Pre-immunity: the Schrödinger's cat of immunology

Marcus Carlsson, Jens Wittsten, Cecilia Söderberg-Nauclér

Abstract

Early 2020, catastrophic consequences of COVID-19 was predicted in the do-nothing scenario, based on mathematical models for epidemiology. As data began to emerge, several scientists noted that growth did not seem exponential, as the models predicted, leading to speculations of pre-existing immunity or immunological dark matter to explain this pattern. On the other hand, reports of choir-rehearsals infecting most members seemed to refute this, and the topic remained inconclusive. We provide a mathematical theory in which both observations are true; on a population level, pre-immunity exists, on an individual level, it doesn't. This theory demonstrates that established formulas relating e.g. R_0 and the herd-immunity threshold are wrong. We derive new mathematical formulas, which applies to any virus whose transmission dynamics is associated with large individual variability in susceptibility to the infection. Contrary to great variability in infectivity, which we show has no bearing on the mathematical modeling, variability in susceptibility actually manifests itself as pre-immunity on a macroscopic scale, thus making pre-immunity a necessity for accurate mathematical modeling.

1 Introduction

Mathematical models for infectious diseases grossly failed to predict the spread of COVID-19. Of course, the mathematical models work in the do-nothing scenario, to quote the famous Imperial College report 9 [10]. Since everybody did something, we (thankfully) do not have the correct answer, and it is still unclear if the mathematical models would have worked in the do-nothing scenario. However, some places employed very modest restrictions to prevent the spread of COVID-19, such as Sweden, and while it is abundantly clear by comparing with neighboring countries that this was a bad idea, the large gap between what happened and what the do-nothing model predicts, leaves doubt to whether it actually has any bearing on reality in the do-nothing scenario.

The fact that COVID-19 spread was much milder than the models predicted did not go unnoticed, leading to worldwide scientific and political disputes. While some analysts claimed that an immunological dark matter or pre-existing immunity kept a majority from becoming infected, this hypothesis got connected with the Great Barrington Declaration (GBD) and most scientist opted for the explanation that substantial man-made variations in the R-value is the only factor needed to explain the disease spread. If this is the case, it would consequently make all mathematical compartmental models pretty useless for any sort of prediction.

The case of Sweden is very interesting because the Non-Pharmaceutical Interventions (NPI's) have remained virtually constant throughout the past winter, and public compliance has also been fairly constant, which in theory means that the mathematical models should work, just with a new (but fixed) R-value adopted to the new situation. However, in [7] we make a major effort to produce refinements of state of the art models that can get near the measured curves for Sweden, without succeeding. That is, unless we add a pre-existing immunity of around 60%, in which case we can model both the second and third wave of Stockholm with surprising precision [8]. Irrespective of what one thinks about GBD and mitigation strategies, it should cause some scientific curiosity that adding one variable solves an otherwise unsolvable puzzle. But, like the famous mathematician J. von Neumann once said: "Give me 4 variables and I can model an elephant, give me 5 and I can wiggle its trunk." In other words, model fit does not imply model assumptions are true.

There are strong arguments against the idea of a protective pre-existing immunity, for example choirs and parties where almost everyone gets infected. If a pre-existing immunity protects say 60% of the people, the probability that some large choir somewhere on the planet consists of mainly unprotected ones, is virtually zero. On the other hand, there are widespread reports of couples where one gets COVID-19 and the other doesn't, along with reports of the virus burning out of fuel before reaching 30% infected, and towns in hard hit Lombardy have unexpectedly high protection from the relatively low attack rate of the first wave [19].

In this article, we put forward a mathematical theory which merges all these phenomena under one umbrella. We argue that it is not super-spreaders which cause the unexpected behavior, it is super-susceptibles, or more precisely, the existence of great individual variations in susceptibility. In other words, we argue that the individual protection against infection to SARS-CoV-2 is highly variable, but that everybody can get infected, especially if exposed to high viral loads.

However, under NPI's and social distancing, the people with low susceptibility will rarely get infected, thus appearing as if pre-immune. We show that great variability in infectivity has no bearing on the mathematical modeling, whereas variability in susceptibility manifests itself as pre-immunity on a macroscopic scale, and hence that pre-immunity is a necessity for accurate mathematical modeling.

2 A mathematical theory for COVID-19 spread dynamics

Compartmental mathematical models for infectious disease spread have their roots in a series of seminal papers by Kermack and McKendrick [14, 15, 16]. These original papers actually provide models that are more advanced than the commonly used SIR and SEIR models, that derive from this work. For example, the age of infection is considered, leading to more complicated equation systems. However, in [7] it is established that such extensions of the SEIR-model have virtually no bearing on the modeled disease progression. Thus, it seems reasonable to use SEIR for modeling COVID-19, which is also the model of choice used in almost all publications on (large outbreaks of) COVID-19.

Compartmental mathematical models for infectious disease spread are usually cast in the mathematical lingo of Ordinary Differential Equations (ODE's). This is unfortunate from a number of perspectives, the key one being that it leaves most non-mathematicians behind, making epidemiological modeling a field that few can grasp and even fewer can question the foundations of. However, if these ODE's are implemented with a step-size of 1 day, then the differential equations turn into simple equations where a certain amount of infected on day 45, say, leads to a certain amount of newly infected on day 46, whilst yet another amount of people recover, and so on. The difference between the output of such difference equations and their differential equation counterparts is marginal, and one can even argue that the former framework is more natural than the latter, since we do not transmit much while sleeping. An even simpler model, which behaves roughly the same on major outbreaks, is to step forward one generation at a time, which anyone with basic understanding of matrix-vector multiplication can understand. A generation is the amount of people in the various compartments, which in the simplest case consists of Susceptibles, Infectives and Recovered (SIR), during the elapse of one serial interval, i.e. the average time from infection of a susceptible to onward transmission, which for COVID-19 equals about a week. Usually a class E for Exposed is also added, but we have chosen to not include this for simplicity and transparency. (In fact, the output of SIR and SEIR models are rather similar.)

Thus, in order to explain how contemporary SIR-models for COVID-19 spread works, we introduce a simple framework where we step forward one generation at a time. Once this has been introduced, we then discuss what needs to be modified to include both super-spreaders and super-susceptibles in a realistic way, which will lead to new models for the spread of infectious diseases. The aim of this exercise is to reach as broad an audience as possible, since the fundamental considerations about equation system design can be introduced without heavy mathematical formalism. In the Supplementary Material however, we provide proofs in the ODE environment using proper mathematical theory.

To understand how a simple SIR-model work on a generation based count, we denote by N the total amount of individuals in a population, $S(t)$ the amount of susceptible in generation t (where $t = 1, 2, 3, \ldots$), $I(t)$ the amount of infective, and $R(t)$ the amount of recovered (not to be confused with neither R_0 nor R_e introduced shortly). Then the equations for the next generation are

$$\begin{cases} S(t+1) &= S(t) - \nu(t) \\ I(t+1) &= \nu(t) \\ R(t+1) &= R(t) + I(t) \end{cases} \tag{1}$$

where $\nu(t)$ is the amount of new infective (sick) week t. (Clearly, the notation $\nu(t)$ is redundant since it equals $I(t+1)$, but we keep this notation since when generalizing to shorter time-steps we will have $I(t+1) = I(t) + \nu(t) - \sigma I(t)$ where $\sigma = 1/T_{generation}$ and $T_{generation}$ the length of the generation time, but since we step forward generation wise we have $T_{generation} = 1$ and the equations simplify.) The computation of new cases $\nu(t)$ is the heart of the algorithm, and in the simplest case we have

$$\nu(t+1) = \frac{S(t)}{N}\alpha I(t), \tag{2}$$

where α is the transmission rate. Since our step-length equals a serial interval, this formula means that each person in $I(t)$ will transmit the disease to α new infective in a fully susceptible population, yielding $I(t+1) = \alpha I(t)$, but as the fraction $S(t)/N$ of susceptible goes down, so does the amount of newly infected per generation. In terms of the classical R_0-value we see that $R_0 = \alpha$, but we prefer to use α for reasons that will soon become apparent. We can also introduce the effective R-value as

$$R_e(t) = \frac{S(t)}{N}\alpha, \tag{3}$$

which starts at $\alpha = R_0$ and then decays until $R_e < 1$ by which time the epidemic starts to recede. Once the epidemic is extinct, we have $S(t) = 1 - R(t)$. In particular, suppose we start (say after a lock-down) with a value $R(t) > 0$ but $I(t) \approx 0$. Setting the right hand side (3) to 1 and solving for $R(t)$ easily leads to the famous formula $1 - 1/R_0$ for the Herd Immunity Threshold (HIT).

2.1 Super-spreaders

At some point there was a high concern of whether super-spreaders provided the missing link to better understanding of the unexpected COVID-19 spread dynamics, but it turns out that mathematically, at least when it comes to deterministic models, super-spreaders have no effect what so ever on disease progression. This has been pointed out in [5, 7] and we shall also demonstrate it below. Here, one has to separate between super-spreaders in the sense that they have an extremely high viral load (possibly combined with lack of symptoms), and *social* super-spreaders, which are persons whose work or behavior puts them in contact with an unusually high number of people. The latter category may indeed have some effect on the disease progression, as they will be more inclined to become infected early on and thus being depleted in the early stages of a pandemic. This was modeled in [4] and also in [7], where it was argued that while the effect is there, it does not drastically alter the output of the model as long as parameters are chosen realistically.

At this late stage of the pandemic, surely most social superspreaders have been exposed to the virus, so we here focus on the former category, those whose viral load can be over a thousand times higher than the average infective individual (see e.g. [13]). To model this mathematically we divide each generation of infectives $I(t)$ in subgroups $I_1(t)$, $I_2(t)$ and $I_3(t)$. The first sub-group consists of, say, 10% superspreaders, the second subgroup of 60% normal infectives, and then the last 30% have so low viral load that they hardly transmit onward at all. Say that an average person has $a = 10$ infectious contacts during the elapse of the disease, and that the first group can infect all 10, the normal group only can infect 1, and the last group (on average) only 0.2. This means that the amount $\nu(t+1)$ of new infectives is given by the formula

$$\nu(t+1) = \frac{S(t)}{N} a(I_1(t) + 0.1 I_2(t) + 0.02 I_3(t)). \tag{4}$$

If we allocate weights $w_1 = 0.1$, $w_2 = 0.6$ and $w_3 = 0.3$ corresponding to the fraction in each respective group, then $I_j = w_j I(t)$ and the above formula reads

$$\nu(t+1) = \frac{S(t)}{N} a(w_1 + 0.1 w_2 + 0.02 w_3) I(t) = \frac{S(t)}{N} a(0.1 + 0.06 + 0.006) I(t) = \frac{S(t)}{N} 1.66 I(t). \tag{5}$$

But now we have only retrieved (2) with an average $\alpha = R_0$ of 1.66. In other words, the super-spreaders become completely invisible in the mathematical model, and hence are unlikely to be the main key to understanding COVID-19 spread dynamics.

2.2 Super-susceptibles

Curiously, the same phenomenon does not occur when we add variation in susceptibility. We first discuss this from a mathematical and then from a medical angle. For simplicity we use the same weights $w_1,\ w_2$ and w_3 as before, but now we divide $S(t)$ into three subgroups $S_1(t)$, $S_2(t)$ and $S_3(t)$ where the first group S_1 has a very good protection against infection with SARS-CoV-2, the second group is normal and the last group consists of the unfortunate who are easily infected, representing the Super-susceptibles. We use the same distribution 10/60/30 as above, so 10% are lucky, 60% average and 30% being unfortunate. Let us say that 100% of all potentially infectious contacts with the super-susceptible group leads to transmission. The amount of contacts between an infective and this group becomes $aw_3 = a(S_3(0)/N)$. Hence the formula for new infections in this group becomes

$$\nu_3(t) = (S_3(t)/N) a I(t),$$

since the amount of encounters leading to infection will decrease as immunity builds up in S_3. In the average group, say only 10% as many get infected, i.e., $\nu_2(t) = 0.1 S_2(t) a I(t)/N$ and in the well-protected group we get only 2%; $\nu_1 = 0.02 S_1(t) a I(t)/N$. Of course, a flaw in the present set-up is that model constants will in practice be unknown (which seems common in mathematical epidemiology, see e.g. [3, 4] or [5] that contain SEIR-type models precisely along the above lines). Here, our overall goal is to move further and eventually replace such systems by more realistic ones, not relying on an abundance of constants picked by the user.

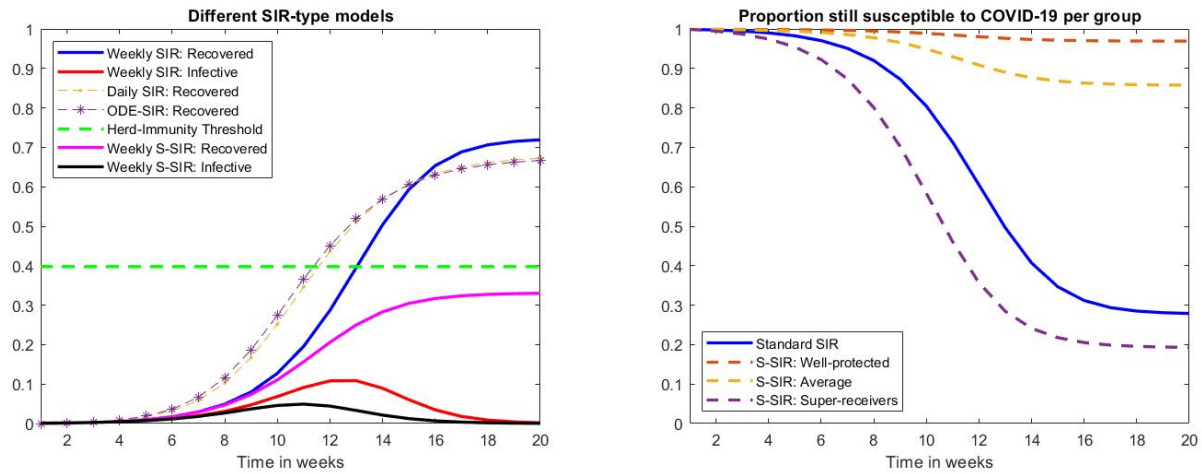

Figure 1: Left, different incarnations of SIR and S-SIR. Right, fraction remaining susceptibles in each subgroup of S-SIR.

Back to the topic, using vector notation we can arrange our equations as follows

$$\begin{bmatrix} \nu_1(t) \\ \nu_2(t) \\ \nu_3(t) \end{bmatrix} = \begin{bmatrix} p_1 S_1(t) \\ p_2 S_2(t) \\ p_3 S_3(t) \end{bmatrix} \frac{aI(t)}{N} \tag{6}$$

where (p_1, p_2, p_3) in the above example equals $(0.02, 0.1, 1)$, and now the total amount of newly infected becomes

$$\nu(t) = \nu_1(t) + \nu_2(t) + \nu_3(t) = \frac{p_1 S_1(t) + p_2 S_2(t) + p_3 S_3(t)}{N} aI(t), \tag{7}$$

which, in opposition to (5), does not reduce to the simplistic (2). Interpreting the coefficient in front of $I(t)$ as the effective R-value, we get $R_0 = (p_1 w_1 + p_2 w_2 + p_3 w_3)a$ and

$$R_e = \frac{p_1 S_1(t) + p_2 S_2(t) + p_3 S_3(t)}{N} a = \frac{p_1 S_1(t) + p_2 S_2(t) + p_3 S_3(t)}{(p_1 w_1 + p_2 w_2 + p_3 w_3)N} R_0, \tag{8}$$

which should be compared with (3).

The key difference with this new model is the fact that ν_1 gets withdrawn from S_1, ν_2 gets withdrawn from S_2 and ν_3 gets withdrawn from S_3. Since the fraction $\nu_3(t)/S_3(t)$ is larger than $\nu_1(t)/S_1(t)$, this means that the super-susceptibles will become depleted faster. But the super-susceptibles contribute to the main part of the coefficient in front of $I(t)$ in (7), and hence it turns out that the disease starts to recede when S_3 becomes sufficiently low, way ahead of what is expected from the (standard homogenous) SIR-model (1)-(2). In other words, while a standard SIR model needs to get above the herd-immunity threshold $1 - 1/R_0$ before it dies out, the model here can die out as soon as the super-susceptibles are depleted. Thus, if we sacrifice an unknown group of unexpected super-susceptibles, the disease can die out by itself at fractions much lower than predicted by HIT, although such a strategy is of course questionable from an ethical standpoint. Super-susceptibles should not be confused with risk-groups, since these rarely transmit the disease onwards.

Incidentally, at the end of each paper [14, 15, 16], Kermack and McKendrick stress that a weakness in their model is that they assume uniform susceptibility, which they consider unrealistic in many cases. However, it seems that they never got around to correct this issue, and we have not found a rigorous analysis of how to model this mathematically elsewhere in the literature either. A reason for this could be the problem with picking realistic constants, mentioned earlier, which we will address later in this paper. In summary, the formula $1 - 1/R_0$, which stems from their seminal papers, may very well be inaccurate for COVID-19.

These ideas are illustrated in Figure 1, where we use $R_0 = 1.66$. The blue curve is the amount of recovered in a standard SIR model (1)-(2) with a weekly step-length. To underline the claims made earlier that this gives a good indication of how SIR-with daily step-length behaves, which in turn is almost identical to the corresponding ODE-version of SIR, we include the graphs of recovered also in these models, along with the herd immunity threshold. While the weekly SIR clearly deviates, all three exhibit a similar behavior and level out at around 70% infected, way above the HIT

at around 40%. The red curve displays the amount of infective per week, which typically looks like a slightly distorted Gaussian (i.e., bell-shaped) curve. Finally, pink and black show recovered/infective for the Susceptibility-stratified SIR model (6)-(7) (S-SIR). Looking at the black curve we see that R_e hits 1 around week 10, whereas in the first case (red curve) this happens at week 12 (since $R_e = 1$ is equivalent with I reaching its peak). Despite that these are not too far apart, the effect on the overall spread is huge; the amount of infective each week in the S-SIR is much lower and the pandemic levels out clearly below the HIT, with a final size of the pandemic roughly half that of the standard SIR. The right graph displays the fraction of susceptibles left in each of the 3 compartments as a function of time (dashed), as well as the corresponding graph for the full population in the standard SIR model (blue solid). Clearly, the group of super-suseptibles have a huge attack rate of around 80% (compared with 72% for the whole population in the standard SIR), but as a consequence, the other two groups have attack rates of around 15% and 5%, respectively.

To give the reader a feeling for how this equation system works, we include the flipped situation where only 10% are super-susceptibles and 30% have very good protection against SARS-CoV-2. Surprisingly, not much changes, the tiny group of super-susceptibles still protect the rest. This group now have a 92% attack rate (as opposed to previously 80%), and when the disease calms down in this compartment, it calms down in the other as well. Looking at the embedded image with the infective/recovered, we also see that the recovered compartment looks almost as either of the curves in Fig. 1 (left), only the final size of the pandemic changes to 23%, from 33% for the previous S-SIR and 72% for standard SIR, respectively. Also curves for infectives I are almost identical except for their magnitude.

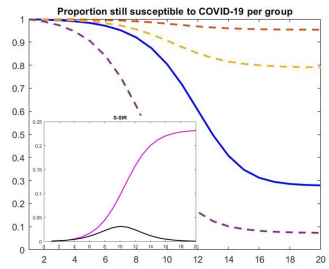

Figure 2: S-SIR with only 10% super-susceptibles.

This is frustrating, since intuitively, it should be possible to find a setting of parameters where the disease burns out among the super-susceptibles, but lives on a bit longer in the average group. One of the key findings of this paper is that this intuition is wrong, which also will explain why these curves look so similar apart for magnitude, see Section 4.

To round up this section, we remark that it is possible to create a more realistic model which merges the framework of the two previous sections by considering various infectivity groups $I_1, \ldots, I_K$ and susceptibility groups $S_1, \ldots, S_J$, along with transmission probabilities $p_{j,k}$ for an encounter between a member of S_j and a member of I_k. However, it turns out that such a model naturally collapses (in a similar manner as was described in Section 2.1) to the same model considered in this section. We provide a proof of this in SM, Section 6.

3 Does variable susceptibility exist?

Many things are still unknown about the SARS-CoV-2 virus but one thing is certain, it affects different individuals in different ways. Whereas some people become terminally ill, the majority has light symptoms, whereas yet others get longlasting symptoms of an unprecedented variety, even if they have not been very ill in the acute phase of the disease. Likewise the ability to spread is highly individual, it has been demonstrated that some people get a viral load over a million times higher than others [13]. In addition, this seems uncorrelated to how sick they become; many individuals with very high viral loads are even asymptomatic. In light of this, the most probable assumption is that also the way the virus enters the human is subject to large individual variations.

It has by now been established that SARS-CoV-2 primarily spreads via airborne transmission [11]. While a certain exposure to air containing SARS-CoV-2 may be enough to infect a highly susceptible individual, a much higher load may be needed to infect an average individual and possibly, due to e.g. an immunological cross-reactivity to other virus, some people may in practice be immune, unless they expose themselves to a very high viral load for an extended period of time, as happens for instance during choir-rehearsal (note that it has been established that singing emits orders of magnitude more virus in the air than normal talking [1]).

Several papers have proposed mechanisms that could lead to a higher level of protection against SARS-CoV-2, we mention e.g. [9, 17]. In a separate manuscript, we provide evidence of a protective effect by cross-reactive antibodies between a peptide in neuraminidace of Influenza A H1N1 and the binding domain the spike protein of SARS-CoV-2 that interacts with the angiotensin-converting enzyme 2 (ACE2), [18].

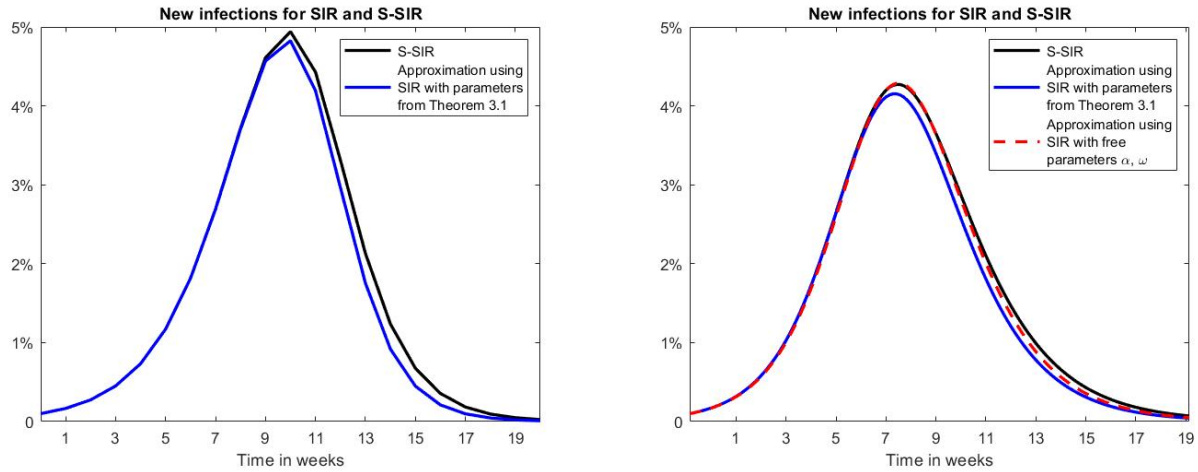

Figure 3: Left, curves for I produced by S-SIR (6)-(7) and corresponding ones for SIR (1)-(2) with parameters from Theorem 4.1, yielding a pre-immunity of 57%. Right, same curves for the corresponding systems using a step-length of 1 day instead of 1 week, plus a fitting with free parameters α and ω. Pre-immunity for this curve was set to 45%.

4 Reduction to SIR-modeling with pre-immunity

Incorporating pre-immunity into a basic SIR-model is simple, one just replaces the initial condition $S(0) = N$ with $S(0) = \omega N$ where ω is the fraction of the population that is not pre-immune. In [7] we showed that such modification to the slightly more advanced SEIR-model is capable of very accurately fitting the observed data from cities like Stockholm and Manaus, as well as the country of India. However, the only drawback is that pre-immunity seems to be an unpopular theory that few people believe in, and admittedly, just because a mathematical model fits with observed data, it doesn't prove that the interpretation of the model is right. The key message of this paper is that anyone believing in variable susceptibility should use pre-immunity when modeling COVID-19 spread, the reason being the following:

Theorem 4.1. *Given any dynamical system of the type* (6)-(7)*, there is a corresponding system of the type* (1)-(2) *such that, if pre-immunity is incorporated, the curves of total susceptible, infective and recovered become very similar*[1]*. A choice of* ω *and* α *(recall* (2)*) which yields this is*

$$\omega = \frac{\left(w_1 p_1 + w_2 p_2 + w_3 p_3\right)^2}{w_1 p_1^2 + w_2 p_2^2 + w_3 p_3^2}. \tag{9}$$

and

$$\alpha = a \frac{w_1 p_1^2 + w_2 p_2^2 + w_3 p_3^2}{w_1 p_1 + w_2 p_2 + w_3 p_3} \tag{10}$$

We provide the details in the Supplementary Material. Note that we also have $R_0 = \alpha\omega$ by the computation preceding (8). Thus, given α and pre-immunity θ, the apparent R_0-value in the beginning of an epidemic becomes

$$R_0 = \omega\alpha = (1-\theta)\alpha \tag{11}$$

(which is why we kept α and R_0 separate in (2)). Hence, one has to be careful with the interpretation of α as R_0 in (2), since this only holds when $\theta = 0$. Even more important is to be careful with the interpretation of θ as fraction of people with protective pre-immunity, since there is, in reality, not a division of θN immune and $(1-\theta)N$ susceptible. To avoid confusion, we introduce the term "population pre-immunity" for θ.

4.1 Examples

We now show some illustrative examples. Figure 3, left, shows the I-curves obtained by running the model (6)-(7) with the same parameters as were used in Figure 1, (so the black curve is the same as the black curve from Figure 1, but displayed

[1] the precise meaning of which is made clear in the SM

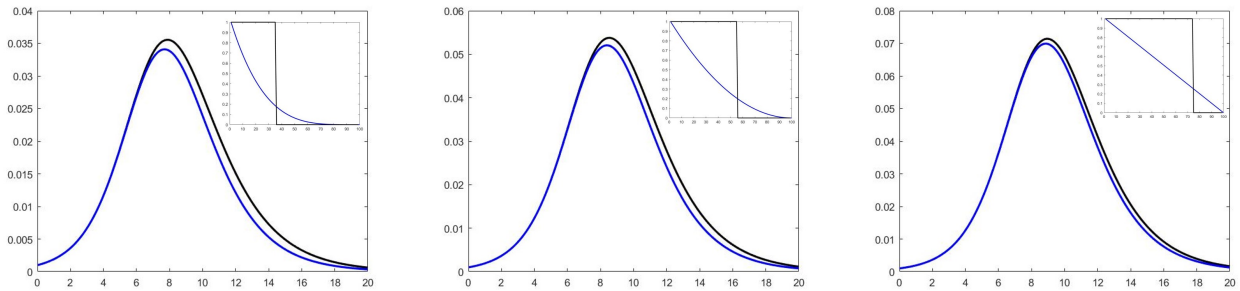

Figure 4: Various p-curves and corresponding wave of infectious for S-SIR and SIR.

on a different scale). In the same figure we see the corresponding curve for basic SIR (1)-(2) with a pre-immunity level of 57%. To the right, we see the same curves from the corresponding model with step-length of one day, as opposed to one generation (i.e. a week). Clearly, there is a non-negligible difference so henceforth we will use this step-length instead in the remainder of the paper. (Again, the simpler system was introduced for easy accessibility of the key ideas.) The red dashed curve was obtained by freely choosing ω and α to obtain best possible fit, which shows that it is clearly possible to do better than the choice in Theorem 4.1. In practice, where anyways the right-hand sides of (10) and (9) are unknown, this means that we can always get an almost perfect fit. This conclusion is based on trying a large amount of different weights and probabilities.

Next, we divide the population into 100 (equal size) subgroups (as opposed to 3), and assign probabilities according to the formula $p_j = \left(1 - \frac{j-1}{100}\right)^{\rho}$ where ρ is a parameter to be chosen. This means that we assume a monomial distribution of individual transmission probabilities, and see which exponent ρ that seems to give a good fit. The paper [7] contains several examples of how to pick α and θ to match a given observed curve of cases, and by studying India, Manaus and Stockholm we conclude that a pre-immunity in the range 40-65% gives a good fit between modeled and observed data.

In figure 4 we show the curves for $\rho = 4,\ 2$ and 1 (p_j-curve in embedded picture) with respective values of pre-immunity by Theorem 4.1 equal to $64\%,\ 45\%$ and 25% respectively, which give further support to the hypothesis that the group of individuals who are very easily infected by SARS-CoV-2, is relatively small. Note that basic SIR with pre-immunity becomes a special case of S-SIR if we choose a function p that is either 0 or 1 with a sharp drop at 100ω. These curves are also seen in the embedded pictures.

4.2 Variants of concern

The variants of concern are known for higher infectivity or higher R_0-value, but how to model them? If an epidemic with known values of $R_0 = \alpha_1\omega_1$ that seems to be over, suddenly starts to expand again due to a new VoC (as happened already once in Manaus, Stockholm, the UK and India), the new R_0-value at that time can be measured, let us denote it $\tilde{R}_0$. But looking at (11), it is clear that the same value can be achieved by either increasing ω or α, or both. Which option we pick leads to a different size of the new outbreak.

To see which option is correct, we return to (6), supposing that this choice of parameters had somehow been measured. The value of p_1 is 1, meaning all encounters between infected and super-susceptibles lead to infection. This value can clearly not grow further. Instead, the values p_2 and p_3 need to be adjusted upwards, due to the ability of the VoC to infect people who had a good protection against the original strain. Looking at the formulas (9) and (10) we see that such adjustments affect both ω and α in a non-linear way, and hence both need to be updated. Therefore, anytime there is a new VoC, the only way to estimate these new parameter-values is to wait long enough until the profile of the new wave becomes clear enough.

4.3 The herd-immunity threshold

As noted in Section 2, the classical formula for the herd-immunity threshold is

$$HIT = 1 - \frac{1}{R_0}. \tag{12}$$

In light of Theorem 4.1, this is not a good approximation of reality. In the model (1)-(2) with $S(0) = \omega N$, we have $R_0 = \omega\alpha$ and the disease will start to die out as soon as $\frac{S(t)}{N}\alpha = 1$. The HIT equals the difference between $S(t)/N$ at

this point and $S(0)/N$, i.e.

$$HIT = \frac{S(0) - S(t)}{N} = \omega - \frac{1}{\alpha} = \omega(1 - \frac{1}{R_0}), \tag{13}$$

which implies that the classical estimate are grossly over-estimating the HIT. How much is impossible to say, but the modeling from India, Manaus and Stockholm in [7] puts ω in the range 0.35-0.6, (that is, of course, under the current NPI's), which puts the herd-immunity threshold at somewhere between 20% and 40% depending on whether we put $R_0 = 2$ or $R_0 = 3$, which is precisely the levels at which it has been observed that the spread calms down. This most likely also includes Manaus, see [12], which argues that the measured seroprevalence of 44% [6] most likely does not reflect the true seroprevalence, which is rather estimated to somewhere around 26%.

4.4 Further extensions

So far, we have limited the discussion to only include variable infectivity and susceptibility in the SIR-framework, with the aim of increasing transparency to non-mathematical readers. In the supplementary material we work with more advanced SEIR models taking variable interaction patterns between age groups, as well as variations social activity-level, into account, following [4]. In short, we conclude that there is only a marginal difference between SEIR and SIR, whereby the results reported above extend to this setting as well. Furthermore, the same pattern repeats itself, where other heterogeneities due to age and activity-level become virtually equivalent with basic SIR upon including pre-immunity. The amount of population pre-immunity that can be attributed to variable susceptibility versus these other types of population heterogeneities remain unclear, but it seems unlikely that the levels of pre-immunity needed in order to accurately model e.g. Stockholm [8] could be reached without a substantial part coming from variable susceptibility. See Section 8 for a more thorough discussion.

5 Discussion

Almost a century ago, the physical community were struggling to understand the seemingly contradictory observations of the wave-particle duality of light and odd illustrations of the inherently probabilistic nature of quantum mechanics. In 2020, experts on immunology and virology faced a similar situation, leading to one camp claiming pre-immunity existed and another one claiming the opposite. We here provide a mathematical theory that says that in a way, both camps were right.

In the single-particle variation of the famous double-slit experiment of quantum mechanics, individual particles (such as photons) are sent through the double-slit apparatus and detected as single particles on a detection screen. However, as more and more individual particles are sent through in this way, an interference pattern emerges on the screen that can only be explained by single particles also having wave-like properties at the same time. Analogously, in our mathematical framework, it is possible that no person has pre-immunity to SARS-CoV-2, yet when observing a large population, around 50% can be observed to have it. This does not necessarily mean that 50% are immune to SARS-CoV-2, but including this in the model is the only way to accurately model the observed data, i.e., how the virus spreads in society. We call it "population pre-immunity" to avoid confusion with sterilizing individual pre-immunity.

There could be many reasons for why certain people are more susceptible than others to infection by SARS-CoV-2, ranging from innate and adaptive immunity to cross-reactive protective immunity from another known virus. But does sterilizing pre-immunity exist, i.e. are there individuals which are completely immune to SARS-CoV-2 without ever having had the virus? Nobody will ever design an experiment to find out if some people are really pre-immune to SARS-CoV-2, since it would include exposing a large number of seemingly healthy unvaccinated people to high viral loads, to see if they all develop COVID-19 or not. But suppose that we were to conduct such an experiment with a group of individuals that are believed to be good candidates for pre-immunity. If all fall ill one can argue that the group was not correctly chosen, if some do not fall ill one can argue that they would have fallen ill with a higher dose. The question of whether pre-immunity really exists is reminiscent of Schrödinger's famous thought experiment of whether one can determine if a cat in a box is dead or alive, without opening the box, with the popular scientific interpretation that prior to observation, the cat is both dead and alive at the same time.[2]

Similarly, it seems like pre-immunity both exists and does not exist at the same time. More precisely, there is no way of really determining whether sterilizing pre-immunity exists with the current tools of immunology, but on a population scale it is clear that some people act as if pre-immune, which is evident by considering the epidemic data from Stockholm County [7, 8]. The key point of this study is that individual pre-immunity is not needed in order to observe pre-immunity

[2]The analogy with physics is only meant to be illustrative, and should not be taken literally.

on a population level, we only need moderate variation in susceptibility in order to observe a substantial pre-immunity of on a population level. Moreover, other types of population heterogeneities, such as variable activity level, also manifest themselves as pre-immunity, although to a lesser degree. Based on this and basic immunological principles, we believe that the most likely hypothesis regarding SARS-CoV-2 transmission, is that everybody can get COVID-19 but that there are great individual variations in the dose required for infection. The findings in this paper do not limit themselves to SARS-CoV-2, but basically shows that classical formulas for the herd-immunity threshold and the models for spread of infectious diseases with roots in the famous paper by Kermack and McKendrick [14] are inapt to model any infectious disease subject to large variability in susceptibility.

Acknowledgement

We thank Erik Wahlén for fruitful discussions.

References

[1] M Alsved, A Matamis, R Bohlin, M Richter, P-E Bengtsson, C-J Fraenkel, P Medstrand, and J Löndahl. Exhaled respiratory particles during singing and talking. *Aerosol Science and Technology*, 54(11):1245–1248, 2020.

[2] Zvi Artstein. Continuous dependence on parameters: On the best possible results. *Journal of Differential Equations*, 19(2):214–225, 1975.

[3] Fred Brauer, Carlos Castillo-Chavez, and Zhilan Feng. *Mathematical models in epidemiology*. Springer, 2019.

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

[5] Tom Britton, Pieter Trapman, and Frank Ball. The risk for a new covid-19 wave–and how it depends on r_0, the current immunity level and current restrictions. *arXiv preprint arXiv:2010.04574*, 2020.

[6] Lewis F Buss, Carlos A Prete, Claudia MM Abrahim, Alfredo Mendrone, Tassila Salomon, Cesar de Almeida-Neto, Rafael FO França, Maria C Belotti, Maria PSS Carvalho, Allyson G Costa, et al. Three-quarters attack rate of sars-cov-2 in the brazilian amazon during a largely unmitigated epidemic. *Science*, 2020.

[7] Marcus Carlsson, Gad Hatem, and Cecilia Söderberg-Nauclér. Mathematical modeling suggests pre-existing immunity to sars-cov-2. *medRxiv*, 2021.

[8] Marcus Carlsson and Cecilia Söderberg-Nauclér. Indications that stockholm has reached herd immunity, given limited restrictions, against several variants of sars-cov-2. *medRxiv*, 2021.

[9] Kieran Dee, Daniel M Goldfarb, Joanne Haney, Julien AR Amat, Vanessa Herder, Meredith Stewart, Agnieszka M Szemiel, Marc Baguelin, and Pablo R Murcia. Human rhinovirus infection blocks sars-cov-2 replication within the respiratory epithelium: implications for covid-19 epidemiology. *Journal of Infectious Diseases*, 2021.

[10] Neil Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, et al. Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. *Imperial College London*, 10(77482):491–497, 2020.

[11] Trisha Greenhalgh, Jose L Jimenez, Kimberly A Prather, Zeynep Tufekci, David Fisman, and Robert Schooley. Ten scientific reasons in support of airborne transmission of sars-cov-2. *The lancet*, 397(10285):1603–1605, 2021.

[12] Daihai He, Yael Artzy-Randrup, Salihu S Musa, and Lewi Stone. The unexpected dynamics of covid-19 in manaus, brazil: Herd immunity versus interventions. *medRxiv*, 2021.

[13] Terry C Jones, Guido Biele, Barbara Mühlemann, Talitha Veith, Julia Schneider, Jörn Beheim-Schwarzbach, Tobias Bleicker, Julia Tesch, Marie Luisa Schmidt, Leif Erik Sander, et al. Estimating infectiousness throughout sars-cov-2 infection course. *Science*, 2021.

[14] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of London. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[15] William Ogilvy Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics ii. the problem of endemicity. *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character*, 138(834):55–83, 1932.

[16] William Ogilvy Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics iii. further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 141(843):94–122, 1933.

[17] Kevin W Ng, Nikhil Faulkner, Georgina H Cornish, Annachiara Rosa, Ruth Harvey, Saira Hussain, Rachel Ulferts, Christopher Earl, Antoni G Wrobel, Donald J Benton, et al. Preexisting and de novo humoral immunity to sars-cov-2 in humans. *Science*, 370(6522):1339–1343, 2020.

[18] Cecilia Söderberg-Nauclér et. al. Unpublished observations, manuscript in preparation.

[19] Marco Vinceti, Tommaso Filippini, Kenneth J Rothman, Silvia Di Federico, and Nicola Orsini. Sars-cov-2 infection incidence during the first and second covid-19 waves in italy. *Environmental research*, page 111097, 2021.