medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

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

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

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

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

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.

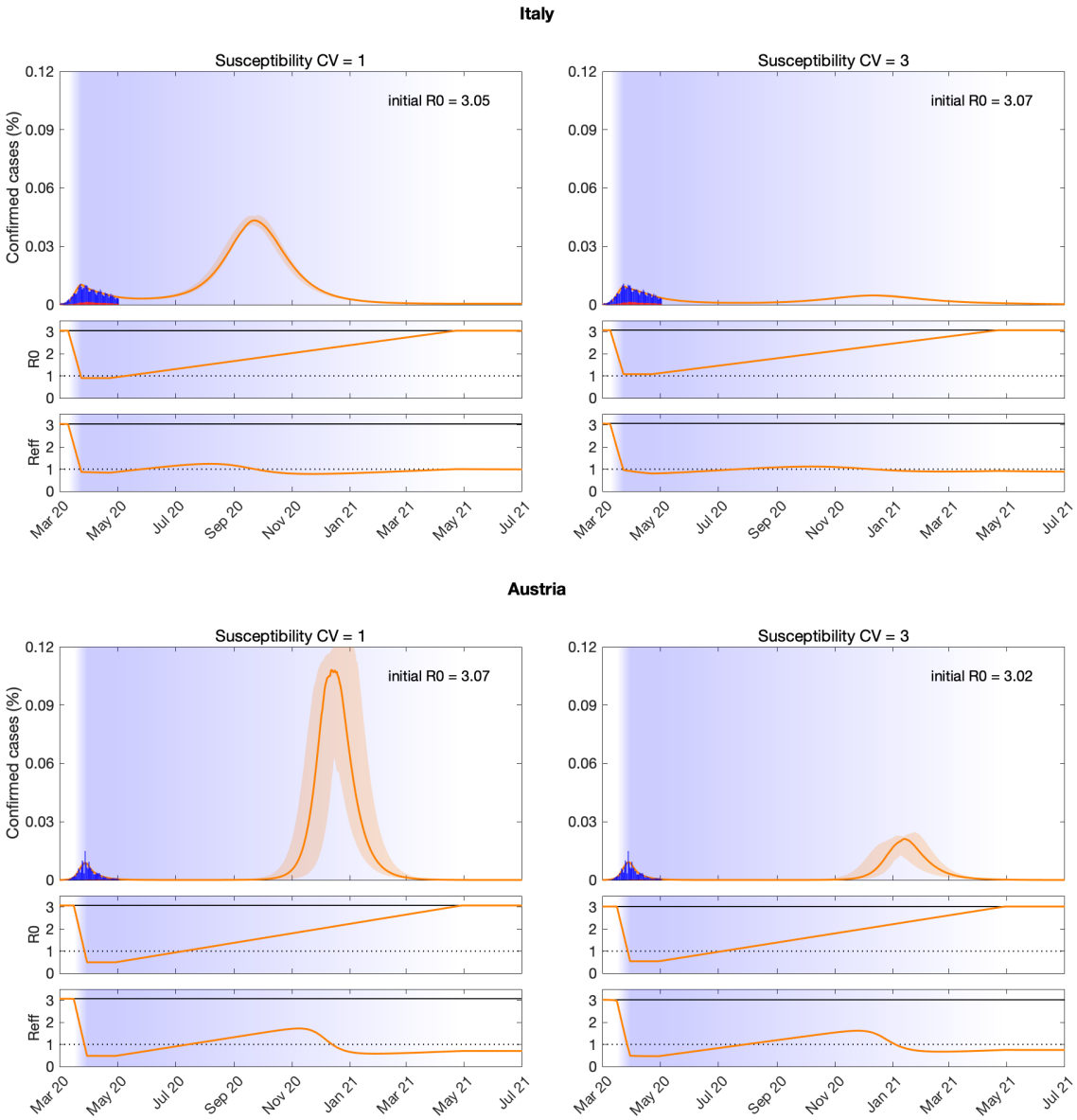

Figure 1: The effect of variation in susceptibility to infection on the size of epidemics. Suppressed wave and subsequent dynamics in Italy and Austria. Blue bars are confirmed new cases and overlaid red bars represent deaths. Basic (R_0) and effective (R_{eff}) reproduction numbers are displayed on 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

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

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.

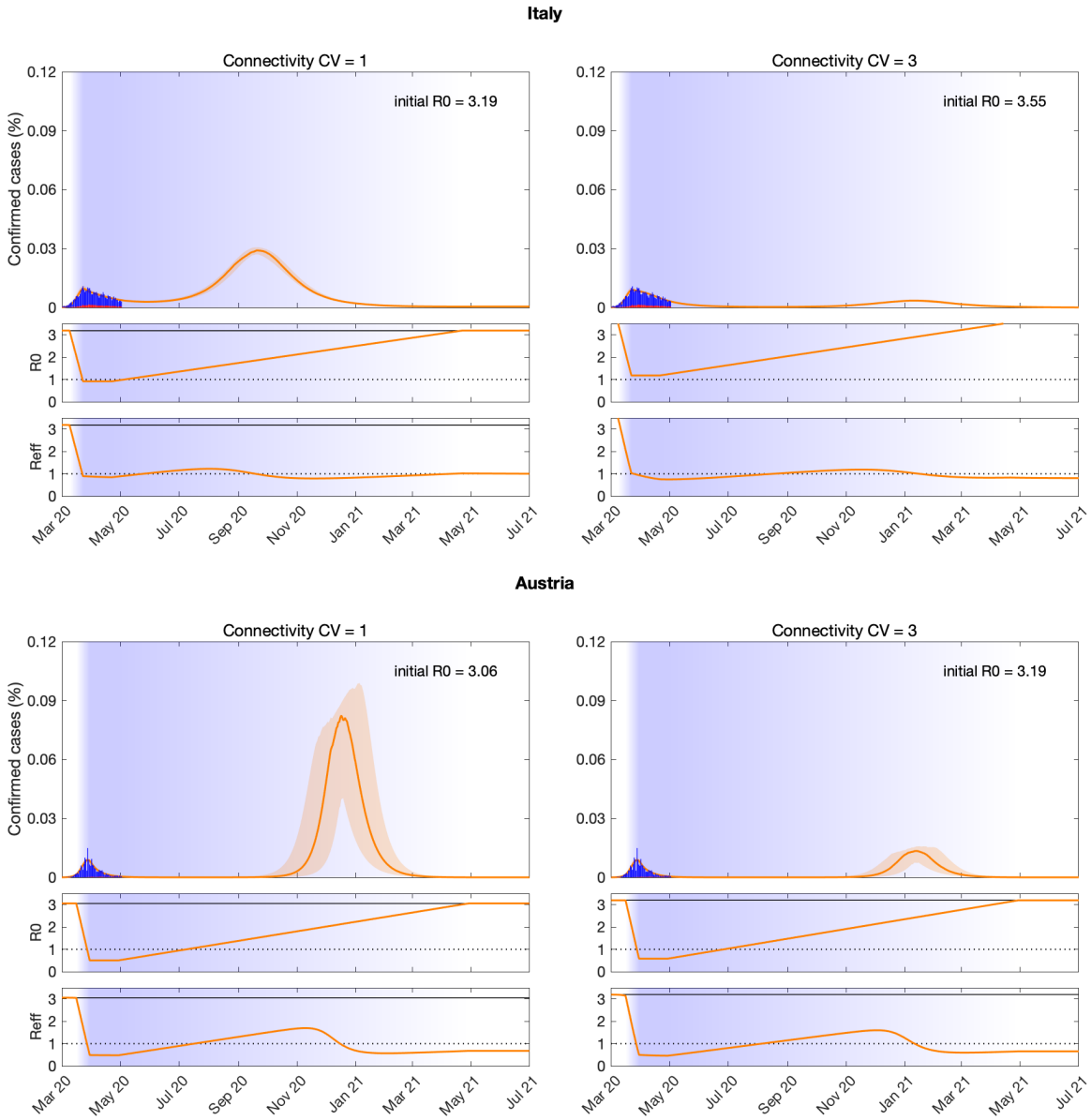

Figure 2: The effect of variation in exposure to infection on the size of epidemics. Suppressed wave and subsequent dynamics in Italy and Austria. Blue bars are confirmed new cases and overlaid red bars represent deaths.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 $\rho = 0.5$. 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.

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.

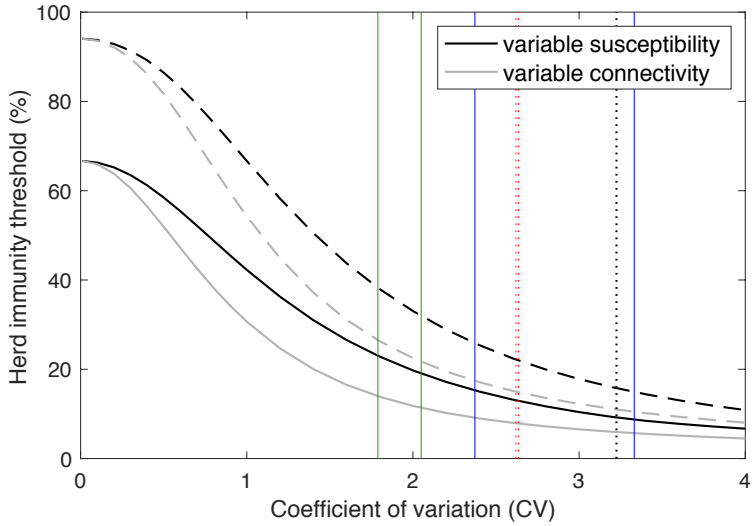

Figure 3: Herd immunity threshold with variation in susceptibility and exposure to infection. Curves generated with the model (Equation 1) with gamma distributed susceptibility (black) or connectivity (gray) assuming $R_0 = 3$: (solid) herd immunity threshold; (dashed) final size of uncontrolled epidemic. 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

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.27.20081893.this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.