5 Vaccination

Vaccines don't save lives — vaccinations do.
– ID Epi Proverb

Broadly speaking, the immune system as a whole provides protection through two important kinds of systems. The **innate immune system**, also called the nonspecific immune system, provides a set of responses to infection that are not particular to any pathogen, but are instead a broad set of defensive responses. You can roughly think of the innate immune system as "pre-programmed."

In contrast, the **adaptive immune system** consists of highly specific responses to particular pathogens. The adaptive immune system is not pre-programmed, but instead uses lymphocytes, B-cells, T-cells and antibodies to learn strategies to combat infections by particular pathogens. Because immunity is acquired over time, the adaptive immune system may also be referred to as the acquired immune system. **Antigens** are substances that elicit an adaptive immune response. Immune responses by either system may be classified as **humoral** or **cellular**. Cellular immunity is mediated by cells, including cells you may have heard of such as T-cells and macrophages. Humoral immunity is mediated by non-cellular macromolecules such as antibodies and complements.

The goal of vaccination is to create an adaptive immune response in the body, such that the body acquires protection against infection, disease, or both. In earlier eras, vaccination was accomplished by **inoculation** in which the microbe or virus is implanted into the body with the goal of stimulating an adaptive immune response. A good example of inoculation was the intentional infection with cowpox to create an immunity to smallpox. Today, vaccination is accomplished by introducing the pathogen's antigens more directly, without introducing a competent pathogen. The various vaccines for COVID-19 include injections of a killed virus, of antigen-presenting nanoparticles, and of mRNA that causes the short-term production of antigens. In all cases, the adaptive immune system learns how to respond to antigens, and thus, immunity is acquired.

Vaccinology and immunology are extremely complicated and fascinating. However, for the purposes of this course, we're going to focus on the primary goal of vaccination: to reduce the burden of disease. Depending on the pathogen, disease, and vaccine, the impact of vaccination may take place via different mechanisms, including (i) preventing infection, (ii) preventing transmission, and/or (iii) preventing progression of the infection to disease. Consequently, our models for vaccination, as with all our models, will have to be explicit about the assumptions made about the vaccine's impacts and mechanism.

5.1 Model: a perfectly protective vaccine

Suppose that a vaccinated person is perfectly immune to infection. In a compartmental modeling framework, this means that we will create a fourth compartment V. Any individual in V will be unable to become infected. The simplest version of this model is one in which individuals are vaccinated (or not) outside of the dynamics of an epidemic. In other words, vaccination is treated as an initial condition. A flow diagram for this model is shown in Figure 1.

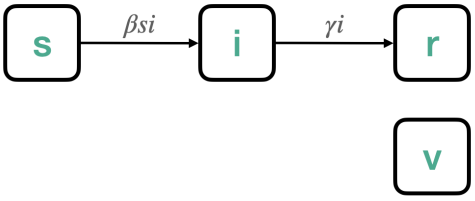

Figure 1: A flow diagram for the SIR+V model with a perfectly protective vaccine, and vaccination as an initial condition.

Notice that in this model, we are assuming that those who are vaccinated are unable to participate in the dynamics at all. One cannot become vaccinated or become unvaccinated through any means. As a result, our typical SIR model equations are identical to our very first models, but with two modifications. First, our total population is now $N = S + I + R + V$, and thus $1 = s + i + r + v$. Second, our dynamics for V are given simply by $\dot{V} = 0$, and thus $\dot{v} = 0$.

These equations help us to understand that, when a vaccine provides perfect protection against infection, the vaccinated compartment participates in the dynamics only insofar as vaccinated individuals are *not* in the other compartments. One way of stating this is to observe that every person who is vaccinated in a person who is not susceptible. As a result, one way to ensure that $s \leq s_{\text{peak}}$ is to ensure that $v \geq 1 - s_{\text{peak}}$, leading us to a statement about vaccination and herd immunity. An epidemic cannot take place in this model (i.e., $\dot{I} < 0$) provided that

$$v \geq 1 - \frac{1}{R_0} \,. \tag{1}$$

Put in plain words, if enough people are vaccinated, there aren't enough susceptibles for an epidemic to take off.

5.2 Model: a partially protective vaccine

No vaccine is perfect in its protection. As a result, most vaccines will have an associated efficacy and/or effectiveness, VE, typically verbalized as a percentage between 0 and 100, and written mathematically with $0 \leq VE \leq 1$. When speaking technically, **vaccine efficacy** refers to a vaccine's protective effects from a controlled clinical trial, while **vaccine effectiveness** refers to an estimate of its protective effects in the field under real-world conditions.[1]

To model a vaccine with a given VE, we're going to introduce two models, called the "All or Nothing" model and "Leaky" model. In the way that they model an imperfect vaccine, it is important that we emphasize the these are models of reality, and not reality itself.

[1] Why might vaccine effectiveness differ from vaccine efficacy? Knowingly creating a false comparison, which is a better measure?

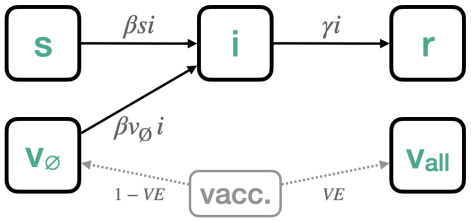

Figure 2: A flow diagram for the SIR+V model with an All-or-Nothing vaccine, and vaccination as an initial condition. The light grey arrows show the proportions by which initial vaccinations are distributed between v_{all} and $v_{\varnothing}$, where we use the $\varnothing$ symbol to denote "nothing."

5.2.1 The All-or-Nothing Model

If a vaccine has a given $VE < 1$, the All-or-Nothing model (ANM) makes the following assumption: vaccination provides perfect protection with probability VE, and provides no protection with probability $1 - VE$. In other words, after vaccination, you either get all the protection or none of the protection; thus, the name. A flow diagram for this model is shown in Figure 2.

Modeling an imperfect vaccine with an ANM assumption results in the following equations,

$$\begin{aligned} \dot{s} &= -\beta s i \\ \dot{i} &= \beta s i + \beta v_{\varnothing} i - \gamma i \\ \dot{r} &= \gamma i \\ \dot{v}_{\varnothing} &= -\beta v_{\varnothing} i \\ \dot{v}_{\text{all}} &= 0 \end{aligned} \tag{2}$$

where the new terms are highlighted in red. Note that VE is nowhere to be found. Instead, it is included *only* as an initial condition! Note that the equations for $\dot{s}$ and $v_{\varnothing}$ are structurally identical, which means that, in effect, "susceptibles" and "nothings" are indistinguishable in their effect on the dynamics. Put differently, for every 100 people who are vaccinated, $(1 - VE) \times 100$ remain effectively susceptible.

How many people must now be vaccinated to achieve herd immunity? From Eq. (2) we can see that

$$\dot{i} \leq 0 \qquad \text{when} \qquad s + v_{\varnothing} \leq \frac{1}{R_0} \, .$$

Suppose that a fraction of the population v is initially vaccinated and the remaining $1 - v$ are susceptible. This implies that $v_{\varnothing} = v(1 - VE)$. Substituting into the equation above, we get

$$v\,VE \geq 1 - \frac{1}{R_0} \qquad \text{or} \qquad v \geq \frac{1}{VE}\left(1 - \frac{1}{R_0}\right) \qquad \text{or} \qquad VE \geq \frac{1}{v}\left(1 - \frac{1}{R_0}\right) \, .$$

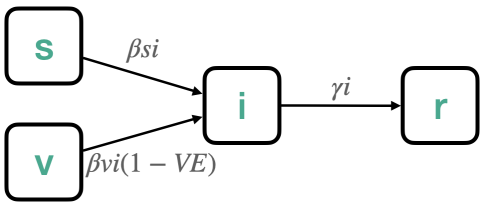

Figure 3: A flow diagram for the SIR+V model with a Leaky vaccine, and vaccination as an initial condition.

Note that each of these inequalities, while mathematically identical, can help us to answer a different question. On the left, we can see clearly that the proportion of the population who is *effectively vaccinated* is simply $v\,VE$. The middle equation can help us answer questions of the form "How many people must be vaccinated to achieve herd immunity if the vaccine has effectiveness VE?" The right equation can help us answer questions of the form "How good does our vaccine need to be if uptake will be only v?"

5.2.2 The Leaky Model

If a vaccine has a given $VE < 1$, the Leaky model (LM) makes the following assumption: vaccination provides everyone with partial protection VE such that the probability of transmission upon exposure decreases from p to $p\,VE$. In other words, after vaccination, everyone gets the same level of protection, whose effect is to provide an imperfect (leaky) barrier to infection. A flow diagram for this model is shown in Figure 3.

Modeling an imperfect vaccine with an LM assumption results in the following equations,

$$\begin{aligned}
\dot{s} &= -\beta si \\
\dot{i} &= \beta si + \beta vi(1-VE) - \gamma i \\
\dot{r} &= \gamma i \\
\dot{v} &= -\beta vi(1-VE)
\end{aligned} \tag{3}$$

where the new terms are again highlighted in red. This time, we can clearly see VE in the equations, as VE affects the rate at which exposure causes vaccinated people to become infected. One interpretation of our new equations is that vaccination causes β to be replaced by $\beta(1-VE)$.

How many people must now be vaccinated to achieve herd immunity? From Eq. (3) we can see that

$$\dot{i} \leq 0 \qquad \text{when} \qquad s + v(1-VE) \leq \frac{1}{R_0}\,.$$

Once more assuming a scenario where everyone is either susceptible or vaccinated (i.e. $s = 1 - v$), we can

see that the conditions for herd immunity are

$$v\,VE \geq 1 - \frac{1}{R_0} \qquad \text{or} \qquad v \geq \frac{1}{VE}\left(1 - \frac{1}{R_0}\right) \qquad \text{or} \qquad VE \geq \frac{1}{v}\left(1 - \frac{1}{R_0}\right) . \tag{4}$$

Your eyes do not deceive you. Equation (4) is identical to what we developed above, meaning that the herd immunity threshold with vaccination does not depend on whether we consider ANM or LM conceptualizations of vaccination!

5.2.3 All-or-Nothing vs Leaky

We have two fundamentally different conceptualizations of an imperfect vaccine (ANM and LM) and yet they both have the same herd immunity threshold. We can also see that they are structurally different. Where might we observe differences in practice?

The most important difference is that the ANM ensures that those who gain protection are protected *no matter what*. They can never be infected, even if the force of infection is extremely high. In contrast, the LM provides no such protection—it attenuates the force of infection, but does not eliminate it. Thus, for high R_0 values and moderate VE values, the ANM will show better effective protection than the LM.[2]

5.3 Model: a three-factor vaccine

The ANM and LM are simple in the sense that the effect of the vaccine is to reduce the risk of infection. However, many vaccines have measurable effects on various endpoints other than infection.[3] These endpoints include mortality, severe disease, hospitalization, symptomatic disease, transmission potential, and more. In other words: a vaccine's effects may be complicated! How can we model a more complex vaccine?

For our purposes, we will consider three important effects that a vaccine may have:

1. A vaccine may reduce one's risk of becoming infected, given that one is exposed. We will term this effectiveness as VE_s.
2. A vaccine may reduce one's risk of transmitting the virus onward, given that one is infected. We will term this effectiveness as VE_i
3. A vaccine may reduce one's risk of a particular disease outcome (e.g. symptoms, severe disease, death), given that one is infected. We will term this effectiveness as VE_p.

[2] Can you think of a way to show this mathematically?

[3] In epidemiology, an **endpoint** is an event or outcome that can be measured objectively to determine whether an intervention (here, a vaccine) is beneficial.

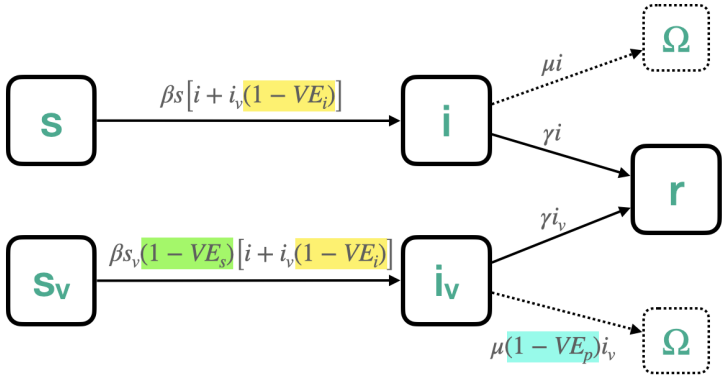

Figure 4: A flow diagram for the SIR+V model with a three-factor vaccine, and vaccination as an initial condition. Small v subscripts represent vaccinated versions of people who are otherwise susceptible and infected, respectively. The compartment Ω represents death. Terms with the three factors of vaccine effectiveness are highlighted.

In other words, our vaccine's effects will now be written out separately for how it affects susceptibility (VE_s), infectiousness (VE_i), and disease progression (VE_p). Note that we will need to include a disease outcome in the model, and so we introduce the state Ω for death. A flow diagram for this model is shown in Figure 4. Modeling an imperfect vaccine with this type of three-factor model results in the following equations, which are broken into three sets of two with new terms highlighted.

$$\begin{aligned}
\dot{s} &= -\beta s\left[i + (1 - VE_i)i_v\right] \\
\dot{i} &= \beta s\left[i + (1 - VE_i)i_v\right] - \left[\gamma + \mu\right] i \\
\\
\dot{s}_v &= -\beta s_v(1 - VE_s)\left[i + (1 - VE_i)i_v\right] \\
\dot{i}_v &= \beta s_v(1 - VE_s)\left[i + (1 - VE_i)i_v\right] - \left[\gamma + \mu(1 - VE_p)\right] i_v \\
\\
\dot{r} &= \gamma(i + i_v) \\
\dot{\Omega} &= \mu i + \mu(1 - VE_p)i_v
\end{aligned} \tag{5}$$

It can be shown[4] that the herd immunity threshold is when the fraction vaccinated v satisfies

$$v\left(1 - \frac{\gamma + \mu}{\gamma + \mu(1 - VE_p)}(1 - VE_s)(1 - VE_i)\right) \geq 1 - \frac{1}{R_0}\,, \tag{6}$$

[4]We'll have the required machinery to do this in a few weeks!

where $R_0 = \frac{\beta}{\gamma+\mu}$ due to the mortality μ associated with the disease.

What can we learn from this equation? First of all, the right hand size is identical to our equations for the ANM and LM calculations, but the left hand side appears different. However, the left hand side bears many similarities to our previous left-hand side with the correct amount of squinting. For example, suppose for simplicity that the vaccine has no additional protective effects against disease progression. In this scenario, the RHS simplifies considerably,

$$v\left(1 - \frac{\gamma+\mu}{\gamma+\mu(1-VE_p)}(1-VE_s)(1-VE_i)\right)\Bigg|_{VE_p=0} \quad = \quad v\Big(1-(1-VE_s)(1-VE_i)\Big)\,.$$

Directly comparing this expression with Eq. (4) asks us to compare the LM/ANM's VE with the three-factor model's $1-(1-VE_s)(1-VE_i)$. Subtracting one from each yields the following comparison

$$1 - VE \sim (1-VE_s)(1-VE_i)\,.$$

This expression has a satisfying interpretation: whereas a LM/ANM vaccine attenuates the transmission process by a single rate $1-VE$ which decreases susceptibility to infection, the three-factor model attenuates the transmission process by decreasing both susceptibility to infection by $(1-VE_s)$ *and* potency of infectiousness by $(1-VE_i)$.

The three-factor vaccine model is more complex, and yet it is useful in particular because its three endpoints have the possibility of being measured in observational studies. As a result, it is possible to anchor the parameters of this model in real data for a variety of infectious diseases.[5]

5.4 Recovering the simpler models

Whenever we make a model more complex by introducing a new parameter, it is valuable to perform a check: if we set the new parameter to zero, do we recover the less complex model? In that spirit, note the following:

First, if we set $VE_p = 0$ and $VE_i = 0$, then we reduce the three-factor model to a one-factor model. In fact, this one-factor model is the leaky model, exactly! Next, if we set $VE = 1$ in the leaky model *or* in the all-or-nothing model, those who are vaccinated are perfectly protected, and we recover the perfect vaccine model, exactly! Finally, if we set $VE = 0$ in the leaky model *or* in the all-or-nothing model, all the protections of the vaccine disappear, and we recover the basic SIR model, exactly!

More complex models of protection and immunity are possible, but the four models introduced here are sufficiently sophisticated that they continue to be used in a wide variety of theoretical and policy applications.

[5] Among the three factors, which do you think is the easiest to estimate from data? Which is the hardest? Why? Hint: think about trying to design these studies yourself, and ask what you would have to measure in each!

5.5 What good are these models?

So far, we've developed the SIR model, and then bolted on various modification, including (i) a return to susceptibility, (ii) birth and death, and (iii) four models of vaccination. In each case we primarily focused on equilibrium solutions, their stability, R_0, epidemic final size, and herd immunity via vaccination or recovery. Especially for our three-factor vaccine model, one could fairly ask what good we derive from this kind of complex model.

One class of questions has to do with timing of vaccination. For instance, in a study of the seasonal human mobility patterns in Niger and their association with measles, Bharti et al discuss the relative timing of targeting vaccination campaigns into areas where population density is increasing (driving increases in β and thus the potential for transmission). This work is particularly important when we think about not only seasonal patterns, but also large unexpected migrations including refugee crises.

Another class of questions has to do with vaccine prioritization, i.e. asking to whom vaccinations should be made available when supply is limited. For instance, in a study of various potential rollouts of COVID-19 vaccines, Bubar et al analyze various strategies for prioritization, including strategies that aggressively pursue herd immunity by prioritizing those who drive transmission, as well as strategies that directly prioritize the most vulnerable older adults to protect them from disease. This work is particularly important when we think about ensuring that policy is rooted in rigorous modeling and theory, while also directly confronting all the uncertainties associated with VE values, rollout, model misspecification, and parameter misspecification.[6]

These classes of questions are not yet possible to address with the Lecture content thus far because they require that we include forms of heterogeneity. One form of heterogeneity is seasonality which changes a model's parameters over time (Bharti et al). Another form of heterogeneity is population structure with different age classes that interact in complex ways (Bubar et al). These heterogeneities naturally allow us to ask questions about tradeoffs, because now we can ask *when* or *whom* to vaccinate. Without heterogeneity, there is no choice, and without choice the policy questions are less interesting. Stay tuned!

[6]Here we use misspecification in a colloquial way. In statistical analyses including large regression models, *statistical model misspecification* refers to including or failing to include relevant variables, i.e. the model doesn't account for everything it should, and/or it does account for things it shouldn't. For our purposes, model misspecification means using a model whose structure (equations) are too rough an approximation for reality, while parameter misspecification means that the numbers we use for our parameters aren't right. This is the grownup version of the tired quotation from George Box, that "All models are wrong. Some are useful." Perhaps we could say that "All models are misspecified. Some are useful." A good model and clever modeler are able to answer their questions using the simplest yet least misspecified available model—a difficult and artful multi-objective optimization.