Why herd immunity to COVID-19 is reached much earlier than thought - update

I showed in my May 10th article [Why herd immunity to COVID-19 is reached much earlier than thought](#) that inhomogeneity within a population in the susceptibility and in the social-connectivity related infectivity of individuals would reduce, in my view probably very substantially, the herd immunity threshold (HIT), beyond which an epidemic goes into retreat. I opined, based on my modelling, that the HIT probably lay somewhere between 7% and 24%, and that evidence from Stockholm County suggested it was around 17% there, and had been reached. Mounting evidence supports my reasoning.[1]

I particularly want to highlight an important paper published on July 24th "Herd immunity thresholds estimated from unfolding epidemics" (Aguas et al.).[2] The author team is much the same as that of the earlier theoretical paper (Gomes et al.[3]) that prompted my May 10th article.

Aguas et al. used a SEIR compartmental epidemic model modified to allow for inhomogeneity, similar to the model I used although they also considered further variants. They fitted their models to scaled daily new cases data from four European countries for which disaggregated regional case data was also readily available. In all cases they found a better fit from their models incorporating heterogeneity to the standard homogeneous assumption SEIR model. They found that:

> *"Homogeneous models systematically fail to fit the maintenance of low numbers of cases after the relaxation of social distancing measures in many countries and regions".*

Aguas et al. estimate the HIT at between 6% and 21% for the countries in their analysis – very much in line with the range I suggested in May. They also found that their HIT estimates were robust to various changes in their model specification. By contrast, if the population were homogeneous or were vaccinated randomly, the estimated HIT would have been around 65% –80%, in line with the classical formula, $\{1 - 1/R_0\}$, where R_0 is the epidemic's basic reproduction number.[4]

Aguas et al.'s Figure 3, reproduced below, shows how the HIT reduces with increasing variation either in susceptibility (given exposure) or in connectivity, which affects both an individual's susceptibility (via altering exposure to infection) and infectivity. The coloured dots and vertical lines show the inferred position of each of the four countries they analysed in each of these (separately modelled) cases.
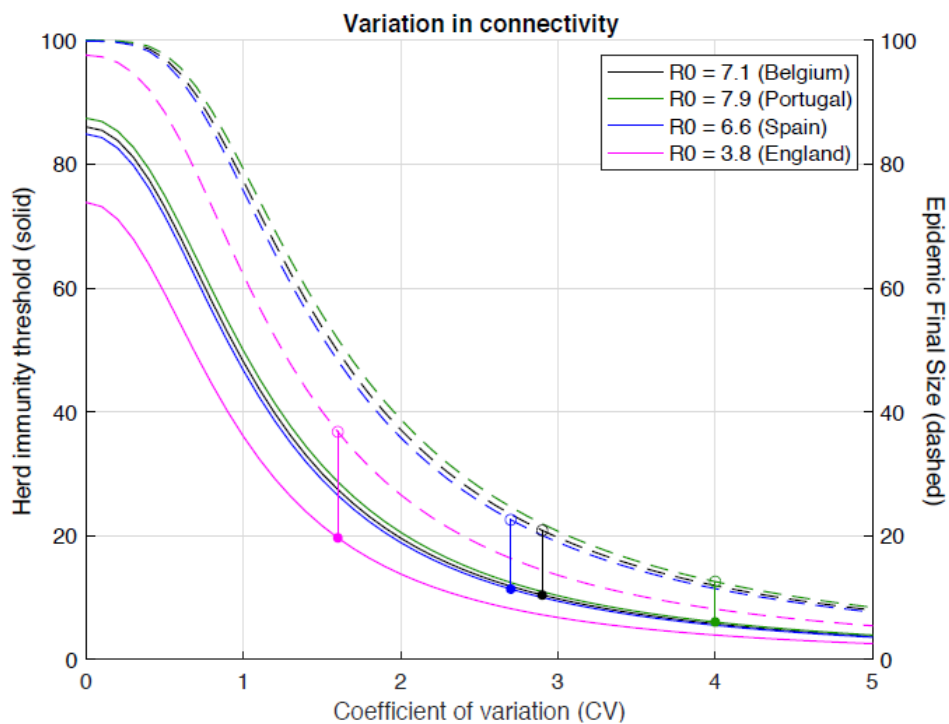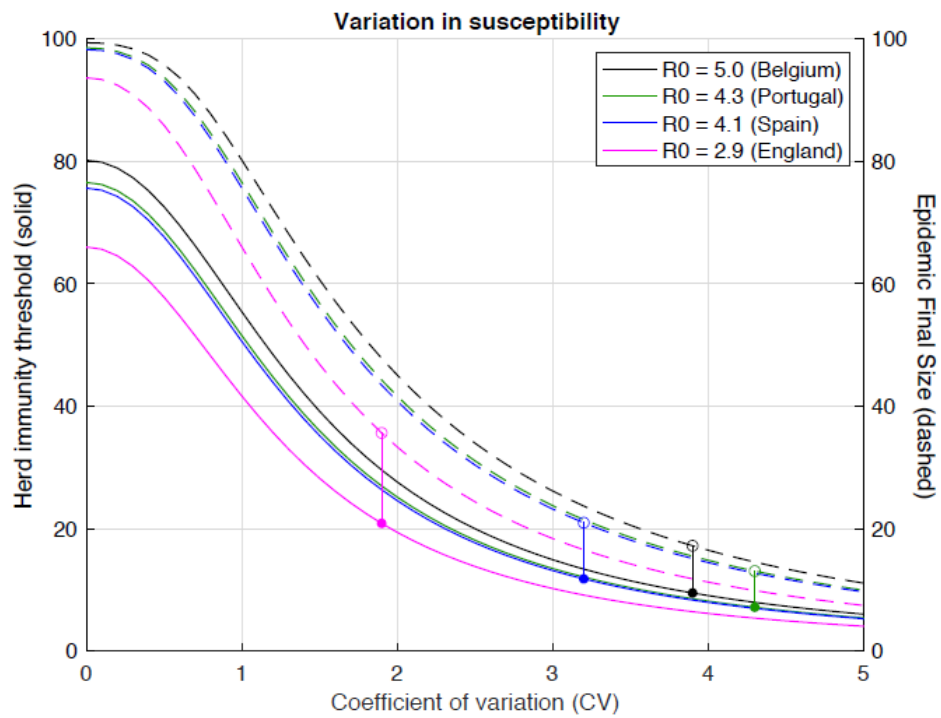
Aguas et al. Fig. 3 **Herd immunity threshold with gamma-distributed susceptibility (top) or connectivity related exposure to infection (bottom)**. Curves generated with the SEIR model (Equation 1-4) assuming values of R0 estimated for the study countries assuming gamma-distributed: susceptibility [top]; connectivity (and hence exposure to infection) [bottom]. Herd immunity thresholds (solid curves) are calculated according to the formula $1 - (1/R_0)^{1/(1 + CV^2)}$ for heterogeneous susceptibility and $1 - (1/R_0)^{1/(1 + 2\,CV^2)}$ for heterogeneous connectivity. Final sizes of the corresponding unmitigated epidemics are also shown (dashed).

As Aguas et al. say in their Abstract:

These findings have profound consequences for the governance of the current pandemic given that some populations may be close to achieving herd immunity despite being under more or less strict social distancing measures.

The underlying reason for the classical formula being inapplicable is, as they say:

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 by natural infection, heterogeneous populations require less infections to cross their herd immunity threshold than suggested by models that do not fully account for variation.

The Imperial College COVID-19 model (Ferguson et al.[5]) is a prime example of one that does not adequately account for variation in individual susceptibility and connectivity.

Aguas et al. point out that consideration of heterogeneity in the transmission of respiratory infections has traditionally focused on variation in exposure summarized into age-structured contact matrices. They showed that, besides this approach typically ignoring differences in susceptibility given virus exposure, the aggregation of individuals into age groups leads to much lower variability than that they found from fitting the data. The resulting models appeared to differ only moderately from homogeneous approximations.

A key reason for variability in susceptibility to COVID-19 given exposure to the SARS-CoV-2 virus causing is that the immune systems of a substantial proportion (35% to 80%) of unexposed individuals have T-cells, circulating antibodies or other components that are cross-reactive to SARS-CoV-2 and can be expected to provide substantial resistance to it.[6][7][8][9] Such components likely arise from past exposure to common cold or other coronaviruses, or to influenza.[10] Not being specific to SARS-CoV-2, and typically not being antibodies, such immune system components are not normally detected in seroprevalence or other tests for immunity to SARS-CoV-2.

I will end with a follow up to my June 28th article focusing on Sweden. In it, I concluded that it was likely the HIT had been surpassed in the three largest Swedish regions, and in the country as a whole, by the end of April notwithstanding that COVID-19-specific antibodies had only been detected in 6.3% of the population.[11] I also projected, based on their declining trend, that total COVID-19 deaths would likely only be about 6,400. Subsequent developments support those conclusions. Swedish COVID-19 deaths have continued to decline, notwithstanding a return to more travel and less social distancing, and are now down to 10 to 15 a day. According to the latest Financial Times analysis,[12] excess mortality in Sweden over 2020 to date was 5,500, or 24%. That is only about half the excess mortality percentage for the UK (45%), Italy (44%) and Spain (56%), and is also lower than for France (31%), the Netherlands (27%) and Switzerland (26%), despite Sweden not having imposed a lockdown or shut primary schools. Moreover, total mortality in Sweden over the last 24 months is now lower than over the previous 24 months, despite an upward trend in the old age population.

Nicholas Lewis   27 July 2020

Further update 31 July 2020

Another important paper has now been published on the role of inhomogeneity within a population in the social-connectivity related susceptibility and infectivity of individuals and in biological susceptibility: "Persistent heterogeneity not short-term overdispersion determines herd immunity to COVID-19" (Tkachenko et al.)[13]. The paper's mathematical/statistical analysis is excellent.[14] Their method of estimating the role of population inhomogeneity in lowering the herd immunity threshold seems reasonable in principle.[15]

However, they estimated the effect of inhomogeneity during lockdowns, and assumed that the effect is the same in other circumstances. But a key effect of social distancing measures, including public events bans, bar and restaurant closures, etc. as well as full lockdowns, is to heavily reduce the number of contacts by the most connected people that are capable of transmitting infection. For people with few social connections, such measures will have a proportionately much smaller effect. So the effect in more normal circumstances of population inhomogeneity in social-connectivity, which appears to be more important than inhomogeneity in biological susceptibility, is bound to be underestimated, quite possibly substantially, by their approach.

Nevertheless, their best fit to New York City COVID-19 data during lockdown gives an estimate of an inhomogeneity factor[16] λ of 4.5.[17] An alternative estimation method based on a cross-sectional regression across US States gives a λ estimate of 5.3.[18]

A middle of the range λ value of 4.9 implies a HIT of 20% if $R_0 = 3.0$ (16.4% if $R_0 = 2.4$; 24.6% if $R_0 = 4$; 28.0% if $R_0 = 5$). It also equates, if all the inhomogeneity is social-connectivity related, to a coefficient of variation (CV)[19] of 1.4 – which is the geometrical mean of the two CV values (1 and 2) that I used in my original article.

Estimating λ from fits to the NYC or Chicago data prior to lockdown implies much higher CV estimates, in the range 2.4 to 2.9 if all inhomogeneneity is social-connectivity related, in non-lockdown circumstances. The corresponding estimates for nine of the worst hit US States range from 1.9 to 3.4.[20]

---

[1] One example, further supporting my superspreader-based evidence of variability in social connectivity, is Miller et al: Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel *medRxiv* 22 May 2020 https://doi.org/10.1101/2020.05.21.20104521 This paper shows that 1-10% of infected individuals caused 80% of infections. That points to variability in social connectivity related susceptibility and infectivity quite likely being higher than I modelled .

[2] Aguas, R. and co-authors: Herd immunity thresholds estimated from unfolding epidemics" *medRxiv* 24 July 2020 https://doi.org/10.1101/2020.07.23.20160762

[3] Gomes, M. G. M., et al.: Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *medRxiv* 2 May 2020. https://www.medrxiv.org/content/10.1101/2020.04.27.20081893v1

[4] The basic reproduction number of an epidemic, R_0, measures how many people, on average, each infected individual infects at the start of the epidemic. If R_0 exceeds one, the epidemic will grow, exponentially at first. But, assuming recovered individuals are immune, the pool of susceptible individuals shrinks over time and the current reproduction number falls. The proportion of the population that have been infected at the point where the current reproduction number falls to one is the 'herd immunity threshold' (HIT). Beyond that point the epidemic is under control, and shrinks.

[5] Neil M Ferguson et al.: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team Report 9, 16 March 2020, https://spiral.imperial.ac.uk:8443/handle/10044/1/77482

[6] Grifoni, A.et al.: Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* 11420, 2020 https://doi.org/10.1016/j.cell.2020.05.015

[7] Braun, J., et al.: Presence of SARS-CoV-2 reactive T cells in COVID-19 patients and healthy donors. *medRxiv* 22 April 2020 https://www.medrxiv.org/content/10.1101/2020.04.17.20061440v1.

[8] Le Bert, N. et al.: Different pattern of pre-existing SARS-COV-2 specific T cell immunity in SARS-recovered and uninfected individuals. *bioRxiv* 27 May 2020. https://doi.org/10.1101/2020.05.26.115832

[9] Nelde, A. et al.: SARS-CoV-2 T-cell epitopes define heterologous and COVID-19-induced T-cell recognition. *ResearchSquare* 16 June 2020. https://www.researchsquare.com/article/rs-35331/v1

[10] Lee, C., Koohy, H., et al.: CD8+ T cell cross-reactivity against SARS-CoV-2 conferred by other coronavirus strains and influenza virus. *bioRxiv* 20 May 2020. https://doi.org/10.1101/2020.05.20.107292.

[11] Such seroprevalence is likely to significantly understate the proportion of the population who have had COVID-19, since asymptomatic or mild disease often results in undetectably low antibody levels (Long, Q. X. et al.: Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med*. 18 June 2020 https://doi.org/10.1038/s41591-020-0965-6 . Such patients will nevertheless be immune to reinfection (Sekine, K. et al.: Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *bioRxiv* 29 June 2020 https://doi.org/10.1101/2020.06.29.174888).

[12] https://www.ft.com/content/a26fbf7e-48f8-11ea-aeb3-955839e06441. Data updated to 13 July

[13] Tkachenko, A.V. et al.: Persistent heterogeneity not short-term overdispersion determines herd immunity to COVID-19. *medRxiv* 29 July 2020 https://doi.org/10.1101/2020.07.26.20162420

[14] I think its title gives a slightly misleading impression, although that issue is not central to their paper. It is in fact "persistent" heterogeneity that causes "short-term" overdispersion, albeit that over a short period random variability will have a significant influence. I'm unconvinced by their argument that estimating social-connectivity related susceptibility and infectivity from overdispersion in transmission statistics is likely to lead to significant bias, provided that estimation is based on large-scale transmission and not just a few superspreader events.

[15] Doing so involves dependency on an estimate of the infection fatality rate, but their IFR-inferred proportion of the New York City population that had been infected by early June 2020 looks reasonable, based on the NYS survey suggesting 22.7% of NYC residents had been infected by late March and the ratio of cumulative COVID-19 deaths 23 days later.

[16] They term their λ an "immunity factor", but it is only partly related to biological immunity. If the causative inhomogeneity is related to biological immunity, $\lambda = 1 + CV^2$, whereas if it is related to social connectivity (which affects infectivity as well as susceptibility) $\lambda = 1 + 2CV^2$.

[17] They also obtian a similar estimate for Chicago, but based on a much narrower range of data.

[18] Or $\lambda = 4.7$ for a selected subset of States.

[19] The ratio of the standard deviation to the mean of a random variable.

[20] I have excluded NY State data, as their curve for that State shows abnormal behaviour, quite likely due to the early epidemic data being strongly dominated by NY City