AMSI VACATION RESEARCH SCHOLARSHIPS 2019–20

EXPLORE THE MATHEMATICAL SCIENCES THIS SUMMER


# Ethical Considerations in Vaccine Hesitancy

Daniel Roberts

Supervised by Dr. Joel C. Miller

La Trobe University

Vacation Research Scholarships are funded jointly by the Department of Education and Training and the Australian Mathematical Sciences Institute.

**Abstract**

In recent years, there has been an increase in anti-vaccination sentiment in many parts of the world, which has raised concerns about the potential for the resurgence of many vaccine-preventable diseases due to lowered immunization rates (Dube, Vivion, and MacDonald 2015). These low immunization rates will not only put vaccine refusers at risk, but also other members of the community who do not have immunity due to vaccine failure or inability to obtain vaccination. This raises an ethical question about the acceptability of refusing vaccination (Jamrozik, Handfield, and Selgelid 2016). To investigate this, we construct a model to determine the risks of refusing vaccination, both to those who refuse, and to the people to whom they might transmit diseases. It was determined that these risks vary highly depending on the importation rate and effective reproduction number of the disease. Probabilities of infection as high as 3% are possible even when the effective reproduction number is less than 1, and the expected number of further cases caused by someone's decision not to vaccinate could be as high as 0.3.

## 1 Introduction

Recently, mathematical models have become increasingly significant for studying infectious diseases. These include many deterministic differential equation-based models, such as Bartsch et al. (2016), as well as stochastic models (Ball et al. 2019), and network-based models (Fransson and Trapman 2019). They have mostly focused on population-scale disease dynamics to inform disease control strategies.

One of the most successful disease control strategies available is vaccination. Mathematical modelling has often shown that it is among the most effective ways of reducing the prevalence of infectious diseases (Harris et al. 2016; Bartsch et al. 2016; Penny et al. 2016). However, some people still refuse vaccination. This raises an ethical issue, as, in so doing, these people put not only themselves, but also others, at risk, since they are more likely to become infected and spread the disease. There are always some members of a community who cannot be vaccinated, including infants, immunosuppressed individuals, and those who have encountered barriers to vaccine access. These people rely on the vaccination of others for protection from diseases.

There is a significant amount of literature on the ethics of vaccination, mostly focused on ethical vaccine policy (Hendrix et al. 2016; Adewale et al. 2019). Some also considers what should be considered ethically responsible behaviour for an individual (Jamrozik, Handfield, and Selgelid 2016; Pierik 2018; MacDonald et al. 2018). Is it ethically acceptable to refuse vaccination when this will endanger other people? To effectively answer this question, we need to know the level of risk involved, both to the vaccine refuser and to those around him.

We investigated the risks of refusing vaccination, both to oneself and to others, using a simple model of local disease spread. Sections 2 and 3 focus on diseases that have been eliminated from a community, and occur only as a result of importations from outside. (These would mostly occur when someone who was infected in another location travels to the community.) In Section 4, this model is adjusted so that it can be used in epidemics.

### 1.1 Modelling small disease outbreaks with probability generating functions.

Investigations into individual risks requires knowledge of small-scale disease transmission among individuals, rather than population-scale modelling. Small-scale transmission can be modelled using probability generating functions (PGFs).

We construct the model by assuming that each individual has a fixed set of other individuals that we call her *offspring*. If she becomes sick, she will definitely transmit the disease to all these offspring, and to no-one else. We assume that the number of offspring each individual has follows a distribution, known as the *offspring distribution*, based on the disease characteristics seen in previous outbreaks. In reality, a person's offspring do not exist until he becomes infected, but it can be seen that either option gives the same result.

It is often best to work with the PGF of the offspring distribution, $\mu(x)$, which is defined as:

$$\mu(x) = \sum_m r_m x^m \tag{1}$$

where $r_m$ is the probability of transmitting the disease to $m$ people.

In this model, *transmission sets*, (see figure 1), will be used extensively to understand disease spread. Someone's transmission set consists of all the people who, if he became infected, he would pass the disease on to, either directly or via other people.

Four theorems related to offspring PGFs are important in this investigation. They are all proved by Miller (2018).

Theorem 1.1 concerns the *effective reproduction number*, the average number of people one infected person transmits to.

**Theorem 1.1** *The effective reproduction number, $R_e$, of a disease is given by:*

$$R_e = \sum_m m r_m = \mu'(1) \tag{2}$$

Theorem 1.2 is used extensively when considering distributions of the number of descendants in generations beyond the first.

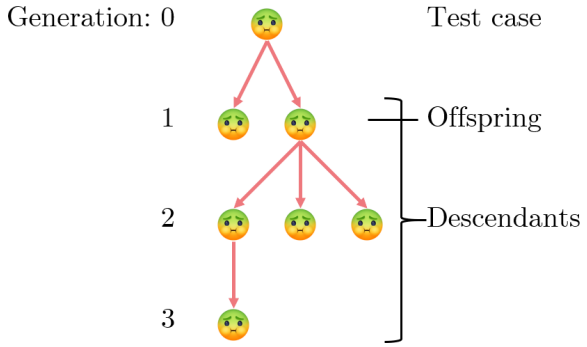



Figure 1: A transmission set showing all the *descendants* (people transmitted to, either directly or via someone else) of a particular person. The *generation* of a descendant refers to how many transmissions there are between the test case and this person.

**Theorem 1.2** *Let $f(x)$ and $g(x)$ be PGFs. If we take a number from the distribution given by $f(x)$ and then find the sum of that many numbers from the distribution given by $g(x)$, the result follows the distribution given by the PGF $f(g(x))$.*

Theorem 1.3 is used primarily to estimate the distribution of total outbreak sizes.

**Theorem 1.3** *Let $\mu(x)$ be the PGF for the offspring distribution. $\Omega_g(x)$, the PGF for the distribution of the total number of cases in the first $g$ generations, is given by the iterative formula:*

$$\Omega_g(x) = x\mu(\Omega_{g-1}(x)), \qquad \Omega_0(x) = x \tag{3}$$

It is possible to use this iterative formula directly to estimate $\Omega_1(x)$, the PGF for the total transmission set size, but it is computationally easier to use theorem 1.4.

**Theorem 1.4** *Let $\Omega_1(x) = \sum_m \omega_m x^m$, i.e. let $\omega_m$ be the probability that an outbreak contains $m$ cases. It follows that, for any $R \in (0,1]$:*

$$\omega_m = \int_0^1 \frac{\Omega_m(Re^{2\pi iu})}{R^n e^{2n\pi iu}} du \qquad \frac{1}{N}\sum_{n=0}^{N-1} \frac{\Omega_m(Re^{2\pi in/N})}{R^n e^{2n\pi in/N}} \tag{4}$$

*where $N$ is sufficiently large.*

$\Omega_m(Re^{2\pi in/N})$ can be calculated recursively using Theorem 1.3.

## Statement of Authorship

The workload for this investigation was divided as follows:

Daniel Roberts performed the mathematical investigations that gave all results except where stated otherwise, with guidance from Dr. Joel C. Miller. He also developed the Python code used to perform all calculations and wrote this report.

Dr. Joel Miller supervised the project, assisted with and guided the mathematical investigations and development of code, and proofread this report. He also had previously developed much of the theory used in this investigation.

## 2 Risk to oneself

We are interested first in the risk that any particular vaccine refuser, whom we will refer to as the test person, poses to herself by her decision not to vaccinate. This risk is considered the probability that, given she is not vaccinated, she is infected with a particular disease. It will depend on two factors, how effectively the disease transmits within a community, i.e. the effective reproduction number, and how often it enters the community from an external source, which we will call the *importation rate*. Since the community is modelled as an infinite population, the importation rate is the probability that any unvaccinated individual is infected from somewhere outside the community.

### 2.1 Susceptibility sets

This model uses *susceptibility sets*. The test person's susceptibility set consists of all the people who, if they become sick, will eventually transmit the disease to her, directly or via others. (See figure 2.) Again, this should be considered from the assumption that the transmissions that each individual causes are fixed, as outlined in section 1.1.

Let the PGF for the distribution of the number of parents (*parent distribution*) be $\varepsilon(x)$. Since every transmission has a parent and an offspring, there are the same total number of parents as offspring, and therefore the same expected number of parents per person as offspring. This expected number is the effective reproduction number, $R_e$.

For most diseases, transmissions to the test person will be independent of each other, so the presence of any particular parent should have no effect on the probability that the test person has other parents. Or, mathematically:

$$P(\text{At least 2 parents} | \text{At least 1 parent}) = P(\text{At least 1 parent}) \tag{5}$$

The probability distribution that has this property is the Poisson distribution, which, for an

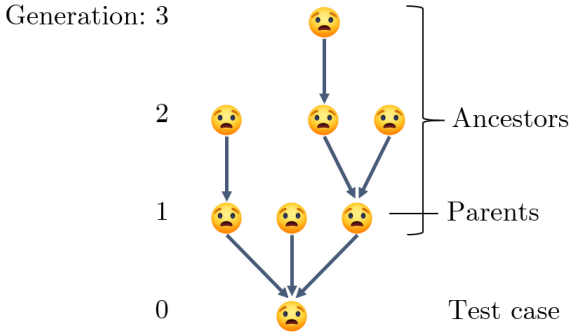



Figure 2: The test person's susceptibility set is all the people who will pass on the disease to her if they become sick, either directly or through a number of transmissions. The individuals who will transmit directly to her are called her *parents*, and anyone above her in the susceptibility set is her *ancestor*. Note that the test person is included in the transmission set.

expected value of $R_e$, has the PGF:

$$\varepsilon(x) = e^{R_e(x \quad 1)} \tag{6}$$

Theorems 1.3 and 1.4 can be used to estimate the PGF for the distribution of the total susceptibility set size, which we will call $\Theta_1(x)$. However, Miller (2018) showed that the probability that a transmission set is of size $m$ when the offspring follow a Poisson distribution with effective reproduction number $R_e$ is $\frac{(mR_e)^{m \quad 1}}{m!e^{mR_e}}$. The susceptibility set distribution is mathematically identical to this, so the PGF for the susceptibility set is:

$$\Theta_1(x) = \sum_m \theta_m x^m = \sum_m \frac{(mR_e)^{m \quad 1}}{m!e^{mR_e}} x^m \tag{7}$$

## 2.2 The total probability of becoming infected

Using this PGF and the importation rate, $p$, the probability that the test person becomes infected can be calculated. Firstly, if the test person has a susceptibility set of size $m$, the probability that she does not become infected is $(1 \quad p)^m$, i.e. the probability that no-one in her susceptibility set becomes infected. However, we have no knowledge of the size of the test person's susceptibility set. The probability that her susceptibility set is of size $m$ AND she does not become infected is $\theta_m(1 \quad p)^m$. Her total probability of avoiding infection, which we will call $s$, is therefore the sum of these probabilities over all values of $m$:

$$s = \sum_m \theta_m(1 \quad p)^m = \Theta_1(1 \quad p) \tag{8}$$

So, we can state that:

**Theorem 2.1** *If the PGF for the distribution of susceptibility set sizes is $\Theta_1(x)$ and the importation rate is $p$, the probability that the test person is infected, $p_T$, is given by:*

$$p_T = 1 \quad s = 1 \quad \Theta_1(1 \quad p) \tag{9}$$

It is worth noting that, since we assume that the number of parents is Poisson-distributed, complete knowledge of the offspring distribution is not needed to perform these calculations. Only the effective reproduction number is needed.

### 2.3 Methods of calculation

There are three ways to calculate $s$. Firstly, we can calculate $\Theta_1$, using either Theorems 1.3 and 1.4, or equation 7, and then use this to calculate $s$ directly. Unfortunately, using Theorems 1.3 and 1.4 is computationally expensive, and using equation 7 often creates overflow errors due to the large factorials.

The second method, which is both reliable and relatively fast, is to estimate $s = \Theta_1(1 \quad p)$ by iterating the condition given in Theorem 1.3, with $x = 1 \quad p$. We have:

$$\Theta_g(x) = x\mu(\Theta_{g \ 1}(x)), \qquad \Theta_0(x) = x$$

This translates to the equation for $s$:

$$s_g = (1 \quad p)\mu(s_{g \ 1}), \qquad s_0 = 1 \quad p$$

This method can be represented with the cobweb diagram in figure 3. It can be shown that the upper bound of the error using this process is given by theorem 2.2. This theorem is based on the use of the secant shown in figure 3, and we rigorously prove it in Appendix A.1.

**Theorem 2.2** *If $s$ is estimated by some value $s_g$ that is calculated by iterating equation 20, the error is bounded by:*

$$\mathit{Error}(s) < \frac{(s_g \quad s_{g \ 1})^2}{s_{g \ 2} \quad 2s_{g \ 1} + s_g} \tag{10}$$

The third method of calculating $s$ is to make use of the fact that, since we can estimate $s$ using the iteration $s_g = (1 \quad p)\varepsilon(s_{g \ 1})$, we know that $s = (1 \quad p)\varepsilon(s) \ =) \ (1 \quad p)\varepsilon(s) \quad s = 0$. This equation can now be solved using a more rapidly converging iterative technique, such as Newton's method. It is proven in Appendix A.2 that, by iterating from an initial estimate $s_0 = 0$, the equation will always converge to the correct value when using Newton's method.

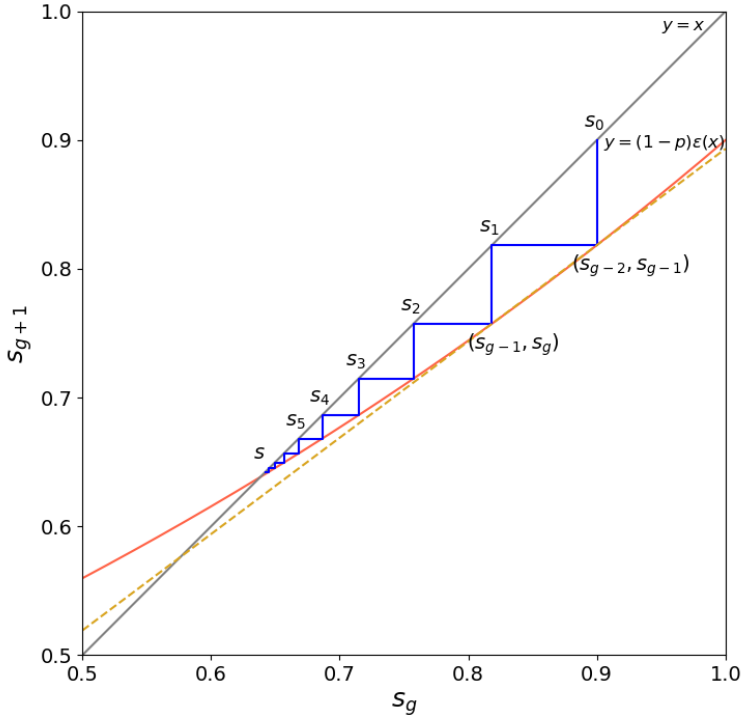



Figure 3: A cobweb diagram. Moving between $y = x$ and $y = (1 \quad p)\varepsilon(x)$ visually demonstrates the iterative process used to find $s$. The secant shown allows for the calculation of an error bound for $s_2$.

There is no clear error bound for this calculation method, but it can easily be determined that the error is less than $\delta$ if $(1 \quad p)\varepsilon(s_n \quad \delta) \quad s_n \quad \delta < 0$ or $s_n + \delta > 1$, which is also shown in Appendix A.2.

Hence, the most efficient method of calculating $s$ is using the recurrence relation that Newton's method provides, which is:

$$s_{n+1} = s_n + \frac{f(s_n)}{f^{\emptyset}(s_n)} = \frac{(R_e s_n \quad 1)(1 \quad p)e^{R_e(s_n \quad 1)}}{R_e(1 \quad p)e^{R_e(s_n \quad 1)} \quad 1} \tag{11}$$

## 3 Risk to others

We consider the risk imposed on other people by considering the distribution of the number of people who would have been saved had the test person chosen to vaccinate. In order to accurately determine this, some adjustments need to be made to the transmission set. Consider the transmissions show in figure 4. X and Y are individuals who become sick from separate introductions. The transmission set of X includes all individuals A-F. C and F, however, would still have become sick even if X had been vaccinated, as Y would have transmitted to them in a separate outbreak. So, we must consider the *e ective transmission set*. This is the set of all people who would not have been infected, had the test person chosen to vaccinate. It is highlighted in yellow.

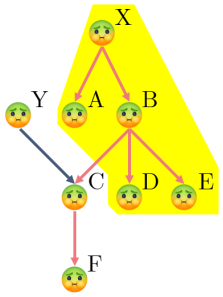



Figure 4: A demonstration of the difference between the transmission set (everyone who X transmits to, directly or via others) and the effective transmission set (the people who would have been saved had X been vaccinated).

Of course, if the test person does not become sick, their effective transmission set is 0. For the moment, we will assume that she is infected, and then make adjustments at the end when this distribution is used to calculate risks.

### 3.1 The effective transmission set size distribution.

Firstly, we will discuss the *e ective o spring distribution*, the distribution of the number of offspring who would not have become infected had the test person vaccinated. Due to the Poisson nature of the parent distribution, the fact that someone has one parent has no influence on the probability that he has other parents. So, we can state that:

$$P(\text{Received a second transmission} \textit{j} \text{Became infected}) = P(\text{Became infected}) \tag{12}$$

We calculated the probability of not becoming infected in section 2 and called it $s$.

We can now consider a very simple PGF, $1 \quad s + sx$. This is effectively the number of cases a particular offspring adds to the effective offspring. If this person would not have been infected otherwise, which happens with probability $s$, this offspring results in 1 person, namely himself, being added to the effective offspring. Otherwise, with probability $1 \quad s$, he would have become sick anyway, so adds no-one to the effective offspring.

By composing this equation with $\mu(x)$ (theorem 1.2), we obtain the PGF for the effective offspring distribution, $S(x)$. So we have:

**Theorem 3.1** *If a disease has an o spring distribution PGF $\mu(x)$, and there is a probability $s$ that each of these o spring received no transmissions from anyone else, $S(x)$, the distribution of the number*

*of o spring that would have been saved had the test person vaccinated, is given by:*

$$S(x) = \mu(1 \quad s + sx) \tag{13}$$

With no specific knowledge about the nature of the offspring distribution, the only way to calculate the PGF for the effective transmission set size, which we will call $\Psi_1(x)$, is using the iterative process outlined by theorems 1.3 and 1.4. In specific situations, this may not be necessary, however, due to the existence of analytic formulas for the transmission set PGFs of many common offspring distributions (Miller 2018).

### 3.2 Adjustments to the effective transmission set PGF to reflect the risk to others.

Once we have the PGF for the effective transmission set size, some more adjustments are needed to obtain something that is meaningful for determining the risks to others. Firstly, since we are interested in the risks that the test person poses to other people, not to herself, we need to remove her from the transmission set. So, the total number of other people whose infection the test person is responsible for has a distribution with PGF $\Psi_A(x) = \frac{\Psi_1(x)}{x}$ (Dividing by $x$ reduces every exponent, which represents the number of cases, by 1.)

We must then adjust the PGF to reflect the fact that there is a probability $s$ that the test person never becomes infected, so never transmits to anyone. So, with probability $s$, she infects no-one, and with probability $1 \quad s$, she infects a number of people following a distribution with PGF $\Psi_A(x)$. So, we can state:

**Theorem 3.2** *The number of cases caused by the test person follows a distribution:*

$$\Psi_B(x) = s + (1 \quad s)\frac{\Psi_1(x)}{x} \tag{14}$$

*where $\Psi_1(x)$ is the PGF of the transmission set size distribution.*

### 3.3 The expected number of people saved by a decision to vaccinate

Often, the full distribution of the number of people saved by the test person's decision to vaccinate is not needed. Only the expected value of this distribution is relevant. When this is the case, there are easier calculation methods. The expected number of effective offspring is $R_e s$. So the expected number of *e ective descendants*, i.e. descendants that would have been saved had the test person vaccinated, in any particular generation $g$ is $(R_e s)^g$. To prove this, consider that each of the $R_e s$ (average) effective infections in the first generation is expected to cause $R_e s$ effective infections in the

second generation, giving an expected total of $(R_e s)^2$ infections in generation 2, and so on. So, the expected total number of infections, by making use of the geometric series, can be shown to be:

$$E(\Psi_1(x)) = \sum_{m=0}^{1} (R_e s)^m = \frac{1}{1 \quad R_e s} \tag{15}$$

This same method can be used to show that the expected value of the transmission and susceptibility sets is $\frac{1}{1 \quad R_e}$, as shown by (Miller 2018).

The expected number of other people who would have been saved had the test person vaccinated, given she became infected, is one less than the expected total effective transmission set size, so:

$$E(\Psi_A(x)) = \frac{1}{1 \quad R_e s} \quad 1 = \frac{R_e s}{1 \quad R_e s} \tag{16}$$

Then, if we do not yet know the that test person will become infected, the expected number of people who would have been saved had she vaccinated is this value multiplied by the probability that she becomes infected:

$$E(\Psi_B(x)) = \frac{(1 \quad s) R_e s}{1 \quad R_e s} \tag{17}$$

Again, it is worth noting that this expected value depends only on the effective reproduction number and importation rate, not on the full offspring distribution, $\mu(x)$. However, $\mu(x)$ is necessary to calculate $\Psi_B(x)$.

## 4 Adjustments for the epidemic model

If a disease has an effective reproduction number greater than 1, outbreaks will not always die out. They can become epidemics. In this case, the outbreak will not die out until a high enough proportion of the community is immune that the effective reproduction number drops below 1.

Since we model the population as infinite, the proportion who are immune in our community does not change, and $R_e$ is constant. This means that the epidemic will never die out. So, there is a very clear difference in this model between a small outbreak that dies out stochastically and an epidemic.

In an infinite population where each individual has a probability $p$ of importing the disease from outside, there will be infinitely many importations. This means that, if $R_e$ is greater than 1, so that every outbreak has some probability of causing an epidemic, an epidemic will certainly occur.

The only exception to this is if the importation rate is 0. In this case, it is possible that the disease is not present, and with no importation, it never will be, making a trivial case.

However, having an importation rate of 0 does not necessarily mean that the disease is not present. It could already be present in the community, and be spreading as an epidemic. In this case, we are still interested in the risks when the importation rate is 0.

If $R_e > 1$, some transmission sets and some susceptibility sets are infinite. Any individual who has an infinite susceptibility set will certainly become sick, as someone in that infinite set will surely be infected. Those who have a finite susceptibility set may or may not become infected, based on whether anyone in this set imports the disease.

### 4.1 Risk to oneself from an epidemic disease

Let $\eta$ be the probability that the test person's susceptibility set is finite. Then, the susceptibility set size distribution is given by the PGF:

$$\Theta_1(x) = \sum_m \theta_m x^m + (1 \quad \eta)x^1 \tag{18}$$

$x^1$ is considered to be 0 when $x$ 2 $[0, 1)$ and 1 when $x = 1$.

Since everyone with an infinite susceptibility set is infected, we can calculate $s$, the probability of not catching the disease, in a similar way to the non-epidemic case:

$$\begin{aligned} s &= \sum_m \theta_m (1 \quad p)^m \\ &= \Theta_1(1 \quad p) \qquad \text{if} \quad 1 \quad p \not= 1 \quad ) \quad p \not= 0 \end{aligned}$$

If $p = 0$, then everyone with an infinite susceptibility set becomes sick, and everyone with a finite susceptibility set does not become sick, so $s$ is the probability of having a finite susceptibility set, or $\eta$.

According to Miller (2018), $\eta$ satisfies the equation:

$$\eta = \varepsilon(\eta) \quad ) \quad \varepsilon(\eta) \quad \eta = 0 \tag{19}$$

If $p \not= 0$, all the same methods can be used to calculate $s$ as in the non-epidemic case, since the same condition applies that $s = \Theta_1(1 \quad p)$. However, when $p = 0$, using $\Theta_1$ directly will simply return $s = \Theta_1(1) = 1$, rather than $s = \eta$. The fixed point iteration would begin with $s_0 = 1$, and the iteration would become $s_g = \varepsilon(s_{g \ 1})$. $\varepsilon(1) = 1$, so this method would also return $s = 1$. So, when $p = 0$, these methods calculate the probability of infection when the disease is not present.

As discussed in Appendix A.2, however, Newton's method does find the correct value $s = \eta$, giving this method another distinct advantage over the others.

### 4.2 Risk to others from an epidemic disease

The risk that the test person poses to other people by refusing vaccination for an epidemic disease can be calculated in exactly the same way as a non-epidemic disease, using the effective offspring distribution $S(x) = \mu(1 \quad s + sx)$. Confirming that this works, however, requires that we prove that the effective transmission set is never infinite, so that cases of infinite risk do not arise.

The expected number of effective offspring is $E(S) = R_e s$. For a given $R_e$, the largest value of $s$ possible is found when $p = 0$. In this case, $s = \eta$. Hence, if we prove that $R_e\eta < 1$ for all $R_e$, it follows that the expected number of effective offspring is always less than 1.

The value of $\eta$ depends on $\varepsilon(x) = e^{R_e(x \quad 1)}$. It also satisfies equation 19, so:

$$\eta = e^{R_e(\eta \quad 1)} \implies R_e = \frac{\ln \eta}{\eta \quad 1}$$

Substituting this value of $R_e$ into $E(S) = R_e\eta$ gives:

$$E(S) = \frac{\eta \ln \eta}{\eta \quad 1}$$

Let $g(\eta) = \frac{\eta \ln \eta}{\eta \quad 1}$, $0 < \eta < 1$. Note that this is a continuous function. We can use L'Hôpital's rule to find the limit of $g(\eta)$ as $\eta \to 1$:

$$\lim_{\eta \to 1} g(\eta) = \lim_{\eta \to 1} \frac{\frac{d}{d\eta}(\eta \ln \eta)}{\frac{d}{d\eta}(\eta \quad 1)} = \lim_{\eta \to 1} \frac{\ln \eta + 1}{1} = 1$$

Differentiating $g(\eta)$ gives:

$$g'(\eta) = \frac{\eta \quad \ln \eta \quad 1}{(\eta \quad 1)^2} = \frac{1}{\eta \quad 1} \quad \frac{\ln \eta}{(\eta \quad 1)^2}$$

Hence, $g(\eta)$ is differentiable on its entire domain. By converting to a Taylor series, we obtain:

$$\begin{aligned}
g'(\eta) &= \frac{1}{\eta \quad 1} \quad \frac{\sum_{k=1}^{1} \frac{(\quad 1)^{k+1}}{k}(\eta \quad 1)^k}{(\eta \quad 1)^2} \\
&= \frac{1}{\eta \quad 1} \quad \sum_{k=1}^{1} \frac{(\quad 1)^{k+1}}{k}(\eta \quad 1)^{k \quad 2} \\
&= \sum_{k=2}^{1} \frac{(\quad 1)^k}{k}(\eta \quad 1)^{k \quad 2} \\
&= \sum_{k=0}^{1} \frac{(\quad 1)^k}{k+2}(\eta \quad 1)^k \\
&= \sum_{k=0}^{1} \frac{1}{k+2}(1 \quad \eta)^k
\end{aligned}$$

Since $0 < \eta < 1$, $0 < 1 \quad \eta < 1$.

So,

$$\frac{1}{k+2}(1 \quad \eta)^k > 0 \qquad \text{for all } k \in \mathbb{N} \setminus \{0\}$$

Hence,

$$g'(\eta) = \sum_{k=0}^{1} \frac{1}{k+2}(1 \quad \eta)^k > 0$$

So, the continuous and differentiable function $g(\eta)$ is an increasing function that approaches 1 as $\eta \to 1$. This means that $g(\eta)$ must be less than 1 for $0 < \eta < 1$. Hence:

$$E(S) = \eta R_e = \frac{\eta \ln \eta}{\eta \quad 1} < 1 \qquad \text{for all } \eta \in (0,1)$$

And hence we can say:

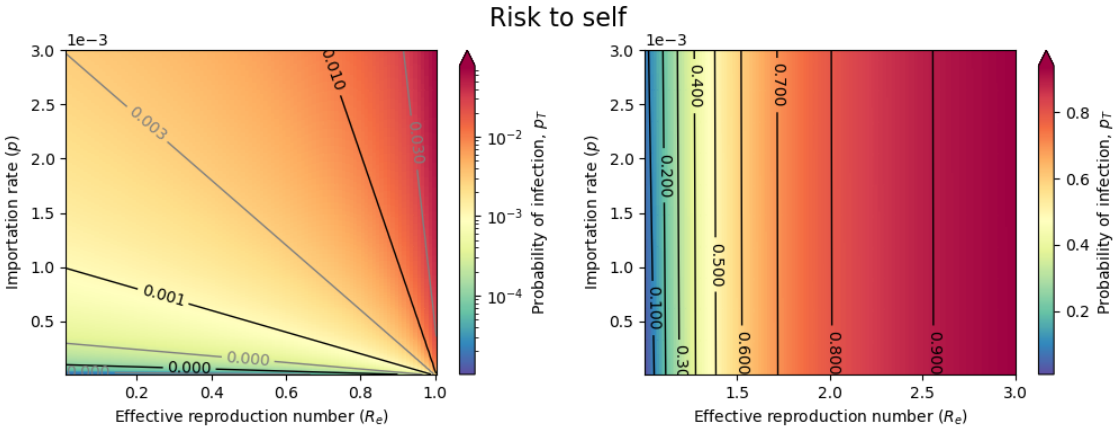
**Theorem 4.1** *If there is an epidemic of a particular disease, the e ective transmission set of every person is nite. Hence, no individual is responsible for an in nite number of other cases.*

With this proof, we have shown that the calculation methods used in the non-epidemic case apply to epidemics as well.

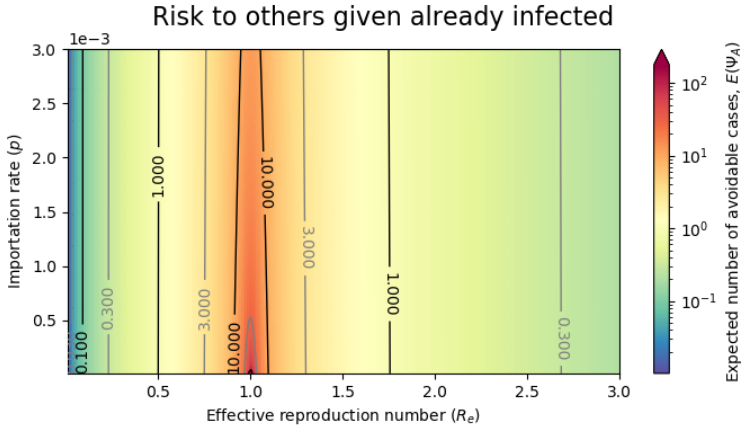
## 5 Results

Code was developed to perform these calculations and visualise the data in Python. Figure 5 shows the risks for importation rates up to 3 in 1000 and effective reproduction numbers up to 3. This importation rate range was selected because measles, one of the diseases of most interest to vaccination ethics research, appears to have an importation rate to Australia of roughly 1.6 in 1000 (i.e. the probability of any unvaccinated individual importing the disease at some point in their lifetime is $1.6 \quad 10^{\ 3}$), based on the number of cases imported in 2009-11 (Chiew et al. 2013), and it was expected that other diseases of interest would be in a similar range. At these importation rates, the probability of infection and the expected number of infections that would have been prevented by a not-yet-infected person's decision to vaccinate vary roughly logarithmically when the effective reproduction number is below 1, but roughly linearly above 1, so these graphs have been split into these two domains, so that the results can be seen clearly.

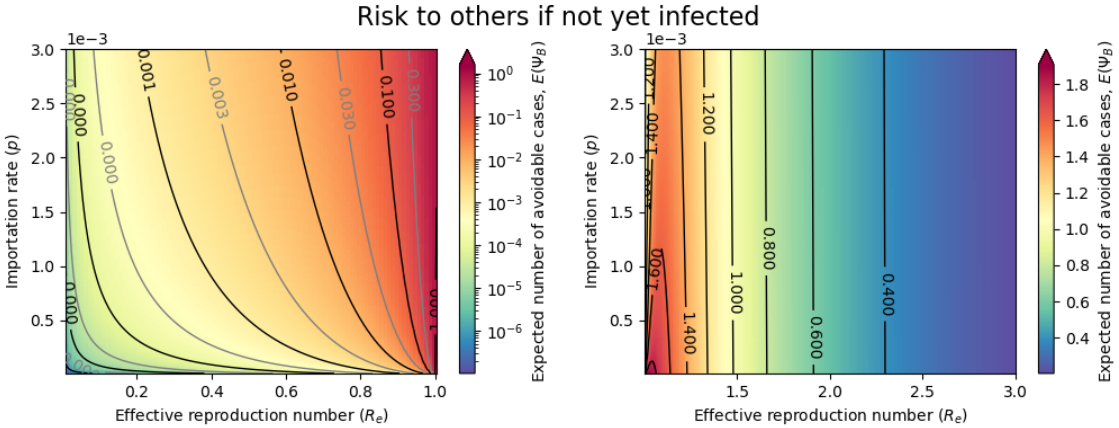
Appendix B shows the results over the full range of importation rates from 0 to 1.




(a) The probability of becoming infected for importation rates relevant to actual diseases.




(b) The number of further infections that could have been avoided, had someone who became sick been vaccinated.




(c) The expected number of other cases that would have been avoided, had someone chosen to be vaccinated, with no foreknowledge about whether he becomes sick.

Figure 5: The risks of refusing vaccination for realistic importation rates.

## 6 Discussion

Application of this model to ethics will require some adjustments. Firstly, the risk that not vaccinating imposes on other vaccine refusers should be ignored, as these people are capable of taking actions to avoid this risk. This adjustment can be made by multiplying $E(\Psi_B)$ by the proportion of the unvaccinated population that are able to vaccinate. The entire distribution of the number of people who cannot be vaccinated who are infected due to a vaccine refuser's actions is given by $\Psi_B(1 \quad h+hx)$, where $h$ is the proportion of unvaccinated individuals that cannot be vaccinated.

The large range of risk values from 1 in 1000 to 0.9 (risk to self) or 1 in 10000 to 1.6 (risk to others) has two important effects. Firstly, this model has to be applied to each disease separately. Secondly, accurate knowledge of the effective reproduction number and the importation rate of a disease are also important. These can be difficult to obtain, and can fluctuate considerably with time. For example, estimates of the effective reproduction number of measles in Australia range from 0.38-0.78 (for 2009-11) (Chiew et al. 2013) to 1.7 (for 2012-13) (Gidding et al. 2018), meaning that, with an importation rate of 3 in 1000, the probability of becoming infected could be anywhere between 0.002 and almost 0.7, and the expected number of others saved by a decision to vaccinate could range from 0.001 to about 1.5. This does not even consider the fluctuations that occur in these values over time. However, even with limited data, this work could be used to provide a likely risk range, which may still have uses due to the knowledge that the risks are above or below a certain threshold considered relevant to a particular policy measure.

Note that this model only applies directly for one person. If an infection passes through two people before infecting a third person, vaccination of either of the first two could save the third, but vaccination of both creates no further benefits, which acts to reduce the risk posed by the two together. On the other hand, two individuals could be part of two different outbreaks that both would have infected a particular individual. If both of these people are vaccinated, an extra person is saved that would not have been saved had either one of them been vaccinated. So, this increases the risk posed by the two people. Either one of these could dominate at any particular value of $R_e$ or $p$, so further research into this effect would be useful.

## 7 Conclusion

By considering the probability distributions of the sizes of transmission and susceptibility sets, a method for determining the risks of refusing vaccination, both to an individual person and to those

around her, has been created. This method is applicable both to eliminated diseases and epidemics. The risk to self is considered to be the probability that someone becomes infected if he refuses vaccination, and the risk to others is considered the expected number of people whose infection would have been avoided had some individual non-vaccinator chosen to vaccinate instead. The risk, both to self and others, varies by 3-4 orders of magnitude across the ranges of importation rates and effective reproduction numbers considered, meaning that accurate knowledge of the importation rates and effective reproduction numbers of a disease is important in order to assess the risks of not vaccinating.

# Appendices

## A Proofs

### A.1 Theorem 2.2

Consider $s_g$ to be the $g^{\text{th}}$ estimate of $s$. It is equal to $\Theta_g(1 \quad p)$, where $\Theta_g$ is the PGF for the number of people in the first $g$ generations of the test person's susceptibility set, equivalent to $\Omega_g$ for the transmission set. $\Theta_g$ is calculated in the same way as $\Omega_g$, recursively using the formula:

$$\Theta_g = x\mu(\Theta_{g \quad 1}(x))$$

We begin with $s_0 = 1 \quad p$, so that the formula becomes:

$$s_g = (1 \quad p)\varepsilon(s_{g \quad 1}), \qquad s_0 = 1 \quad p \tag{20}$$

Figure 3 is reproduced here to aid with this proof (figure 6). In this case, we consider estimating the point of intersection between $y = (1 \quad p)\varepsilon(x)$ and $y = x$, by beginning at the point $(1 \quad p, 1 \quad p)$ and repeatedly moving vertically to $y = (1 \quad p)\varepsilon(x)$ and then horizontally to $y = x$. The probability of avoiding infection by importation from outside, $1 \quad p$, will always be equal to or greater than the probability of avoiding infection altogether, $s$. So, $1 \quad p \quad s$. Hence, the cobweb diagram always begins at or above $s$, the intersection point. As $\varepsilon(x)$ is an increasing function, there is no point where the cobweb diagram will cross to the other side of the intersection point, so $s_g \quad s$ for all $g$ 2 N.

For any two consecutive points where the cobweb line meets $y = (1 \quad p)\varepsilon(x)$ (i.e. two points $(s_{g \quad 1}, s_g)$ and $(s_{g \quad 2}, s_{g \quad 1})$, for any $g$ 2 N), a straight line drawn through these points will always intersect $y = x$ at a value lower than $s$. This is because $(1 \quad p)\varepsilon(x)$ is always concave-up, so will

This assumes that the gradient of the new line is less than 1. When $R_e < 1$, this is a valid assumption, as the gradient

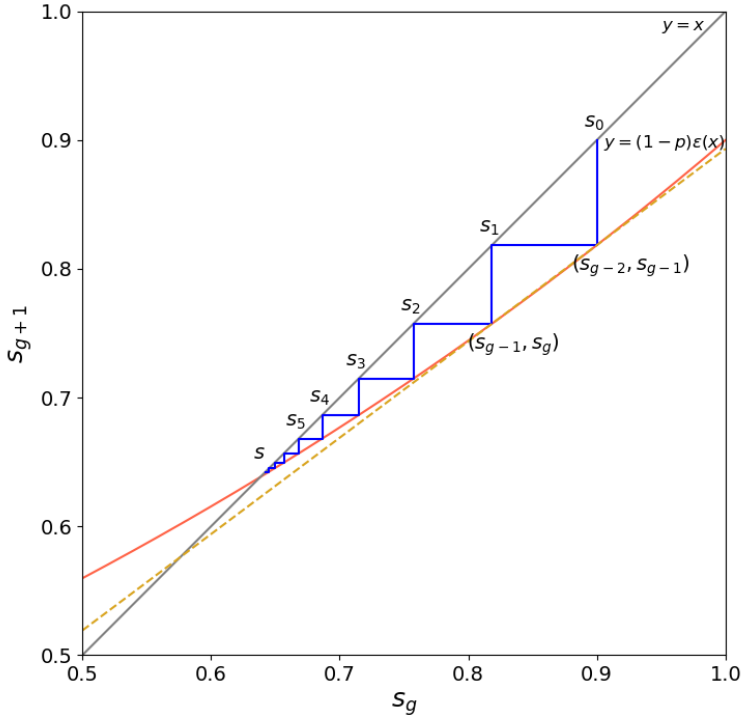



Figure 6: We reproduce figure 3, to assist in this proof.

curve away above the straight line, intersecting $y = x$ at a higher value, as can be seen in figure 3. Hence, we can use the difference between $s_g$ and this intersection point as an error bound for $s_g$.

Some simple calculations show that this distance is given by $\frac{(s_g \quad s_g \quad 1)^2}{s_g \quad 2 \quad 2s_g \quad 1+s_g}$, which concludes the proof.

## A.2 Correct convergence and error bound properties of Newton's method

As before, let $f(x) = (1 \quad p)\varepsilon(x) \quad x$, so that $s$ is a root of $f$. Figure 7 graphically shows the use of Newton's method to calculate $s$ using $f(x)$. We can state that:

$$f(x) = (1 \quad p)e^{R_e(x \quad 1)} \quad x$$
$$f^{\emptyset}(x) = R_e(1 \quad p)e^{R_e(x \quad 1)} \quad 1$$
$$f^{\emptyset\emptyset}(x) = R_e^2(1 \quad p)e^{R_e(x \quad 1)}$$

of $(1 \quad p)\varepsilon(x)$ is less than 1 for $x$ 2 $[0, 1]$, because the gradient increases across this range to a value of $(1 \quad p)R_e < 1$. If $R_e > 1$, the condition that the gradient of the line is less than 1 must be included for this claim to be valid. However, there is always a point close to the intersection point where this holds, as the line $y = (1 \quad p)\varepsilon(x)$ is below $y = x$ when $x > s$, meaning that, in order to approach and intersect at $s$, there must be a range of values directly above $s$ for which the gradient is less than 1.

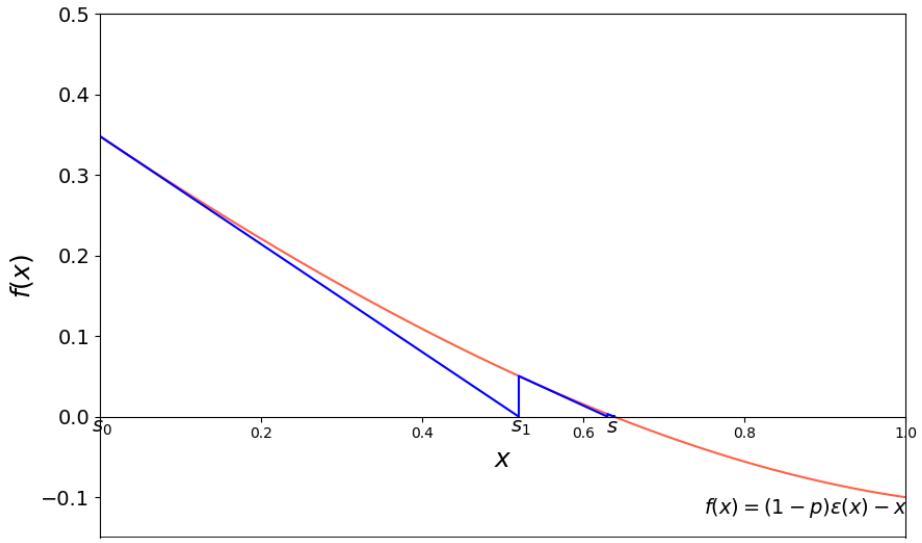



Figure 7: A demonstration of the use of Newton's method to find $s$, using an effective reproduction number of 0.95 and an importation rate of 0.9.

For the moment, we will ignore the cases $p = 0$ and $p = 1$. So, we can state that $0 < p < 1$. We can then determine a few characteristics of $f(x)$ from the equations above:

$f(0) = (1 \quad p)e^{\ R_e} > 0$

$f(1) = \quad p < 0$

$R_e < e^{R_e} = \sum_{n=0}^{1} \frac{R_e^n}{n!} =) \quad R_e e^{\ R_e} < 1$. Also, $1 \quad p < 1$. So, $f^{0}(0) = R_e(1 \quad p)e^{\ R_e} \quad 1 < 0$.

$f^{00}(x) > 0$ for all $x$ 2 R, so $f(x)$ is concave up over its entire domain.

Because $f(0)$ is positive and $f(1)$ is negative, there will always be an odd number of roots on $[0, 1]$, and since $f(x)$ is concave up, there cannot be more than 2 roots in total. So, there must be exactly 1 root on $[0, 1]$, which must therefore correspond to $s$. Since we know that $f(0) > 0$ and that the gradient is negative at 0, by beginning with $s_0 = 0$, we guarantee that the first tangent will give a positive root, i.e. $s_1 > 0$. However, since $f(x)$ is concave up, it will pull upwards away from the tangent line, and will therefore intersect the $x$-axis at a larger value than the tangent (see figure 7). So, $0 < s_1 < s$, meaning that $s_1$ is a better approximation of $s$ than $s_0 = 0$. As long as $s_n < s$, $f(s_n) > 0$, and $f^{0}(s_n) < 0$, as the function has an increasing gradient (concave up) and has a negative gradient at $s$, where it drops below the $x$-axis, meaning that, for any $x$-value less than $s$, the gradient will be negative. Hence, the same conditions apply as for $s_0 = 0$, meaning that $s_n < s_{n+1} < s$. Hence, it can

be seen that the iteration is guaranteed to approach $s$ if we begin with $s_0 = 0$, since each iteration gets closer to it, and cannot pass it.

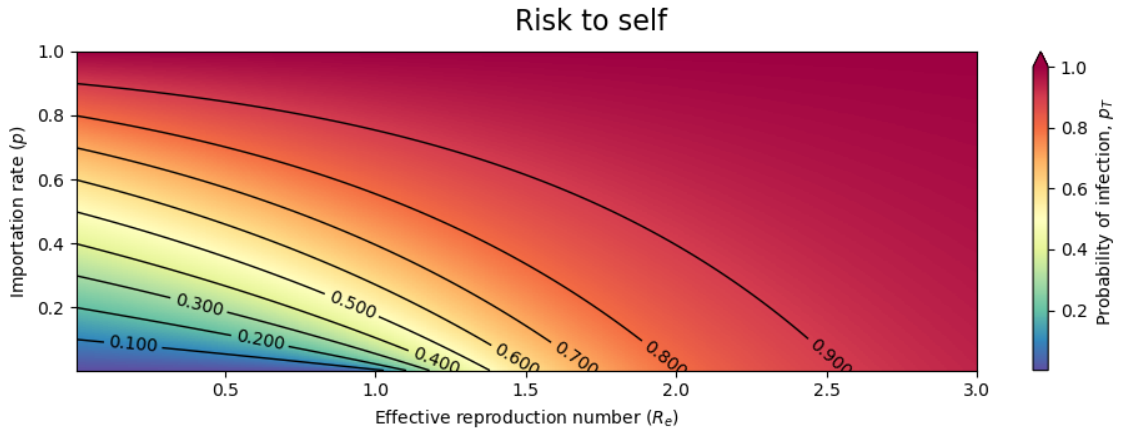
There are two special cases to consider. If $p = 1$, everyone becomes sick from outside, so $s = 0$. In this case, we have $f(x) = \quad x$, meaning that $f(0) = 0$, so the iteration begins at the root, and we have solved it already. If $p = 0$, (i.e. no introductions) no-one will become sick if $R_e < 1$, meaning $s = 1$. If $R_e > 1$, only those with infinite susceptibility sets become sick, so $s = \eta$. [y] In this case we have $f(x) = e^{R_e(x \quad 1)} \quad x$, $f'(x) = R_e e^{R_e(x \quad 1)} \quad 1$ and $f''(x) = R_e^2 e^{R_e(x \quad 1)}$. So, if $R_e < 1$, $f(1) = 0$ and $f'(1) = R_e \quad 1 < 0$. This means that the first x-intercept is at 1, as the function is decreasing at this point, so iterating with Newton's method from 0 will certainly converge to it, by the same reasoning as in the general case. If $R_e > 1$, $f'(1) > 0$, meaning that $f(x)$ crosses the $x$-intercept from negative to positive here. So, somewhere earlier it will have crossed from positive to negative. Since $\eta$ is found using $\varepsilon(\eta) \quad \eta = 0$, it is a root of $f$ when $p = 0$, meaning that this other root must give $\eta$. Since it is the first root, Newton's method starting at $s_0 = 0$ will converge to it.

For any values of $p$ and $R_e$, if we find that, for some $\delta$, $f(s_n + \delta) < 0$, we can say that $s$, which is the one x-intercept on $(0, 1)$, must be somewhere on $(s_n, s_n + \delta)$, since $f$ goes from negative to positive in this domain. This proves the first option for confirming that the error is less than $\delta$. This only applies if $s_n + \delta \quad 1$, as $f(x)$ crosses above 0 after some $x \quad 1$. However, since $s < 1$, $s_n + \delta > 1$ means that the error is less than $\delta$ anyway, giving the second option.

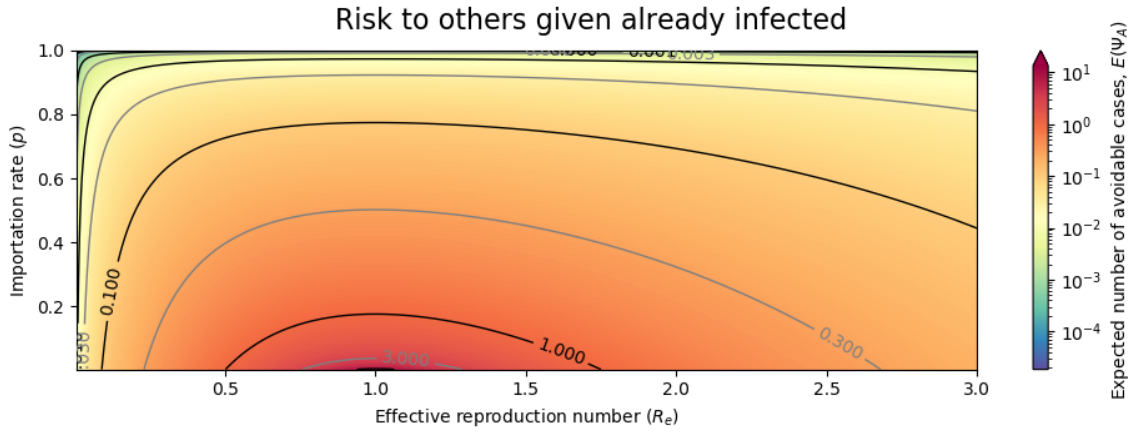
## B Extra results

Figure 8 shows the probability of infection and the expected number of infections caused over the full range of importation rates, from 0 to 1. This is really only included for the sake of interest, as importation rates are generally very low.

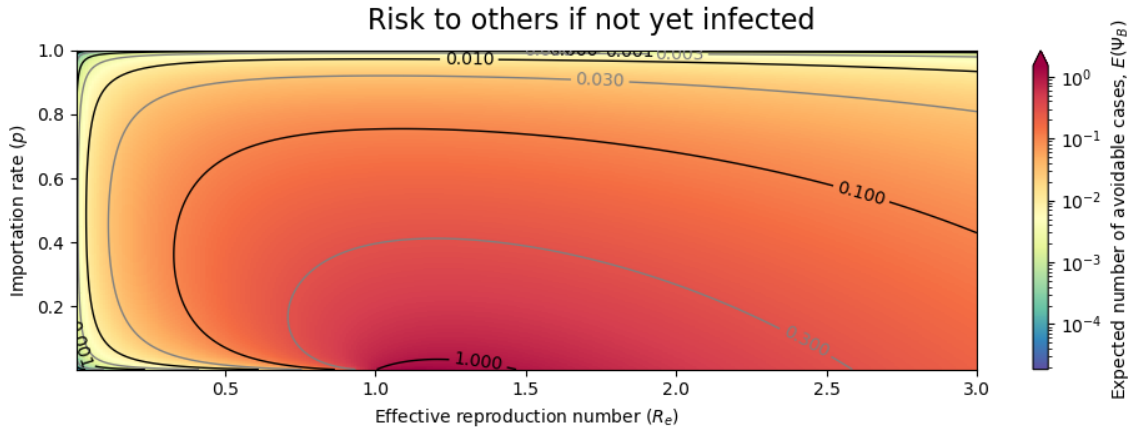
[y] This leaves $R_e = 1$. As $R_e$ drops closer to one, the proportion of the population in the epidemic (i.e. probability of becoming infected with $p = 0$) falls, approaching a limit of 0. When $p = 0$ and $R_e = 1$, $f'(1) = e^0 \quad 1 = 0$, meaning that the graph has a stationary point at $(1, 0)$, which can be the only $x$-intercept, which Newton's method must converge to. So, effectively, when $R_e = 1$, $s = \eta = 0$.




(a) The total probability of catching the disease.




(b) The expected number of people who would have been saved if a non-vaccinator who became sick was vaccinated.




(c) The expected number of people whose infection would have been avoided had the test person been vaccinated, regardless of whether or not the test person becomes sick.

Figure 8: Risks over the full range of importation values.

## C Software

The Python code below was used to generate the data used in the heatmaps.

```
import numpy as np

#----------------------------Accuracy variables----------------------------#

decplaces=12 #The number of decimal places we want to calculate to.
pstop=10**-decplaces #The error bound.

#-------------Imporation rates and effective reproduction numbers-------------#

res=300 #Heatmap resolution. It is only used to generate p and Re
p=np.flip(np.linspace(0,0.003,res)[1:]) #The set of importation rates
Re=np.linspace(0,3,res)[1:] #The set of effective reproduction numbers

#--------------Calculating the probability of not being infected--------------#

#This function calculates s, the total probability of an individual not being
#infected, which is also the probability that an infected individual's
#offspring would have been saved had they been vaccinated.
def scalc(Re,p):
    gmax=1000 #Maximum iteration
    q=1-p
    g=0
    sg=0 #s_g, the current estimate of s
    sgm1=-1 #s_(g-1), the previous estimate of s
    while sg-sgm1>pstop and g<gmax:
        g+=1
        sgm1=sg
        sg=(Re*sgm1-1)*q*np.exp(Re*(sgm1-1))/(Re*q*np.exp(Re*(sgm1-1))-1)
    if g==gmax:
        raise Exception('The system has not converged fast enough.')
    return sg,g


#----------------------------The main function----------------------------#

#This function takes the vectors containing the effective reproduction numbers
#and importation rates, and calculates the total probability of infection and
```

```
#the expected number of people that would have been saved by a vaccination,
#both given that this person became infected (EPsi) and if they have not been
#infected yet (EPsitot). Re and p are extracted as well so that all data can be
#placed in one tuple, making it easier to construct the heatmaps.

def Repitexpcalc(Re,p):
    numRe=len(Re)
    nump=len(p)
    ptot=np.zeros((nump,numRe))
    s=np.zeros((nump,numRe))
    g=np.zeros((nump,numRe),dtype=np.int_)
    for i in range(nump):
        for h in range(numRe):
            s[(i,h)],g[(i,h)]=scalc(Re[h],p[i])
    ptot=1-s #This gives the total probability of infection
    ES=Re*s #This is the expected number of effective offspring
    EPsi=1/(1-ES)-1
    EPsitot=ptot*EPsi
    return ptot,EPsi,EPsitot,Re,p

data=Repitexpcalc(Re,p)
```

## References


Adewale, Olatanwa et al. (2019). "The Ethics of Parental Refusal to Vaccinate: Costs, Community Safety, and Individual Rights". In: *Journal of the Georgia Public Health Association* 7.2, pp. 98–103.

Ball, Frank et al. (2019). *Stochastic Epidemic Models with Inference*. Springer Nature.

Bartsch, Sarah M et al. (2016). "Modeling the economic and epidemiologic impact of hookworm vaccine and mass drug administration (MDA) in Brazil, a high transmission setting". In: *Vaccine* 34.19, pp. 2197–2206.

Chiew, May et al. (2013). "Estimating the measles effective reproduction number in Australia from routine notification data". In: *Bulletin of the World Health Organization* 92, pp. 171–177.

Dube, Eve, Maryline Vivion, and Noni E MacDonald (2015). "Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications". In: *Expert review of vaccines* 14.1, pp. 99–117.

Fransson, Carolina and Pieter Trapman (2019). "SIR epidemics and vaccination on random graphs with clustering". In: *Journal of mathematical biology* 78.7, pp. 2369–2398.

Gidding, Heather F et al. (2018). "Declining measles antibodies in the era of elimination: Australia's experience". In: *Vaccine* 36.4, pp. 507–513.

Harris, Rebecca C et al. (2016). "Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines". In: *Human vaccines & immunotherapeutics* 12.11, pp. 2813–2832.

Hendrix, Kristin S et al. (2016). "Ethics and childhood vaccination policy in the United States". In: *American journal of public health* 106.2, pp. 273–278.

Jamrozik, Euzebiusz, Toby Handfield, and Michael J Selgelid (2016). "Victims, vectors and villains: are those who opt out of vaccination morally responsible for the deaths of others?" In: *Journal of medical ethics* 42.12, pp. 762–768.

MacDonald, Noni E et al. (2018). "Mandatory infant & childhood immunization: Rationales, issues and knowledge gaps". In: *Vaccine* 36.39, pp. 5811–5818.

Miller, Joel C (2018). "A primer on the use of probability generating functions in infectious disease modeling". In: *Infectious Disease Modelling* 3, pp. 192–248.

Penny, Melissa A et al. (2016). "Public health impact and cost-effectiveness of the RTS, S/AS01 malaria vaccine: a systematic comparison of predictions from four mathematical models". In: *The Lancet* 387.10016, pp. 367–375.

Pierik, Roland (2018). "Mandatory vaccination: an unqualified defence". In: *Journal of Applied Philosophy* 35.2, pp. 381–398.