 subdeterminants of J, c_3 be the negative of the sum of all 3×3 subdeterminants of J, $c_4 = \det J$, and combine these quantities as*

$$q_1 = c_1 c_2 - c_3, \qquad \Delta_1 = c_2 - c_1, \qquad \Delta_2 = c_4 - c_3, \qquad \rho = c_3/c_1.$$

1. The associated equilibrium point is asymptotically stable if and only if

$$c_1 > 0, \qquad c_4 > 0, \qquad q_1 > 0, \qquad c_3 q_1 > c_1^2 c_4. \tag{28}$$

2. If $c_1 > 0$, $c_3 > 0$, and $c_4 > 0$, then $q_2 > 0$ is necessary and sufficient for stability, where

$$q_2 = c_1 c_3 \Delta_1 - c_3^2 - c_1^2 \Delta_2. \tag{29}$$

3. The condition $q_2 > 0$ can be rewritten as

$$\Delta_1 > \rho + \frac{\Delta_2}{\rho}. \tag{30}$$

Proof. Claim 1 is the standard form of the Routh-Hurwitz stability criteria in four variables [9]. The first part of claim 2 is easily seen from the fourth RH criterion, as $q_1 \leq 0$ with $c_3 q_1 > c_1^2 c_4$ is impossible when c_3 and c_4 are both positive. Defining $q_2 = c_3 q_1 - c_1^2 c_4$, we have

$$q_2 = c_1 c_2 c_3 - c_3^2 - c_1^2(c_3 + \Delta_2) = c_1 c_3 \Delta_1 - c_3^2 - c_1^2 \Delta_2,$$

as required. Replacing c_3 with ρc_1 reduces the criterion $q_2 > 0$ to (30). $\qquad \square$

As noted earlier, terms beyond the leading order as $\epsilon \to 0$ can be discounted, provided those leading order terms are not 0. Thus, apparently lower order terms can be safely ignored only after checking that the apparently larger terms are non-zero. Using the notation J_{ij} for the submatrix of J that uses rows i and j and columns i and j, it appears from the structure of the Jacobian that $|J_{12}|$ is $O(\Gamma^2)$, which dominates the other terms that comprise c_2. This is not correct, however, because of cancelation of dominant terms. Hence, we must calculate $|J_{12}|$ carefully.

We begin by introducing some simplifying notation into the Jacobian with

$$s = \mathcal{R}_0 S, \qquad v = \mathcal{R}_0 V, \qquad p = \mathcal{R}_0 P, \tag{31}$$

yielding the result

$$J_{EDE} = \begin{bmatrix} -\Gamma(1-s) & \Gamma s & \Gamma Z & 0 & 0 \\ \Gamma\alpha\xi v & -\Gamma\alpha(1-\xi v-\epsilon a) & 0 & \Gamma\alpha\xi Z & 0 \\ -s & -s & -(1+Z) & 0 & -1 \\ -\sigma v & -\sigma v & 0 & -(1+\sigma Z) & 1 \\ -p & -p & 0 & 0 & -\Gamma\theta-(1+Z) \end{bmatrix}, \tag{32}$$

Using (25), we can simplify the j_{11} and j_{22} entries of this matrix to

$$j_{11} = -\Gamma(1-s) = -\Gamma(1+\epsilon a)\xi v + O(\epsilon),$$

$$j_{22} = -\Gamma\alpha(1-\xi v-\epsilon a) = -\Gamma\alpha[s-\epsilon a(1-\xi v)] + O(\epsilon) = -\Gamma\alpha(1-\epsilon a)s + O(\epsilon).$$

From these formulas, we have

$$|J_{12}| = \Gamma^2\alpha\xi vs(1-\epsilon^2 a^2) - \Gamma^2\alpha\xi vs = O(1).$$

Thus, the term $|J_{12}|$, which appears at first to dominate the other 2×2 subdeterminants, actually makes no contribution at the leading order level of $O(\Gamma)$. With this detail in hand, we can now safely remove all terms that are subdominant in their own row, yielding the simplified matrix

$$J_{EDE} = \left[\begin{array}{cccc|c} -\Gamma\xi v & \Gamma s & \Gamma Z & 0 & 0 \\ \Gamma\alpha\xi v & -\Gamma\alpha s & 0 & \Gamma\alpha\xi Z & 0 \\ -s & -s & -(1+Z) & 0 & -1 \\ -\sigma v & -\sigma v & 0 & -(1+\sigma Z) & 1 \\ \hline 0 & 0 & 0 & 0 & -\Gamma\theta \end{array} \right], \tag{33}$$

We see from the block structure that the matrix decomposes into a 4×4 and a 1×1 block, so we need only consider the larger block, to which Theorem 1 applies. The quantity c_1 is the negative of the trace, which to leading order is

$$c_1 = \Gamma(\alpha s + \xi v).$$

Among the 3×3 subdeterminants, only $|J_{123}|$ and $|J_{124}|$ are of $O(\Gamma^2)$. Calculating these subdeterminants yields the results

$$J_{123} = -\Gamma^2\alpha s Z(s+\xi v) \sim -\Gamma^2\alpha s Z, \quad J_{124} = -\Gamma^2\alpha\sigma\xi vZ(s+\xi v) \sim -\Gamma^2\alpha\sigma\xi vZ.$$

Thus,

$$c_3 = -|J_{123}| - |J_{124}| = \Gamma^2 Z(\alpha s + \alpha\sigma\xi v). \tag{34}$$

Similarly, the determinant calculation for J_{1234} yields the result

$$c_4 = \Gamma^2 Z(\alpha s + \alpha\sigma\xi v + \alpha\sigma Z) = c_3 + \Gamma^2\alpha\sigma Z^2. \tag{35}$$

Using the relation $\alpha\sigma = 1$, these last two formulas simplify to

$$c_3 = \Gamma^2(\alpha s + \xi v)Z = \Gamma Z c_1, \qquad c_4 = c_3 + \Gamma^2 Z^2; \tag{36}$$

hence,

$$\rho = \Gamma Z, \qquad \Delta_2 = \Gamma^2 Z^2 = \rho^2. \tag{37}$$

The $c_4 = c_3 + \Gamma^2 Z^2$, so $\Delta_2 = \Gamma^2 Z^2$. Using the results for ρ and Δ_2, the stability criterion (30) reduces to

$$c_2 > c_1 + 2\Gamma Z. \tag{38}$$

Now examining the 2×2 subdeterminants, and using $\alpha\sigma = 1$, yields the results

$$|J_{13}| + |J_{24}| = c_1 + 2\Gamma Z, \qquad |J_{23}| + |J_{14}| > 0, \qquad |J_{12}| = o(\Gamma), \qquad |J_{34}| = o(\Gamma),$$

which proves that the stability criterion is satisfied. As is typical for straightforward epidemiology models, the EDE is stable whenever it exists, which is when $\mathcal{R}_v > 1$.

To summarize, we have established the needed analytical results for the EDE:

Theorem 2. *There exists a unique and asymptotically stable EDE for the problem* (12) *if and only if* $\mathcal{R}_v > 1$. *This solution is determined to leading order as* $\epsilon \to 0$ *by*

$$P = W, \quad Z = \frac{\mathcal{R}_v - 1}{\sigma + (1-\sigma)\mathcal{R}_0 S}, \quad V = \frac{W}{1 + \sigma Z}, \quad Y = SZ, \quad X = \xi V Z,$$

with S *the unique solution of*

$$(U - S)[\sigma + (1 - \sigma)\mathcal{R}_0 S] = (\mathcal{R}_v - 1)S, \qquad \mathcal{R}_v = \mathcal{R}_0(U + \xi W)$$

in the interval $0 \le S \le \mathcal{R}_v^{-1}U$. *The equilibrium rate of new infections is then*

$$F(W; \mathcal{R}_0, \xi, \sigma) = 1 - S - V.$$

It is interesting to note that the vaccination rate constant θ does not appear in either the formula for $\mathcal{R}_v$ or the EDE formulas; hence, adjusting this parameter in an attempt to account for vaccine unwillingness would make no difference.

4   Results
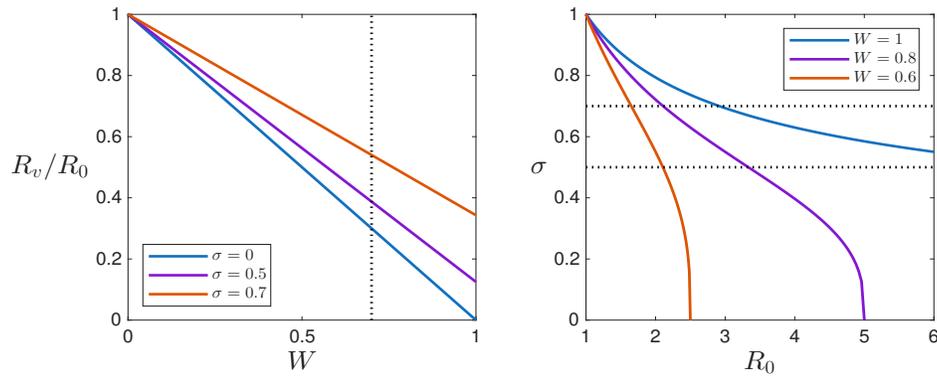
4.1   The Endemic Scenario

Figure 2: Left: Reduction factor from $\mathcal{R}_0$ to $\mathcal{R}_v$; Right: Regions in disease-vaccine space where the DFE is stable (below and left of the curves)

On the endemic time scale, our tasks were to determine the impact of vaccine unwillingness on the basic reproduction number (21) and the equilibrium new infection rate $F(W)$ given in Theorem 2. It is reasonable to reduce the number of parameters by assuming that all vaccination modifications are of equal value, so that $\xi = \sigma^3$. Figure 2 displays the results. The left panel of the figure shows the factor by which the basic reproduction number is reduced by vaccination as a function of the fraction of the population that is willing to be vaccinated. The three curves are for values of σ corresponding to a perfect vaccine, a moderate vaccine ($\sigma = 0.5$), and a marginal vaccine ($\sigma = 0.7$). At first it seems odd that the moderate curve is closer to

the perfect one than the marginal one; the reason for this is that the basic reproduction number is a linear function of the overall epidemiological disease burden of vaccinated individuals, which is $\xi = \sigma^3$, rather than a linear function of the infection reduction factor σ. The vertical dotted line marks $W = 0.7$, which is a reasonable estimate of vaccine willingness. Even with the minimal vaccine, the difference in outcome between $W = 1$ and $W = 0.7$ is large, with an $\mathcal{R}_v : \mathcal{R}_0$ ratio of 0.34 for the former and 0.54 for the latter.

The basic reproduction number itself is not the most important property of an endemic disease model; rather, it is the hypersurface in the parameter space that marks out the stability regions for the DFE and EDE. Several projections of this hypersurface onto the $\mathcal{R}_0\sigma$ plane are shown in the right panel of Figure 2, with three different values of willingness. The region below and to the left of each curve indicates the disease/vaccine combinations where the DFE is stable, given the level of willingness indicated in the legend. Thus, the region between the $W = 1$ and $W = 0.8$ curves shows disease/vaccine combinations for which the assumption of full willingness gives the wrong stability result if the actual willingness is only 80%. The dotted lines mark the approximate range of σ values for the seasonal flu vaccine [4]. Most other vaccines are far more effective, corresponding to lower σ. Even 10% unwillingness would significantly overestimate the impact of vaccination on disease eradication, particularly for a highly effective vaccine.
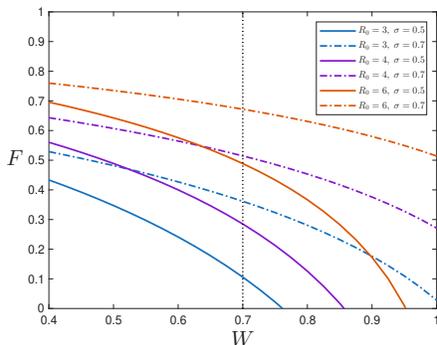

Figure 3: The rate of new infections at equilibrium

When the EDE is stable, the most important model outcome is the rate of new infections at equilibrium. This outcome is illustrated in Figure 3. The impact of assuming $W = 1$ when the actual willingness is $W = 0.7$ can be seen by comparing the values of F on the dotted vertical line with those at the far right of the plot. This impact is significant for a vaccine that is only marginally effective (the dash-dot curves), and grows quickly as the vaccine effectiveness increases. As in the case when the DFE is stable, neglecting vaccine unwillingness yields results that are too far from correct to be useful.

Instead of making the rate diagram more complicated to address vaccine unwillingness, couldn't we instead compensate for unwillingness by replacing the rate constant ϕ with the smaller constant $W\phi$ while retaining the simpler disease structure that assumes full willingness? For endemic scenarios, the answer is an emphatic "no." Changing the rate constant makes no difference to leading order in the endemic model, which is essentially the same as making no difference at all. While our study considers a particularly simple disease profile, there is no reason to think that this conclusion would be any different for a more complicated profile. Given that unwillingness is always a significant issue in any community, all endemic models with vaccination need to have unwillingness built into the rate diagram.

4.2 The Epidemic Scenario

It was surprising that the rate constant θ played no role in the endemic results to leading order. It should play a significant role at leading order in the epidemic model, however. The long-term outcome of the epidemic model is defined as (8), while the short term behavior can be defined as the graphs of $\widehat{I}$ and V. In both cases, our primary goal is to observe the difference made by attempting to account for unwillingness by deprecating the rate constant from θ to $W\theta$ rather than by adding an unwilling susceptible subclass. This comparison depends on the disease and vaccine characteristics $\mathcal{R}_0$ and σ.

Note that it would be a simple matter to compute sensitivity coefficients for various parameters to various outcomes; however, this is overkill given that the model is for a generic disease with parameters estimated

from general averages. Instead, it is more useful to judge impact as "noticeable" if one can see a difference in a graph and "significant" if the curves show a clear difference over a broad range of parameter sets.
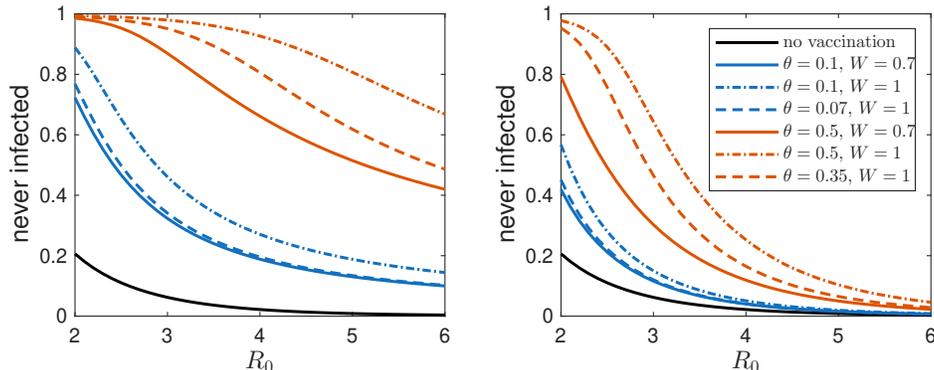

Figure 4: Population fractions that escape infection in the epidemic scenario, with $\sigma = 0$ (left) and $\sigma = 0.7$ (right). The nominal vaccination rate constants are $\theta = 0.1$ (blue) and $\theta = 0.5$ (orange). Solid curves have $W = 0.7$, and dash-dot curves have $W = 1$. Dashed curves are for the model that assumes full willingness but deprecates the vaccination rate constant by a factor of $W = 0.7$.

Figure 4 shows how the graph of the population fraction that avoids infection during the initial epidemic (8) vs $\mathcal{R}_0$ depends on a variety of factors. The left panel is for the perfect vaccine ($\sigma = 0$) and the right for a marginal vaccine ($\sigma = 0.7$). The bottom (black) curve is for the case of no vaccine, the middle group (blue) is for slow vaccination ($\theta = 0.1$), and the top group (orange) is for fast vaccination ($\theta = 0.5$). Within each group, the three curves represent three different assumptions about willingness. The lowest (solid) curve is for the full model (7) with willingness used to partition the susceptible class into subgroups. The top (dash-dot) curves are for the model without subclasses (9) and assuming full willingness. The middle (dashed) curves are the more nuanced simplification that uses the model without subclasses but attempts to correct for unwillingness by using $W\theta$ as the infection rate constant for the model.

Comparison of the solid and dash-dot curves shows the impact of vaccine unwillingness. Significantly more people become infected during the epidemic when 30% of people are unwilling to be vaccinated, with the effect larger for more effective vaccines and/or faster vaccination rate. These details are not surprising, as vaccination should have a greater impact when it is accomplished quickly and when it is more effective.

The overall impact of vaccine effectiveness is seen in the general comparison of the two panels in the figure. A perfect vaccine has a very large impact even if there is significant unwillingness and the vaccination rate is slow. However, the impact of a marginal vaccine, delivered at a slow rate, and with 30% unwillingness (the solid blue curve in the right panel) is still noticeable for diseases of moderate infectiousness ($\mathcal{R}_0 < 4$) and significant for any disease if the vaccination rate is large, even with 30% unwillingness. Of the factors that determine the public health consequences of vaccination, it appears from the figure that disease infectiousness and vaccine effectiveness are the most important, followed by vaccination rate and then unwillingness.

What then of the modeling strategy of accounting for unwillingness by deprecating the rate constant rather than using the more correct compartment diagram (solid curves vs dashed curves)? This depends on the vaccination rate. For slow vaccination (blue curves), the difference is noticeable, but not large. For fast vaccination (orange curves), the difference is always significant, regardless of vaccination effectiveness or disease infectiousness.

For a look at some specific simulation details, we examine graphs of the total infectiousness $\widehat{I}$ and vaccinated population fraction V for several scenarios, shown in Figure 5. The solid curves are for (7), while the dash-dot curves are for the simplified model with the S, S_W, and V equations replaced by (9). All scenarios use modest values of $\sigma = 0.5$ and $W = 0.7$. The left panels have $\mathcal{R}_0 = 3$, while the right panels have $\mathcal{R}_0 = 6$; the top panels have $\theta = 0.1$ and the bottom have $\theta = 0.5$.

Comparing the left and right panels shows the impact of the basic reproduction number. As expected, larger $\mathcal{R}_0$ results in a larger wave of infection. Larger $\mathcal{R}_0$ also decreases the size of the vaccinated population because a larger number of susceptibles become infected before they can be vaccinated. Comparing the
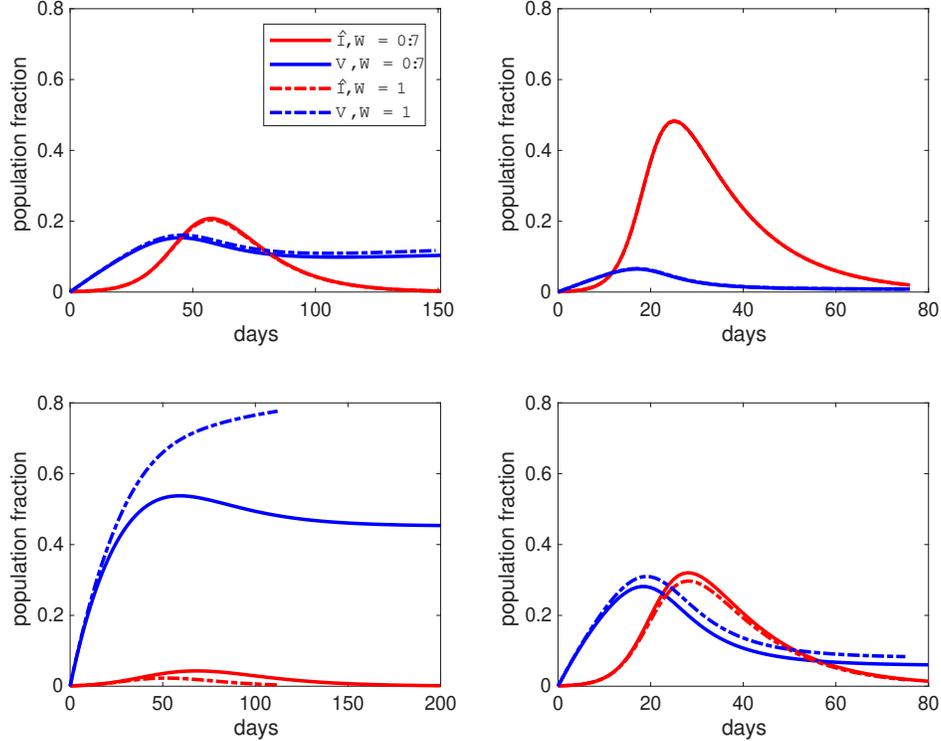
Figure 5: Total population infectiousness $\widehat{I}$ (red) and vaccinated fraction V (blue), with solid curves for the model (7) and dash-dot curves for the simplified model that accounts for unwillingness by decreasing the vaccination rate constant to $W\theta$ rather than modifying the standard SVIR compartment model; $\mathcal{R}_0 = 3$ (left panels), $\mathcal{R}_0 = 6$ (right panels), $\theta = 0.1$ (top panels), $\theta = 0.5$ (bottom panels); all panels have $\sigma = 0.5$ and $W = 0.7$.

top and bottom panels shows the impact of the vaccination rate constant. As expected, faster vaccination reduces the wave of infection in addition to increasing the vaccinated population. Of particular interest is the comparison between the two methods of accounting for unwillingness. When the vaccination rate is slow, there is little difference between accounting for unwillingness by modifying the rate diagram and accounting for it by lowering the vaccination rate constant. When vaccination is fast, there is error in both $\widehat{I}$ and V that is noticeable but not large, with the exception that the error in the size of the vaccinated fraction is large if $\mathcal{R}_0$ is small in addition to the vaccination rate being fast.

5   Discussion

This study presents a simple model for a generic disease with an available vaccine and no mortality. We set out to address the questions of how much error in key epidemiological outcomes is caused by ignoring vaccine unwillingness and the extent to which that error can be corrected by using vaccine unwillingess to deprecate the vaccination rate constant rather than to change the rate diagram for the disease. The results are quite different, depending on the time scale of interest.

For endemic models, whose purpose is to determine whether the final outcome is a stable disease-free equilibrium state or a stable endemic disease equilibrium state, the error caused by neglecting unwillingness is considerable and cannot be corrected by accounting for unwillingness in the vaccination rate constant. On the endemic time scale of demographics, vaccination occurs so rapidly that its exact rate does not affect the key outcomes. Nearly all susceptibles in the long-term scenario are unvaccinated, as willing susceptibles become vaccinated much more quickly than susceptible individuals of either willingness become infected.

The answers to the questions about the importance of correct modeling of vaccination are more subtle

in the epidemic time scale, where the large population of susceptibles means that a significant number of willing susceptibles become infected before they can be vaccinated. The epidemiological outcomes, and also the extent to which incorrect modeling matters, depend on the comparison between the rates by which susceptibles become infected, represented by $\mathcal{R}_0$, and by which willing susceptibles become vaccinated, represented by θ. If a disease is highly infectious and the vaccination rate relatively slow, then vaccination has only a limited impact on the epidemiological outcomes, and consequently the error of ignoring unwillingness is unimportant. In the opposite case of less infectiousness and faster vaccination, the error in neglecting unwillingness becomes significant.

From a modeling perspective, it is clear that correct incorporation of vaccine unwillingness into models intended to determine the effective basic reproduction number, stability of equilibrium outcomes, and ongoing disease burden for endemic disease settings is absolutely vital. For models intended to determine short-term outcomes, such as the population fraction that escapes infection in the initial wave, there is often no great harm in factoring unwillingness into the rate constant. However, the analysis of models on the epidemic time scale is necessarily done by numerical integration of the differential equations. While adding additional classes to an endemic model increases the complexity of the analysis, it makes no real difference to numerical integration whether we add a few additional variables. Hence, it should be preferable to correctly account for vaccine unwillingness for epidemic models as well as endemic models.

References

[1] Annenberg Public Policy Center. Vaccine confidence falls as belief in health misinformation grows. `https://www.annenbergpublicpolicycenter.org/vaccine-confidence-falls-as-belief-in-health-misinformation-grows/`, Nov 2023. Accessed: 2025-09-28.

[2] Eve Dubé, Caroline Laberge, Maryse Guay, Paul Bramadat, Réal Roy, and Julie Bettinger. Vaccine hesitancy. *Human Vaccines and Immunotherapeutics*, 9(8):1763–1773, 2013.

[3] Centers for Disease Control and Prevention. CDC immunization schedule adopts individual-based decision-making for COVID-19 and standalone vaccination for chickenpox in toddlers. `https://www.cdc.gov/media/releases/2025/cdc-immunization-schedule-adopts-individual-based-decision.html`, October 2025. Accessed: 2025-11-02.

[4] Centers for Disease Control and Prevention. CDC seasonal flu vaccine effectiveness studies. `https://www.cdc.gov/flu-vaccines-work/php/effectiveness-studies/index.html`, May 2025. Accessed: 2026-01-30.

[5] Yi Jiang, Kristin M. Kurianski, Jane HyoJin Lee, Yanping Ma, Daniel Cicala, and Glenn Ledder. Incorporating changeable attitudes toward vaccination into compartment models for infectious diseases. *Math. Biosci. Eng.*, 22(2):260–289, 2025.

[6] R Ke, E Romero-Severson, S Sanche, and N. Hengartner. Estimating the reproductive number $\mathcal{R}_0$ of sars-cov-2 in the united states and eight european countries and implications for vaccination. *Journal of Theoretical Biology*, 517, 2021.

[7] Glenn Ledder. Scaling for dynamical systems in biology. *Bull. Math. Biol.*, 79:2747–2772, 2017.

[8] Glenn Ledder. Incorporating mass vaccination into compartment models for infectious diseases. *Math. Biosci. Eng.*, 19(9):9457–9480, 2022.

[9] Glenn Ledder. Using asymptotics for efficient stability determination in epidemiological models. *Math. Biosci. Eng.*, 2025.

[10] PBS NewsHour / Associated Press. RFK Jr. pulls funding for vaccines being developed to fight respiratory viruses. `https://www.pbs.org/newshour/health/rfk-jr-pulls-funding-for-vaccines-being-developed-to-fight-respiratory-viruses`, August 2025. Accessed: 2025-11-02.

[11] Sheryl Gay Stolberg and Christina Jewett. Kennedy advises new parents to "do your own research" on vaccines. `https://www.nytimes.com/2025/04/29/us/politics/kennedy-vaccines-research.html`, April 2025. Accessed: 2025-10-28.

[12] U.S. Department of Health and Human Services. HHS winds down mRNA vaccine development under BARDA. `https://www.hhs.gov/press-room/hhs-winds-down-mrna-development-under-barda.html`, August 2025. Accessed: 2025-11-02.

[13] Paulina van den Driessche and James Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.*, 180:29–48, 2002.