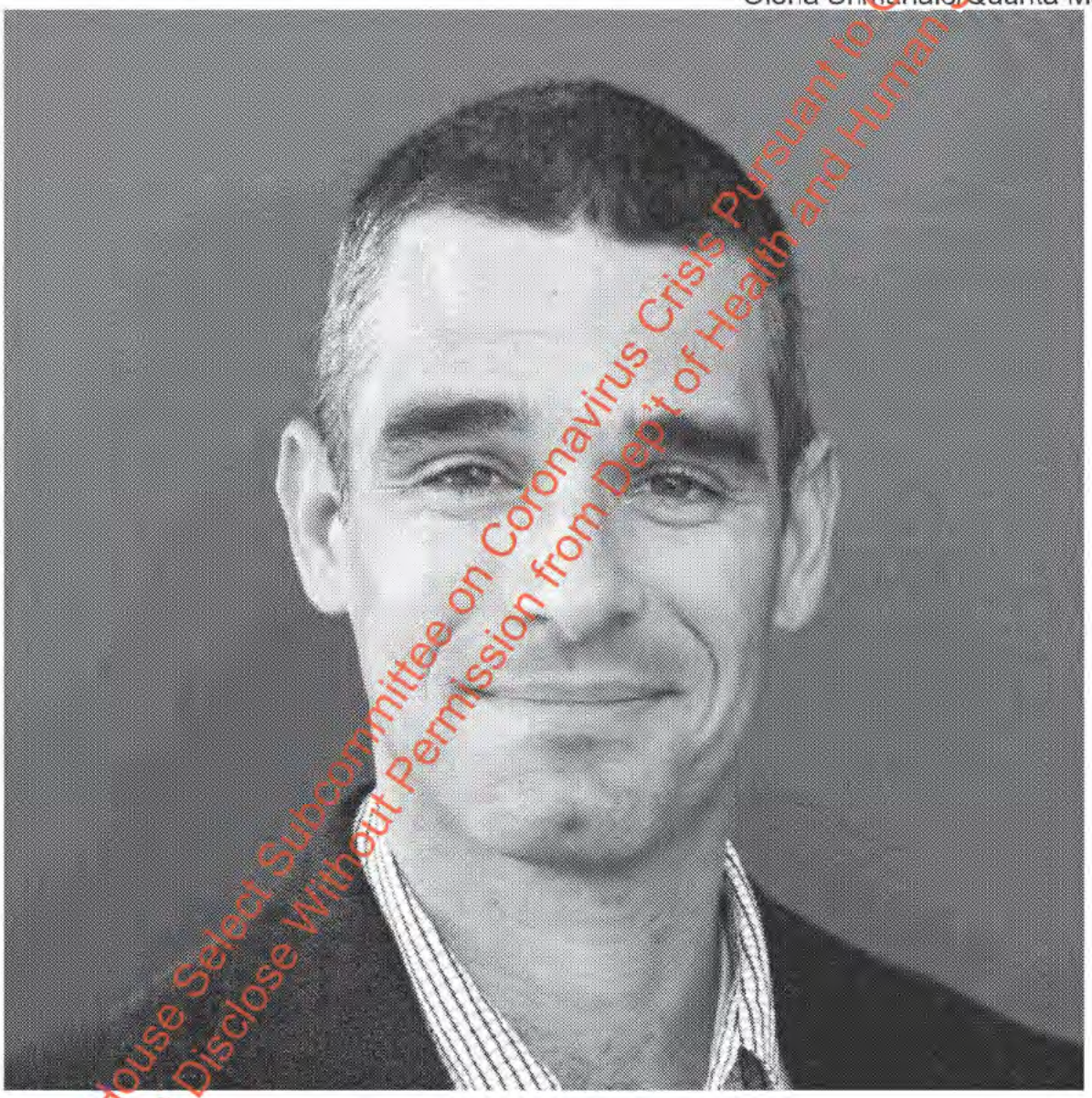
Message

From: Alexander, Paul (HHS/ASPA) [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=BC4EDA8AD333439EB3D296AE0E0F9634-ALEXANDER,]
Sent: 7/26/2020 1:29:11 AM
To: Paul Alexander [redacted]
Subject: FW: NYC 25% antibodies...and aircraft Rosevelt Herd paper
Attachments: Herd paper.pdf

t's not easy to figure out when a disease will stop spreading through a population.

Olena Shmahalo/Quanta Magazine



Kevin Hartnett

Senior Writer

June 30, 2020

VIEW PDF/PRINT MODE

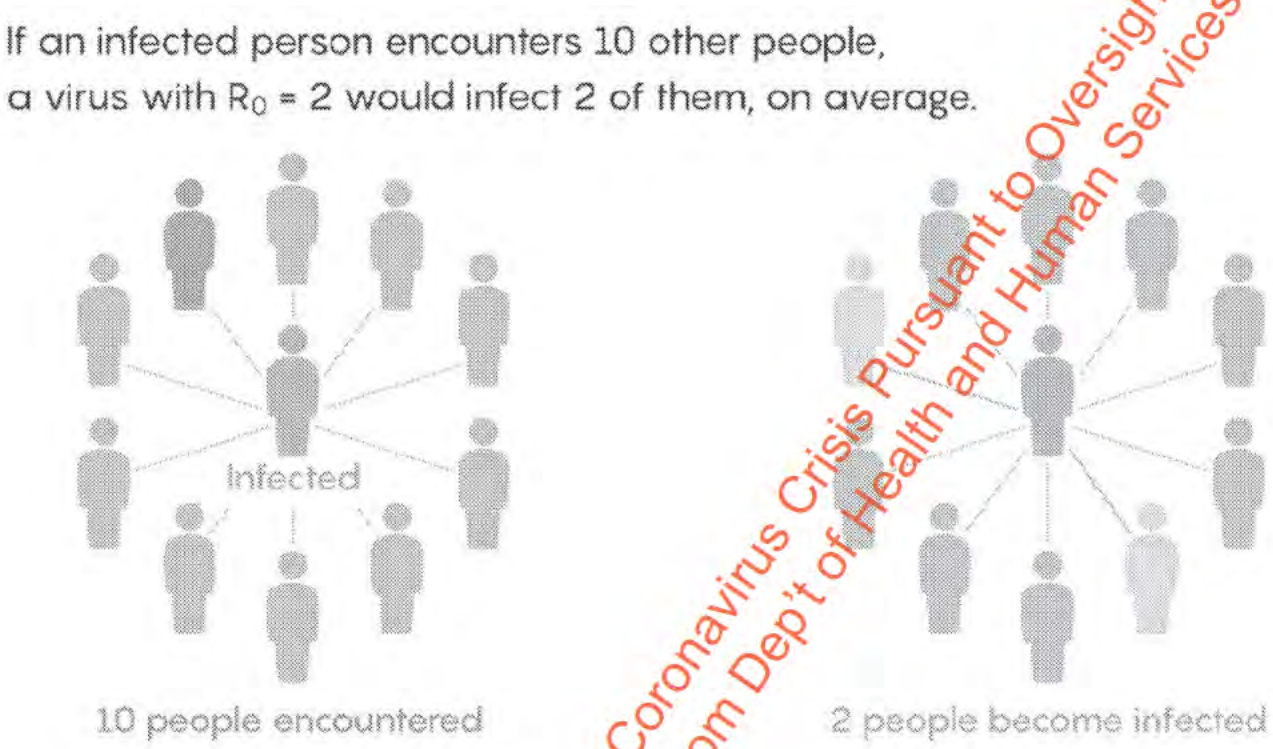
While much about the COVID-19 pandemic remains uncertain, we know how it will likely end: when the spread of the virus starts to slow (and eventually ceases altogether) because enough people have developed immunity to it. At that point, whether it's brought on by a vaccine or by people catching the disease, the population has developed "herd immunity."

"Once the level of immunity passes a certain threshold, then the epidemic will start to die out because there aren't enough new people to infect," said Natalie Dean of the University of Florida.

While determining that threshold for COVID-19 is critical, a lot of nuance is involved in calculating exactly how much of the population needs to be immune for herd immunity to take effect and protect the people who aren't immune.

At first it seems simple enough. The only thing you need to know is how many people, on average, are infected by each infected person. This value is called R_0 (pronounced "R naught"). Once you have that, you can plug it into a simple formula for calculating the herd immunity threshold: $1 - 1/R_0$.

Let's say the R_0 for COVID-19 is 2.5, meaning each infected person infects, on average, two and a half other people (a common estimate). In that case, the herd immunity threshold for COVID-19 is 0.6, or 60%. That means the virus will spread at an accelerating rate until, on average across different places, 60% of the population becomes immune.

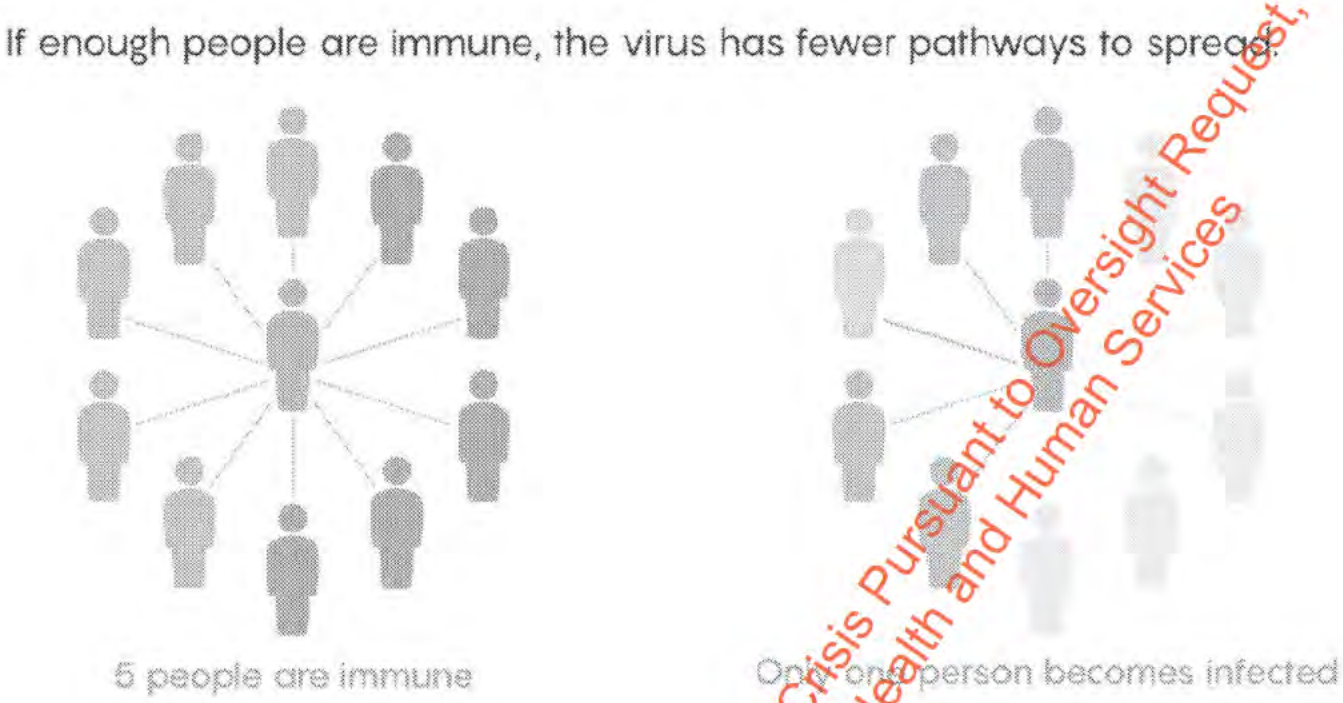


Lucy Reading-Ikkanda/Quanta Magazine

At that point, the virus will still spread but at a decelerating rate until it stops completely. Just as a car doesn't come to a stop the moment you take your foot off the gas, the virus won't vanish the moment herd immunity is reached.

"You could imagine that once 60% of the population is infected, the number of infections starts to drop. But it might be another 20% that gets infected while the disease is starting to die out," said Joel Miller of La Trobe University in Australia.

That 60% is also the threshold past which new introductions of the virus — say, an infected passenger disembarking from a cruise ship into a healthy port with herd immunity — will quickly burn out.



Lucy Reading-Ikkanda/Quanta Magazine

"It doesn't mean you won't be able to start a fire at all, but that outbreak is going to die," said Kate Langwig of Virginia Polytechnic Institute and State University.

However, things quickly get complicated. The herd immunity threshold depends on how many people each infected person actually infects — a number that can vary by location. The average infected person in an apartment building may infect many more people than the average infected person in a rural setting. So while an R_0 of 2.5 for COVID-19 may be a reasonable number for the whole world, it will almost certainly vary considerably on a more local level, averaging much higher in some places and lower in others. This means that the herd immunity threshold will also be higher than 60% in some places and lower in others.

"I think the range of R_0 consistent with data for COVID-19 is larger than most people give credit to," said Marc Lipsitch of Harvard University, who has been advising health officials in Massachusetts and abroad. He cited data indicating it could be more than twice as high in some urban settings as the overall U.S. average.

And just as R_0 turns out to be a variable, and not a static number, the way people acquire their immunity also varies, with important implications for calculating that herd immunity threshold.

Usually, researchers only think about herd immunity in the context of vaccine campaigns, many of which assume that everyone is equally likely to contract and spread a disease. But in a naturally spreading infection, that's not necessarily the case. Differences in social behaviors lead some people to have more exposure to a disease than others. Biological differences also play a role in how likely people are to get infected.

Gabriela Gomes of the University of Strathclyde in Scotland studies how biological and behavioral differences can affect the spread of a virus. She concludes some parts of the world may already be close to reaching herd immunity.
Courtesy of Gabriela Gomes

"We are born different, and then these differences accumulate as we live different experiences," said Gabriela Gomes of the University of Strathclyde in Scotland. "This affects how able people are to fight a virus."

Epidemiologists refer to these variations as the "heterogeneity of susceptibility," meaning the differences that cause some people to be more or less likely to get infected.

But this is too much nuance for vaccination campaigns. "Vaccines are generally not distributed in a population with respect to how many contacts people have or how susceptible they are, because we don't know that," said Virginia Pitzer of the Yale School of Public Health. Instead, health officials take a maximalist approach and, in essence, vaccinate everyone.

However, in an ongoing pandemic with no guarantee that a vaccine will be available anytime soon, the heterogeneity of susceptibility has real implications for the disease's herd immunity threshold.

In some cases it will make the threshold higher. This could be true in places like nursing homes, where the average person might be more susceptible to COVID-19 than the average person in the broader population.

But on a larger scale, heterogeneity typically lowers the herd immunity threshold. At first the virus infects people who are more susceptible and spreads quickly. But to keep spreading, the virus has to move on to people who are less susceptible. This makes it harder for the virus to spread, so the epidemic grows more slowly than you might have anticipated based on its initial rate of growth.

"The first person is going to be likely to infect the people who are most susceptible to begin with, leaving the people who are less susceptible toward the latter half of the epidemic, meaning the infection could be eliminated sooner than you'd expect," Lipsitch said.

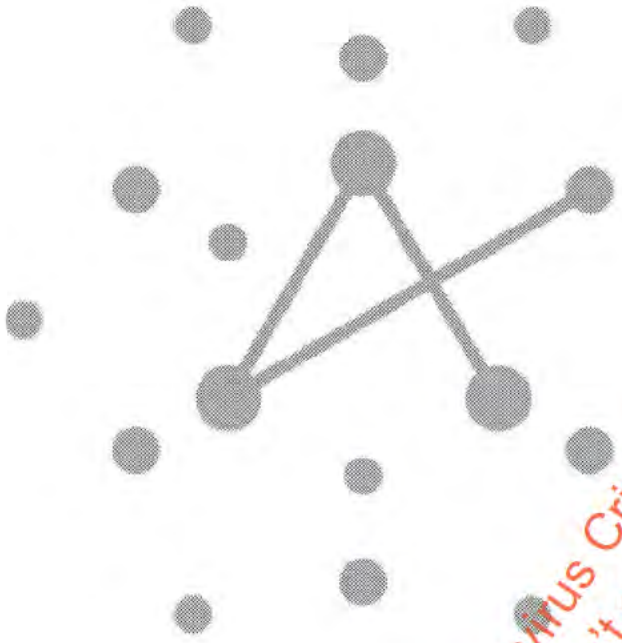
Estimating Heterogeneity

So how much lower is the herd immunity threshold when you're talking about a virus spreading in the wild, like the current pandemic?

According to the standard models, about 60% of the U.S. population would need to be vaccinated against COVID-19 or recover from it to slow and ultimately stop the spread of the disease. But many experts I talked to suspect that the herd immunity threshold for naturally acquired immunity is lower than that.

"My guess would be it's potentially between 40 and 50%," Pitzer said.

Lipsitch agrees: "If I had to make a guess, I'd probably put it at about 50%."



Abstractions

Exploring the world of scientific ideas

Abstractions navigates promising ideas in science and mathematics. Journey with us and join the conversation.

See all Abstractions blog

These are mostly just educated estimates, because it's so hard to quantify what makes one person more susceptible than another. Many of the characteristics you might think to assign someone — like how much social distancing they're doing — can change from week to week.

"The whole heterogeneity problem only works if the sources of heterogeneity are long-term properties of a person. If it's being in a bar, that's not in itself sustained enough to be a source of heterogeneity," Lipsitch said.

Heterogeneity may be hard to estimate, but it's also an important factor in determining what the herd immunity threshold really is. Langwig believes that the epidemiological community hasn't done enough to try and get it right.

"We've kind of been a little sloppy in thinking about herd immunity," she said. "This variability really matters, and we need to be careful to be more accurate about what the herd immunity threshold is."

Some recent papers have tried. In June the journal *Science* published a study that incorporated a modest degree of heterogeneity and estimated the herd immunity threshold for COVID-19 at 43% across broad populations. But one of the study's co-authors, Tom Britton of Stockholm University, thinks there are additional sources of heterogeneity their model doesn't account for.

"If anything, I'd think the difference is bigger, so that in fact the herd immunity level is probably a bit smaller than 43%," Britton said.

Another new study takes a different approach to estimating differences in susceptibility to COVID-19 and puts the herd immunity threshold even lower. The paper's 30 authors, who include Gomes and Langwig, estimate that the threshold for naturally acquired herd immunity to COVID-19 could be as low as 20% of the population. If that's the case, the hardest-hit places in the world may be nearing it.

"We're getting to the conclusion that the most affected regions like Madrid may be close to reaching herd immunity," said Gomes. An early version of the paper was posted in May, and the authors are currently working on an updated version, which they anticipate posting soon. This version will include herd immunity estimates for Spain, Portugal, Belgium and England.

Many experts, however, consider these new studies — not all of which have been peer-reviewed yet — to be unreliable.

> We've kind of been a little sloppy in thinking about herd immunity.
> Kate Langwig, Virginia Polytechnic Institute and State University

In a Twitter thread in May, Dean emphasized that there's too much uncertainty around basic aspects of the disease — from the different values of R_0 in different settings to the effects of relaxing social distancing — to place much confidence in exact herd immunity thresholds. The threshold could be one number as long as a lot of people are wearing masks and avoiding large gatherings, and another much higher number if and when people let their guard down.

Other epidemiologists are also skeptical of the low numbers. Jeffrey Shaman of Columbia University said that 20% herd immunity "is not consistent with other respiratory viruses. It's not consistent with the flu. So why would it behave differently for one respiratory virus versus another? I don't get that."

Miller added, "I think the herd immunity threshold [for naturally acquired immunity] is less than 60%, but I don't see clear evidence that any [place] is close to it."

Ultimately, the only way to truly escape the COVID-19 pandemic is to achieve large-scale herd immunity — everywhere, not just in a small number of places where infections have been highest. And that will likely only happen once a vaccine is in widespread use.

In the meantime, to prevent the spread of the virus and lower that R_0 value as much as possible, distancing, masks, testing and contact tracing are the order of the day everywhere, regardless of where you place the herd immunity threshold.

RELATED:

1. What Other Coronaviruses Tell Us About SARS-CoV-2
2. How Math (and Vaccines) Keep You Safe From the Flu
3. The Animal Origins of Coronavirus and Flu

"I can't think of any decision I'd make differently right now if I knew herd immunity was somewhere else in the range I think it is, which is 40-60%," said Lipsitch.

Shaman, too, thinks that uncertainty about the naturally acquired herd immunity threshold, combined with the consequences for getting it wrong, leaves only one path forward: Do our best to prevent new cases until we can introduce a vaccine to bring about herd immunity safely.

"The question is: Could New York City support another outbreak?" he said. "I don't know, but let's not play with that fire.

Dr. Paul E. Alexander, PhD
Senior Advisor to the Assistant Secretary
For COVID-19 Pandemic Policy
Office of the Assistant Secretary of Public Affairs (ASPA)
US Department of Health and Human Services (HHS)
Washington, DC
[redacted]

Email: [redacted]

From: Alexander, Paul (HHS/ASPA)
Sent: Saturday, July 25, 2020 9:29 PM
To: Caputo, Michael (HHS/ASPA) [redacted]
Subject: RE: NYC 25% antibodies...and aircraft Rosevelt

You got me thinking and I went looking at literature and voila...I found it...top researchers have written that COVID heard can be as low they think as 20%...and NYC is there...and that's why I argue things have slowed...and maybe, just maybe, all we need is 20% antibodies....I could not stop thinking about what I wrote so had to find something...thank god to at least show I aint nuts my man....

See this...read and you will see...see what they wrote in red...thank god!!!!!!!!!!!!!!!!! I did not want to look like a nut ball and if as they think and as I think this may be true, then as they say, several hard hit areas may have hit heard at 20% like NYC...that's my argument....why not consider it????

"If anything, I'd think the difference is bigger, so that in fact the herd immunity level is probably a bit smaller than 43%," Britton said.

Another new study takes a different approach to estimating differences in susceptibility to COVID-19 and puts the herd immunity threshold even lower. The paper's 10 authors, who include Gomes and Langwig, estimate that the threshold for naturally acquired herd immunity to COVID-19 could be as low as 20% of the population. If that's the case, the hardest-hit places in the world may be nearing it.

https://www.quantamagazine.org/the-tricky-math-of-covid-19-herd-immunity-20200630/

The herd immunity thresholds are given by a simple formula (1 − 1/R0) which, in the case of SARSCoV-2, suggests that 60-70% of the population would need be immunized to halt spread considering estimates of R0 between 2.5 and 3. R0 is each person can potentially infect 2.5 people…so to speak. See how it is explained in this paper (in red) being published I found and it damn well backs up my claim…the issue is that the most susceptible people are infected first and this causes the less susceptible people to infection to be left and this reduces the R0 or risk fo spread….if populations are sufficiently heterogeneous….as to susceptibility to infection…that's the key.

A crucial caveat in exporting these calculations to immunization by natural infection is that natural infection does not occur at random. Individuals who are more susceptible or more exposed are more prone to be infected and become immune, which lowers the threshold. In the model, the herd immunity threshold declines sharply when coefficients of variation increase from 0 to 2 and remains below 20% for more variable populations.

Dr. Paul E. Alexander, PhD
Senior Advisor to the Assistant Secretary
For COVID-19 Pandemic Policy
Office of the Assistant Secretary of Public Affairs (ASPA)
US Department of Health and Human Services (HHS)
Washington, DC
[redacted]
[redacted]
Email: [redacted]

From: Caputo, Michael (HHS/ASPA) [redacted]
Sent: Saturday, July 25, 2020 1:40 PM
To: Alexander, Paul (HHS/ASPA) [redacted]
Subject: Re: NYC 25% antibodies...and aircraft Rosevelt

How can this be researched and proven true or false?

Michael R. Caputo

Assistant Secretary for Public Affairs
US Health and Human Services

Work Cell: [redacted]

DELIBERATIVE, PRE-DECISIONAL MATERIALS
FOUO - FOR OFFICIAL USE ONLY

On Jul 25, 2020, at 12:11 PM, Alexander, Paul (HHS/ASPA) [redacted] wrote:

Folks, my initial look now at the final case numbers for the aircraft carrier Rosevelt was about 5000 crew with approximately 1000 infected and 1 death.

This suggests an infection rate of 20%. Question remains why didn't everyone on carrier get infected. And here again this 20% infection rate rears it's head. Seems that this 20% rate is related to COVID. I am doing some more looking into this. The 20% infection rate is important and may suggest that 20% is needed for herd immunity to box this virus in, and not 70%.

We might be able to control this virus if we aggressively lock granny down, maybe even quarantine. Acutely regulate nursing homes for safety, message the younger people to behave responsibly, ensure all the PPE and hospitals are ramped up, and unleash the rest of the healthy society to get to our 20%. We might be able to defeat this even before vaccine is on deck. This 20% is nagging me and I wanted to share my thoughts and provide some evidence to spurn a debate.

Paul

Sent from my iPhone

On Jul 25, 2020, at 7:52 AM, Alexander, Paul (HHS/ASPA) [redacted] wrote:

> The 20% number keeps emerging and it is interesting and there may be something there...
>
> The Diamond Princess which was the first cruise ship that became news, eventually is I remember had 3700 passengers (paying and crew) and 711 got infected...so 20%...why would a closed ship where no one got on or off, why did the virus not infect all of the ship? Seemed only 20% got infected...I think 14 died. But the issue is the virus stopped at 20%.

H1N1 in 2009/2010 under the prior administration infected 60-65 million Americans, 20-25% of population. 20 K died but again, now H1N1 is roughly stable, and influenza virus. Again 20% infected.

I find this 20% number keeps popping up and just as a disease detective, it is something to think about. Herd may be 20% for COVID...I don't know yet. but sharing...

Dr. Paul E. Alexander, PhD
Senior Advisor to the Assistant Secretary
For COVID-19 Pandemic Policy
Office of the Assistant Secretary of Public Affairs (ASPA)
US Department of Health and Human Services (HHS)
Washington, DC
[redacted]

Email: [redacted]

From: Alexander, Paul (HHS/ASPA)
Sent: Friday, July 24, 2020 11:22 AM
To: Caputo, Michael (HHS/ASPA) [redacted]; McKeogh, Katherine (OS/ASPA) [redacted]; Oakley, Caitlin B. (OS/ASPA) [redacted] Hensley, Gordon (HHS/ASPA) [redacted]; Murphy, Ryan (OS/ASPA) [redacted]; Traverse, Brad (HHS/ASPA) [redacted] Stephen Hahn [redacted] Wagner, John (FDA/OC) [redacted]; Brennan, Patrick (OS/ASPA) [redacted] Pratt, Michael (OS/ASPA) [redacted]
Subject: FW: NYC 25% antibodies...raises issue of maybe herd is not 70% but this virus stops/slows at 20-25%

Michael, I am arguing based on this 25% antibody in NYC, not the entire state of NY, that if this is so, that it may be, it just may be, that the 62-70% herd threshold we typically want in aclosed population for the pathogen to be boxed in and shut down, that it may be that this virus needs just 20-25% antibody to be slowed or even stopped...we need to consider this possibility given the dramatic drop in infection in NYC and deaths when it drove the nightmare in April-June...NYC accounts for 40% of deaths in the US...as well it may be that there will be an upsurge but all the ingredients of a spike in NYC is there yet it has not happened and no way a city that is in dissaray public health wise can tamp this down in a few weeks...no way...so something has happened there...and this 25% antibody level if it can be replicated in other samples, then it may be that the herd number is low...and it may be that it will be best if we open up and flood the zone and let the kids and young folk get infected as we acutely lock down the elderely and at risk folk...but use the strong and well in the society to get infected and get to that 25%....maybe as we wait for a vaccine and therapeutics, we may be to get 25% antibodies ourselves by natural immunity....natural exposure....

This cannot be discounted...we have ot think out of the box...

https://www.nydailynews.com/coronavirus/ny-coronavirus-cuomo-20200427-qsk3wltz5vgdjoz2rcif57abti-story.html

Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold

Authors: M. Gabriela M. Gomes[1,2,3]*, Rodrigo M. Corder[4], Jessica G. King[5], Kate E. Langwig[6], Caetano Souto-Maior[7], Jorge Carneiro[8], Guilherme Gonçalves[9], Carlos Penha-Gonçalves[8], Marcelo U. Ferreira[4], Ricardo Aguas[10].

[1] Liverpool School of Tropical Medicine, Liverpool, United Kingdom.

[2] Department of Mathematics and Statistics, University of Strathclyde, Glasgow, United Kingdom.

[3] Centro de Matemática da Universidade do Porto, Porto, Portugal.

[4] Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, Brazil.

[5] Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, United Kingdom.

[6] Department of Biological Sciences, Virginia Tech, Blacksburg, VA, USA.

[7] Laboratory of Systems Genetics, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

[8] Instituto Gulbenkian de Ciência, Oeiras, Portugal.

[9] Unidade Multidisciplinar de Investigação Biomédica, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal.

[10] Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.

*Correspondence to: gabriela.gomes@strath.ac.uk.

Abstract: As severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spreads, the susceptible subpopulation is depleted causing the incidence of new cases to decline. Variation in individual susceptibility or exposure to infection exacerbates this effect. Individuals that are more susceptible or more exposed tend to be infected earlier, depleting the susceptible subpopulation of those who are at higher risk of infection. This selective depletion of susceptibles intensifies the deceleration in incidence. Eventually, susceptible numbers become low enough to prevent epidemic growth or, in other words, the herd immunity threshold (HIT) is reached. Although estimates vary, simple calculations suggest that herd immunity to SARS-CoV-2 requires 60-70% of the population to be immune. By fitting epidemiological models that allow for heterogeneity to SARS-CoV-2 outbreaks across the globe, we show that variation in susceptibility or exposure to infection reduces these estimates. Accurate measurements of heterogeneity are therefore of paramount importance in controlling the COVID-19 pandemic.

One Sentence Summary: Models that curtail individual variation in susceptibility or exposure to infection overestimate epidemic sizes and herd immunity thresholds.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in China in late 2019 and spread worldwide causing the ongoing pandemic of coronavirus disease (COVID-19). As of 06 May 2020, more than 3.5 million cases have been confirmed and almost 250,000 died (*1*). Scientists throughout the world have engaged with governments, health agencies, and with each other, to address this emergency. Mathematical models have been central to important decisions concerning contact tracing, quarantine, and social distancing, to mitigate or suppress the initial pandemic spread (*2*). Successful suppression, however, leaves populations at risk to resurgent waves due to insufficient acquisition of immunity. Models have thus also addressed longer term SARS-CoV-2 transmission scenarios and the requirements for continued adequate response (*3*). This is especially timely as countries begin to relax lockdown measures that have been in place over recent weeks with varying levels of success in tackling national outbreaks.

Here we demonstrate that individual variation in susceptibility or exposure (connectivity) accelerates the acquisition of immunity in populations due to selection by the force of infection. More susceptible and more connected individuals have a higher propensity to be infected and thus are likely to become immune earlier. Due to this *selective immunization*, heterogeneous populations require less infections to cross their herd immunity thresholds (HITs) than homogeneous (or not sufficiently heterogeneous) models would suggest. We integrate continuous distributions of susceptibility or connectivity in otherwise basic epidemic models for COVID-19 and show that as the coefficient of variation (CV) increases from 0 to 4, the herd immunity threshold declines from over 60% (*4*, *5*) to less than 10%. Measures of individual variation are urgently needed to narrow the estimated ranges of HITs and plan accordingly.

SARS-CoV-2 transmission in heterogeneous populations

SARS-CoV-2 is transmitted primarily by respiratory droplets and modelled as a susceptible-exposed-infectious-recovered (SEIR) process.

Individual variation in susceptibility is integrated as a continuously distributed factor that multiplies the force of infection upon individuals as

$$\dot{S}(x) = -\lambda x S(x), \quad \dot{E}(x) = \lambda x S(x) - \delta E(x), \quad \dot{I}(x) = \delta E(x) - \gamma I(x), \qquad (1)$$

where $S(x)$ is the number of individuals with susceptibility x, $E(x)$ and $I(x)$ are the numbers of individuals who originally had susceptibility x and became exposed and infectious, δ is the rate of progression from exposed to infectious, γ is the rate of recovery or death, and $\lambda = (\beta/N)\int[\rho E(x) + I(x)]\,dx$ is the average force of infection upon susceptible individuals in a population of size N. The basic reproduction number is

$$R_0 = \langle x \rangle (\beta/N)(\rho/\delta + 1/\gamma), \qquad (2)$$

where ρ is a factor measuring the infectivity of individuals in compartment E in relation to those in I, and $\langle x \rangle$ is the mean susceptibility factor at epidemic onset. Prior to the epidemic, susceptibility is described by a probability density function $q(x)$ with mean 1 and $CV = \langle (x-1)^2 \rangle$ explored as a parameter. The effective reproduction number (R_{eff}, also denoted by R_e or R_t by other authors) is a time-dependent quantity obtained by multiplying R_0 by the susceptibility of the population over time.

Figure 1 depicts model trajectories fitted to suppressed epidemics in Italy and Austria, assuming coefficients of variation 1 and 3. The difference in the size of second waves between the two levels of variation is substantial. In the case of Italy, where suppression was less successful, the

pandemic appears mostly resolved when $CV = 3$. However, a large second wave (or a series of smaller waves, depending on possible containment strategies) remains in the horizon when $CV = 1$. Countries where suppression of the initial outbreak was more successful, such as Austria, have acquired less immunity and therefore the potential for future transmission in the respective populations remains naturally larger. However, also in these situations, expectations for the potential of subsequent waves is much reduced by variation in susceptibility to infection.

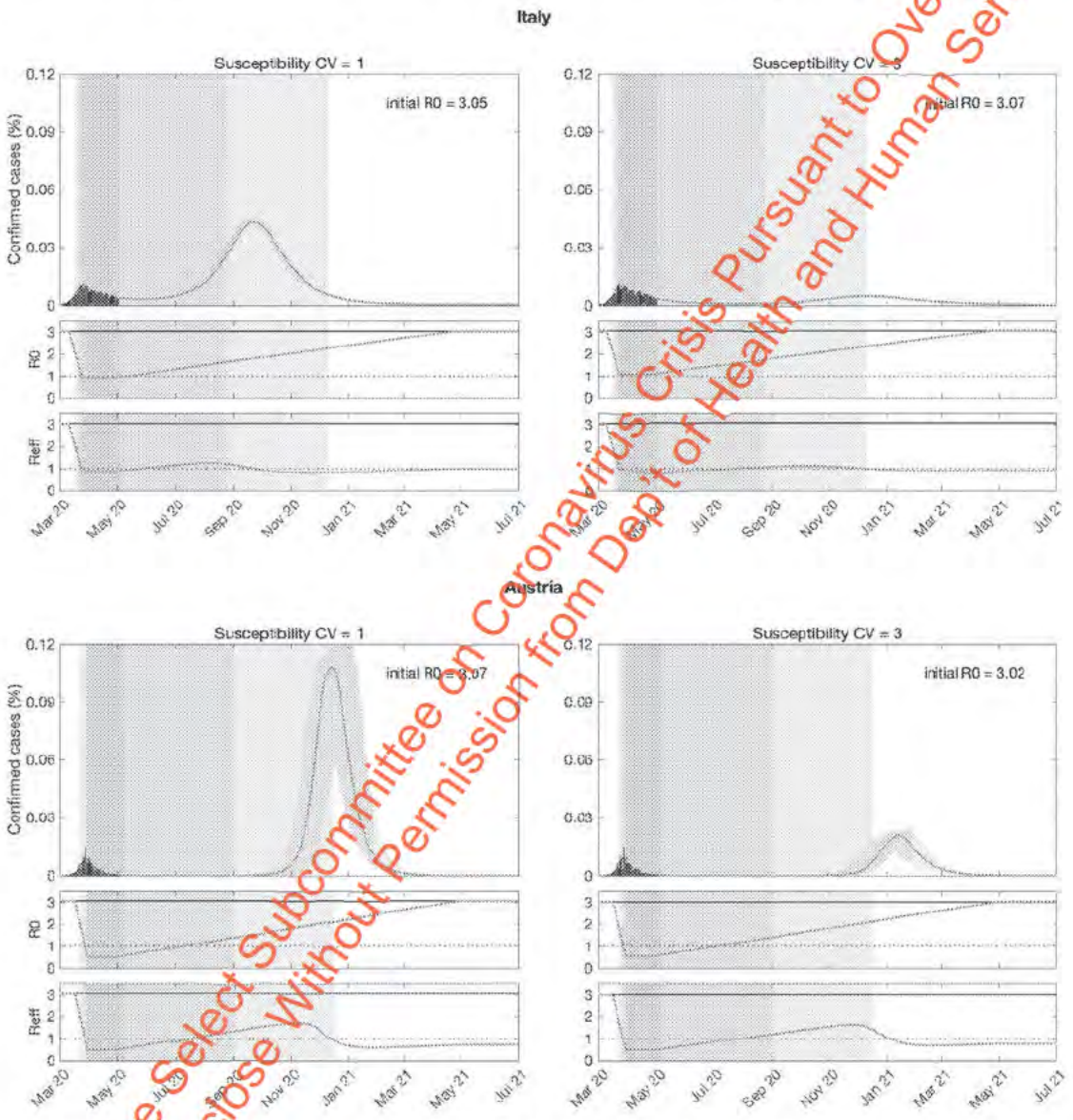


Figure 1: The effect of variation in susceptibility to infection on the size of epidemics. Suppressed wave and subsequent dynamics in Italy and Austria. Blue bars are confirmed new cases and overlaid red bars represent deaths. Basic (R_0) and effective (R_{eff}) reproduction numbers are displayed on bottom panels. Blue shades represent social distancing (intensity reflected in R_0 trends and shade density). Susceptibility factors were implemented as gamma distributions. Consensus parameter values (Materials and Methods): $\delta = 1/4$ per day; $\gamma = 1/4$ per day; and $\rho = 0.5$. Fraction of infected individuals identified as positive (reporting fraction): $p = 0.1$. R_0 and social distancing parameters were estimated by Bayesian inference as described in Supplementary Materials. Curves represent median model predictions from 10^4 posterior samples. Orange shades represent 95% credible intervals.

In a directly transmitted infectious disease, such as COVID-19, variation in exposure to infection is primarily governed by patterns of connectivity among individuals. We incorporate this in the system (Equation 1) by adding variation in infectivity and assuming a positive correlation between susceptibility and infectivity. Formally this corresponds to modifying the force of infection as $\lambda = (\beta/N)(\int x[\rho E(x) + I(x)]\,dx/\int xq(x)\,dx)$ and the basic reproduction number as

$$R_0 = (\langle x^2\rangle/\langle x\rangle)(\beta/N)(\rho/\delta + 1/\gamma), \quad (3)$$

where $\langle x\rangle$ and $\langle x^2\rangle$ are the first and second moments of the distribution $q(x)$ prior to the epidemic. Applying this model to the epidemics in Italy and Austria (Figure 2) leads to similar results to those obtained when variation was in susceptibility to infection.

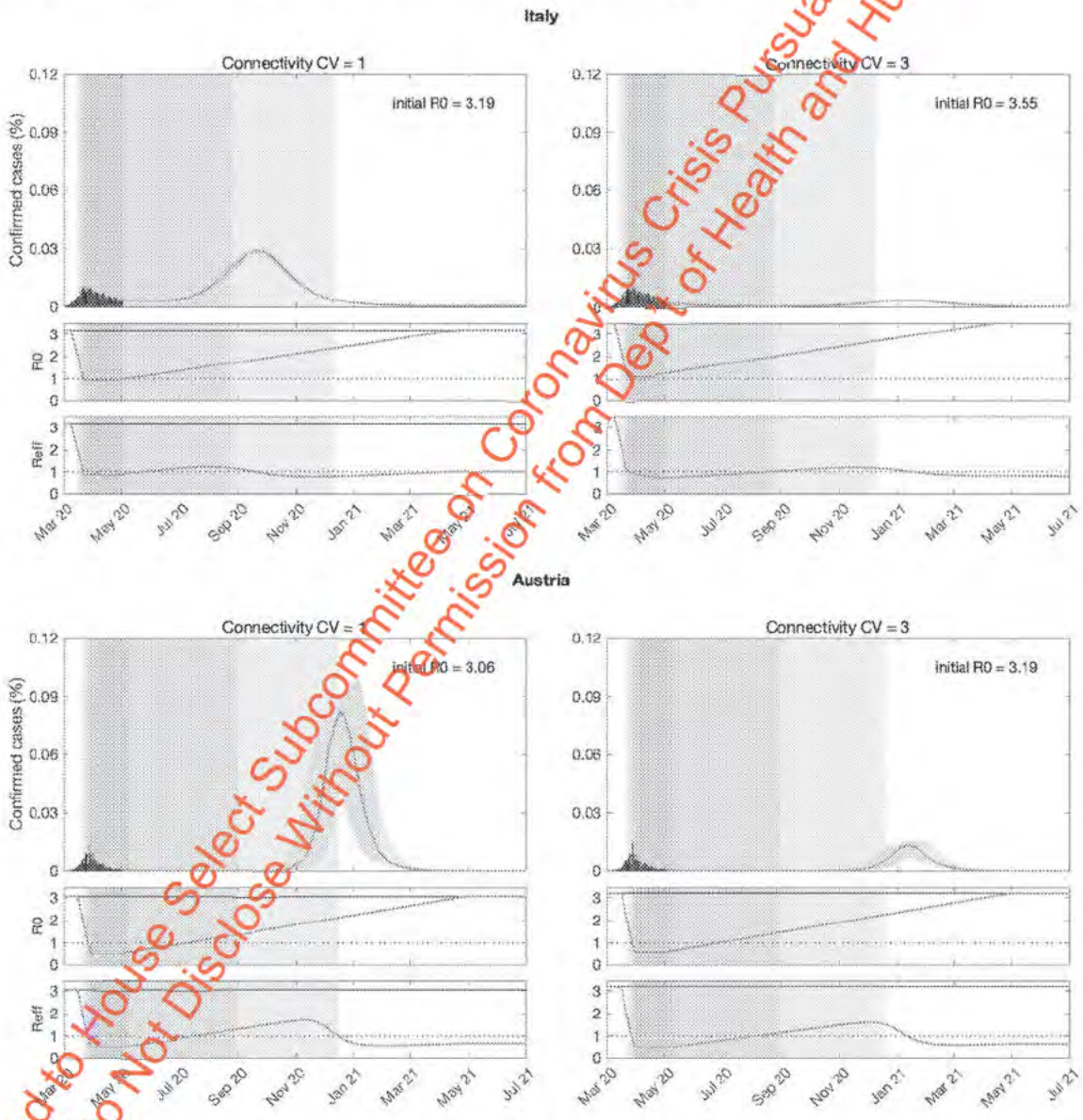


Figure 2: The effect of variation in exposure to infection on the size of epidemics. Suppressed wave and subsequent dynamics in Italy and Austria. Blue bars are confirmed new cases and overlaid red bars represent deaths.

Basic (R_0) and effective (R_{eff}) reproduction numbers are displayed on bottom panels. Blue shades represent social distancing (intensity reflected in R_0 trends and shade density). Connectivity factors were implemented as gamma distributions. Consensus parameter values (Materials and Methods): $\delta = 1/4$ per day; $\gamma = 1/4$ per day; and $p = 0.5$. Fraction of infected individuals identified as positive (reporting fraction): $p = 0.1$. R_0 and social distancing parameters were estimated by Bayesian inference as described in Supplementary Materials. Curves represent median model predictions from 10^4 posterior samples. Orange shades represent 95% credible intervals.

The herd immunity threshold

Individual variation in risk of acquiring infection is under selection by the force of infection, whether individual differences are due to biological susceptibility, physical exposure, or a combination of the two traits. Selection results in the removal of the most at-risk individuals from the susceptible pool as they become infected and eventually recover (some die). This selective acquisition of infection and immunity results simultaneously in decelerated epidemic growth and accelerated induction of immunity in the population. The herd immunity threshold (HIT) defines the percentage of the population that needs to be immune to reverse epidemic growth and prevent future waves. Figure 3 shows the expected downward trends in the HIT for SARS-CoV-2 as the coefficients of variation of the gamma distributed susceptibility or exposure are increased between 0 and 4 (to assess robustness to changing the type of distribution see Figure S22 for equivalent plots with lognormal distributions). While herd immunity is expected to require 60-70% of a homogeneous population to be immune given an R_0 between 2.5 and 3, these percentages drop to the range 10-20% for CVs between 2 and 4. Therefore, a critically important question is: how variable are humans in their susceptibility and exposure to SARS-CoV-2? Hitherto, there is no definite answer to this question.

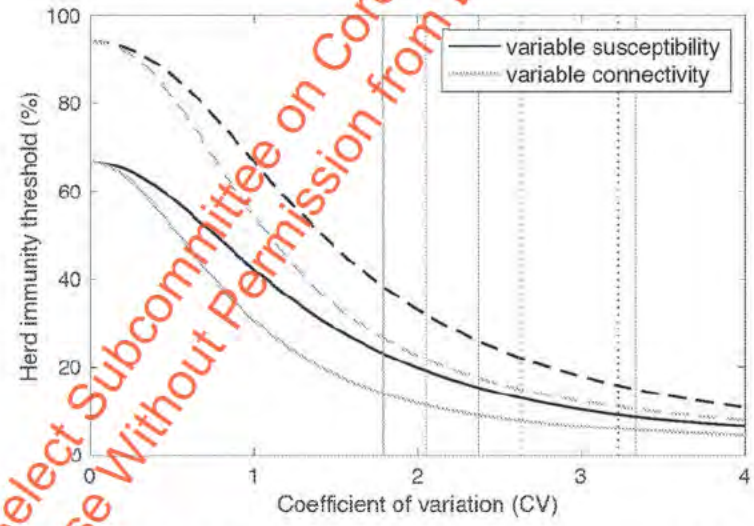


Figure 3: Herd immunity threshold with variation in susceptibility and exposure to infection. Curves generated with the model (Equation 1) with gamma distributed susceptibility (black) or connectivity (gray) assuming $R_0 = 3$: (solid) herd immunity threshold; (dashed) final size of uncontrolled epidemic. Vertical lines indicate coefficients of individual variation for several infectious diseases according to literature: (solid green) susceptibility or exposure to malaria [Amazon 1.79 (*6*), Africa 2.05 (*7*)]; (solid blue) susceptibility or exposure to tuberculosis [Portugal 2.37, Brazil 3.33 (*8*)]; (dotted red) infectiousness for SARS-CoV-1 [Singapore 2.62, Beijing 2.64 (*9*)]; (dotted black) infectiousness for SARS-CoV-2 [3.22 (*10*)].

As the pandemic unfolds evidence will accumulate in support of low or high coefficients of variation, but soon it will be too late for this to impact public health strategies unless we act

pragmatically. We searched the literature for estimates of individual variation in the propensity to acquire or transmit several infectious diseases including COVID-19 and overlaid these estimates as vertical lines in Figure 3. CV estimates are mostly comprised between 2 and 4, a range where naturally acquired immunity to SARS-CoV-2 may place populations over the HIT once as few as 10-20% of its individuals are immune. This depends, however, on which specific transmission traits are variable and how much the trait variants are distributed.

Variation in infectiousness was critical to attribute the scarce and explosive outbreaks to superspreaders when the SARS-CoV-1 emerged in 2002 (*9*), but infectiousness does not respond to selection as susceptibility or exposure do. Models with individual variation in infectiousness perform equivalently to homogeneous versions when implemented deterministically (Figure S21). They diverge when stochasticity is added in the sense that disease extinction becomes more likely and outbreaks become rarer and more explosive (*9-11*), but this an entirely different phenomenon to that presented in this paper.

Among the estimates of individual variation plotted in Figure 3, those corresponding to SARS-CoV viruses, with coefficients of variation in the range 2.6-3.2, have been described as variation in individual infectiousness (*9*, *10*), but the way authors describe superspreaders is suggestive that higher infectiousness may stem from higher connectivity with other individuals who may be susceptible. This would support the scenarios displayed in Figure 2 with $CV = 3$ for connectivity, although little is known about how this might have been modified by social distancing.

Discussion

The concept of *herd immunity* is most commonly used in the design of vaccination programs (*12, 13*). Defining the percentage of the population that must be immune to cause infection incidences to decline, herd immunity thresholds constitute convenient targets for vaccination coverage. In idealized scenarios of vaccines delivered at random and individuals mixing at random, herd immunity thresholds are given by a simple formula $(1 - 1/R_0)$ which, in the case of SARS-CoV-2, suggests that 60-70% of the population would need be immunized to halt spread considering estimates of R_0 between 2.5 and 3. A crucial caveat in exporting these calculations to immunization by natural infection is that natural infection does not occur at random. Individuals who are more susceptible or more exposed are more prone to be infected and become immune, which lowers the threshold (*14*). In our model, the herd immunity threshold declines sharply when coefficients of variation increase from 0 to 2 and remains below 20% for more variable populations. The amplitude of the decline depends on what property is heterogeneous and how it is distributed but the downwards trend is robust (Figures 3 and S22).

Heterogeneity in the transmission of respiratory infections has traditionally focused on variation in exposure summarized into age-structured contact matrices. Besides overlooking differences in susceptibility given exposure, the aggregation of individuals into age groups curtails coefficients of variation with important downstream implications. We calculated CV for the landmark POLYMOD matrices (*15*, *16*) and obtained values between 0.3 and 0.5. Recent studies of COVID-19 integrated contact matrices with age-specific susceptibility to infection (structured in three levels) (*17*) or with social activity (three levels also) (*18*) which, again, resulted in coefficients of variation less than 1. We show that models with coefficients of variation of this magnitude would appear to differ only moderately from homogeneous approximations when compared with those that incorporate CVs between 2 and 3, as estimated for a variety of

infectious diseases (Figure 3) and supported by detailed mobility data in the city of Portland, Oregon, USA (*19*) (we obtained an estimate rounding $CV = 2$ based on data extracted with WebPlotDigitizer). It is therefore crucial that variation in susceptibility and exposure to infection is included in epidemic models at the finest resolution of individuals. This has required agent-based models which are computationally intensive and not amenable to mathematical treatment (*19*). Here, we introduce mathematical formalisms that enable the entire individual variation to be captured while maintaining the analytical tractability of the simplest homogeneous models. This is especially relevant when dealing with major crises such as the current pandemic where optimal strategies rely on a capacity to quickly rationalize the best compromise between protecting health and safeguarding the economy. The larger the individual variation, the more optimistic the public health prognostics and the milder the required health policies.

Interventions themselves have potential to manipulate individual variation. Current social distancing measures may be argued to either increase or decrease variation in exposure, depending on the roles of different functional strata in societies and the compliance of individuals who are normally more highly connected in relation to the average. Datasets that describe connectivity patterns before and during movement restrictions, such as those in (*17*) could, in principle, inform relevant changes in distributions of individual connectivity but surveys must be applied on representative samples of the population and the information cannot be collapsed into age-group averages. A deeper understanding of the putative patterns is crucial not only to develop more accurate predictive models, but also to refine control strategies and to interpret data resulting from prevalence studies and serological surveys.

An analysis of the outbreak on board the Diamond Princess cruise ship reported a cumulative infected percentage of 17% (*20*). Seroprevalences estimated from various settings are currently widespread, but reportedly between less than 1 and just over 20%, including estimates from Kobe, Japan (3.3%) (*21*) and Guilan province, Iran (22%) (*22*). While seropositivity estimates are limited by epidemiological context and current estimates are undoubtedly affected by testing uncertainties, our results suggest that some estimated values may be closer to reaching herd immunity thresholds than otherwise predicted, if populations were sufficiently heterogeneous. Worth nothing, however, that these estimates may have been offset by the social distancing measures.

Given current uncertainties, a high level of pragmatism may be required in incorporating results from serological surveys into policy decisions. We have assumed that infection elicits persistent adaptive immunity. This assumption is justified by encouraging reports on animal models and humans recovered from SARS-CoV-2 infection, even though volatile immunity has not been ruled out yet. Our results are robust as long as recovered individuals remain immune for several months. Any test that allows for retrospectively detecting past infections is therefore a convenient tool for monitoring the prevalence and distribution of individuals who may have acquired immunity. It would be imperative to conduct repeated serological studies in representative samples of the population (*23*) especially as control measures are relaxed, not necessarily to imply that antibodies themselves are neutralizing but to identify past infection and potential for immune protection. Given a percent positivity in an initial survey, the curve traced by subsequent measurements could indicate if and how rapidly a population is moving towards the herd immunity threshold, and simultaneously advise which control measures should be enforced.

References and Notes:

1. World Health Organization, Coronavirus disease 2019 (COVID-19) Situation Report – 107 (WHO, 2020).
2. N. M. Ferguson, *et al.*, Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand (Imperial College COVID-19 Response Team, 2020). 10.25561/77482.
3. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* 10.1126/science.abb5793 (2020).
4. J. T. Wu, K. Leung, G. M. Leung, Nowcasting are forecasting the potential domestic and international spread of the 2019-nCov outbreak originating in Wuhan, China: a modelling study. *Lancet* **395**, 689-697 (2020).
5. K. O. Kwok, F. Lai, W. I. Wei, S. Y. S. Wong, J. Tang, Herd immunity – estimating the level required to halt the COVID-19 epidemics in affected countries. J. Infect. 10.1016/j.jinf.2020.03.027 (2020).
6. R. M. Corder, M. U. Ferreira, M. G. M. Gomes, Modelling the epidemiology of residual Plasmodium vivax malaria in a heterogeneous host population: a case study in the Amazon Basin. *PLOS Comput. Biol.* **16**, e1007377 (2020).
7. D. L. Smith, J. Dushoff, R. W. Snow, S. I. Hay, The entomological inoculation rate and its relation to the prevalence of Plasmodium falciparum infection in African children. *Nature* **438**, 492-495 (2005).
8. M. G. M. Gomes, J. F. Oliveira, A. Bertolde, D. Ayabina, T. A. Nguyen, E. L. Maciel, R. Duarte, B. H. Nguyen, P. B. Shete, C. Lienhardt C, Introducing risk inequality metrics in tuberculosis policy development. *Nat. Commun.* **10**, 2480 (2019).
9. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355-359 (2005).
10. A. Endo, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, S. Abbott, A. J. Kucharski, S. Funk, Estimating the overdispersion in COVID-19 transmission using outbreak sizes outsize China [version 1; peer review: awaiting peer review]. *Wellcome Open Res.* **5**, 76 (2020).
11. C. Donnat, S. Holmes, Modeling the heterogeneity in COVID-19's reproduction number and its impact on predictive scenarios. arXiv:2004.05272 (2020).
12. G. Gonçalves, Herd immunity: recent uses in vaccine assessment. *Expert Rev. Vaccines* **7**, 1493-1506 (2008).
13. P. Fine, K. Eames, D. L. Heymann, "Herd immunity": a rough guide, *Clin. Infect. Dis.* **52**, 911-916 (2011).
14. M. J. Ferrari, S. Bansal, L. A. Meyers, O. N. Bjornstad, Network frailty and the geometry of herd immunity. *Proc. R. Soc. B* **273**, 2743-2748 (2006).
15. J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J.

Edmunds. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Med.* **5**, e74 (2008).

16. K. Prem. A. R. Cook, M. Jit, Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Comput. Biol.* **13**, e1005697 (2017).
17. J. Zhang, *et al.*, Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* 10.1126/science.abb8001 (2020).
18. T. Britton, F. Ball, P. Trapman, The disease-induced herd immunity level for COVID-19 is substantially lower than the classical herd immunity level. arXiv:2005.03085 (2020).
19. S. Eubank, *et al.*, Modelling disease outbreaks in realistic urban social networks. *Nature* **429**, 180-184 (2004).
20. K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell, Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Euro Surveill.* **25**, 2000180 (2020).
21. A. Doi, *et al.*, Seroprevalence of novel coronavirus disease (COVID-19) in Kobe, Japan. medRxiv: 10.1101/2020.04.26.20079822 (2020).
22. M. Shakiba, *et al.*, Seroprevalence of COVID-19 virus infection in Guilan province, Iran. medRxiv: 10.1101/2020.04.13.20079244 (2020).
23. M. F. Silveira, *et al.*, Repeated population-based surveys of antibodies against SARS-CoV-2 in Southern Brazil. medRxiv: 10.1101/2020.05.01.20087205 (2020).
24. W. E. Wei, Z. Li, C. J. Chiew, S. E. Yong, M. P. Toh, V. J. Lee, Presymptomatic Transmission of SARS-CoV-2 — Singapore, January 23–March 16, 2020. MMWR Morb Mortal Wkly Rep [Internet]. 2020 Apr 10 [cited 2020 May 4];69(14):411–5. Available from: http://www.cdc.gov/mmwr/volumes/69/wr/mm6914e1.htm?s_cid=mm6914e1_w
25. K. K. W. To, O. T. Y. Tsang, W. S. Leung, A. R. Tam, T. C. Wu, D. C. Lung, *et al.*, Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect. Dis.* **20**, 565-74 (2020).
26. M. M. Arons, K. M. Hatfield, S. C. Reddy, A. Kimball, A. James, J. R. Jacobs, *et al.*, Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *N. Engl. J. Med.* 10.1056/NEJMoa2008457 (2020).
27. X. He, *et al.*, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* 10.1038/s41591-020-0869-5 (2020).
28. S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, *et al.*, The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* 10.7326/M20-0504 (2020).
29. Q. Li, *et al.*, Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199-1207 (2020).
30. J. Zhang, *et al.*, Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: A descriptive and modelling study. *Lancet Infect. Dis.* S1473-3099(20)30230-9 (2929).

31. H. Nishiura, N. M. Linton, A. R. Akhmetzhanov, Serial interval of novel coronavirus (COVID-19) infections. *Int. J. Infect. Dis.* **93**, 284-6 (2020).

32. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, J. Shaman, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* 10.1126/science.abb3221 (2020).

33. P. Stapor, D. Weindl, B. Ballnus, S. Hug, C. Loos, A. Fiedler, *et al.*, PESTO: Parameter EStimation TOolbox. *Bioinformatics* **34**, 705-707 (2018).

34. C. I. Jarvis, *et al.*, Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. medRxiv: 10.1101/2020.03.31.20049023 (2020).

35. K. Prem, Y. Liu, T. W. Russell, A. D. Kucharski, R. M. Eggo, N. Davies, The effect of control strategies to reduce social mixing on outcomes of COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Health* 10.1016/S2468-2667(20)30073-6.

36. H. Tian, *et al.*, An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 10.1126/science.abb6105.

37. S. Flaxman, *et al.*, Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries (Imperial College COVID-19 Response Team, 2020). 10.25561/77731.

38. Kucharski *et al.*, Effectiveness of isolation, testing, contact tracing and physical distancing on reducing transmission of SARS-CoV-2 in different settings. medRxiv: 10.1101/2020.04.23.20077024 (2020).

39. H. Salje, *et al.*, Estimating the burden of SARS-CoV-2 in France. pasteur-02548181 (2020).

40. L. Di Domenico, G. Pullano, C. E. Sabbatini, P.-Y. Boelle, V. Colizza, Expected impact of lockdown in Île-de-France and possible exit strategies. medRxiv: 10.1101/2020.04.13.20063933 (2020).

Acknowledgements: We thank Jan Hasenauer (Institute of Computational Biology, Helmholtz Zentrum München, München, Germany) for helpful discussions.

Funding: RMC and MUF receive scholarships from the Conselho Nacional de Desenvolvimento Científico e Tecnológio (CNPq), Brazil. CSM was supported by the Intramural Research Program of the NIH, The National Heart Lung and Blood Institute.

Author contributions: MGMG conceived the study and wrote the first draft, RA, RMC and JGK performed the analyses, all authors wrote the paper.

Competing interests: Authors declare no competing interests.

Data and materials availability: All data is available in the main text or the supplementary materials.