Fighting COVID-19: the heterogeneous transmission thesis

Maria Chikina
Department of Computational and Systems Biology
University of Pittsburgh Medical School

Wesley Pegden
Department of Mathematical Sciences
Carnegie Mellon University

Abstract: Minimizing infections and deaths from COVID-19 are not the same thing. While society has some control on the final number of infected individuals through intervention and mitigation strategies, we have much greater control over the age-profile of the final cohort of infected individuals. By ignoring this distinction, strategies which focus on minimizing transmission rates to every extent possible in the entire population could increase deaths among all age groups.

We argue for what we call the ***heterogeneous transmission thesis***: in the response to a highly transmittable infectious disease with highly age-variable mortality rates, death rates (for all age groups) may be minimized by mitigation strategies which *selectively* reduce transmission rates in at-risk populations, while maintaining closer-to-normal transmission rates in low-risk populations.

The heterogeneous transmission thesis

The basic idea of the heterogeneous transmission thesis is simple: at the end of the COVID-19 outbreak, a significant fraction of the world's population (e.g., perhaps 5%-20%, at least) will have been infected. Mitigation strategies can affect this final number of total infections, but probably only by a small multiplicative factor (for example, halving them). On the other hand, mitigation strategies also have the potential to shift the age distribution among eventually infected population. Because the mortality rate from COVID-19 between age groups varies by several orders of magnitude (rather than a small multiplicative factor), this can be a much more powerful approach to reducing overall mortality. In particular, this suggests that strategies which are intended to shift the final age distribution—by reducing the transmission rates among older populations more than among younger populations—could have the potential to save the greatest numbers of lives.

There has already been considerable discussion on the role of overburdened healthcare systems on the potential for a high COVID-19 mortality rate, and the heterogeneous transmission thesis is especially relevant in light of these concerns. Indeed, to minimize hospital overcrowding, the goal is not to minimize the number of COVID-19 *infections* but instead the number of COVID-19 infections *requiring hospitalization*. Like overall mortalities, the COVID-19 hospitalization rate varies orders of magnitude by age group, suggesting that to minimize the the number of hospitalizations, we should seek to shift the age-profile of infections, rather than simply minimizing their total number.

What we are not saying

Realizing that some readers may not digest all of this document before reaching conclusions, we would like in advance to dispel some possible misconceptions:

- **We are not arguing against mitigation efforts.** Indeed, in the mathematical models we present later, **the worst case scenario is when no action is taken to reduce transmission rates**. We do argue that *heterogeneous* mitigation efforts, targeted at a greatest reduction of transmission rates among older populations, while allowing closer-to-normal transmission rates among younger populations, may save the most lives.
- **We are not arguing that the economic costs of mitigation efforts outweigh their benefits**, beyond our Assumption number 3 (discussed below) which posits that mitigation efforts cannot continue indefinitely. While there may be arguments that some mitigation efforts are so economically destructive that in the long term they could cause more excess fatalities than COVID-19 itself, our goal is just to estimate the direct affects of the interventions themselves on COVID-19 deaths.
- **We are not saying that mitigation efforts can ignore the younger population.** Our conclusions concern reducing transmission among the older population, and between the younger and older populations, below the level of transmission within the younger population. But some interactions between younger and older people are unavoidable; for example, some children live with older relatives. Mitigations must account for these interactions. Our model also simplistically assumes that the older population is the only at-risk population. In reality, public policy must be cognizant of the presence of other risk factors across age groups. *Most importantly of all, mitigation strategies for low-risk and high-risk groups must be coordinated; our findings for the success of heterogeneous transmission only hold in regimes where transmission among the high-risk population, and between the low- and high-risk populations, are effectively controlled.*
- **We are not saying that younger people won't die, and thus should be exempt from mitigations.** COVID-19 is a deadly infectious disease which will likely kill many people across a range of age groups, including those at lower risk. However, because the greatest fraction of hospitalizations and deaths will be driven by older populations, and hospital overcrowding will affect severe COVID-19 cases among all age groups, heterogeneous mitigation strategies have the potential to help all groups, by minimizing overall hospitalizations and deaths.
- **We are not saying that mortalities among younger people should be traded for mortalities among older people**. The effect of heterogeneous mitigation strategies on infections in the younger population in our model is primarily to shift the timing of younger infections, so that they occur earlier relative to infections among the older population, rather than to increase their total number. Indeed, depending on the extent to which hospital overcrowding affects mortality rates, our model shows that heterogeneous mitigation strategies, with greater transmission rates for younger populations, can actually minimize mortality for both age groups separately.
- **We are not predicting the future or making specific policy recommendations.** Mathematical models are just that—models. Reality will likely diverge from any specific predictions we make about precise numbers of mortalities, or the precise optimal choices for targeting transmission rates. Our point is a conceptual one: in the presence of an infectious disease with the characteristics of COVID-19, there are parameter regimes where minimizing transmission rates among all age groups will result in more deaths than selectively reducing transmission rates. In particular, our point is simply that *it cannot be taken as a general rule that any step taken to reduce transmission rates will reduce COVID-19 deaths*. Instead, because of how they may affect the age distribution among infections, *some* efforts to reduce transmission rates are likely to actually increase total COVID-19 mortalities.

Acknowledgments

This document has benefited from valuable feedback from many individuals, including Boris Bukh, Forrest Collman, Jordan Ellenberg, Ryan O'Donnell, Russell Schwartz, among several others, who have helped us not only improve the presentation of our results, but suggested a wide range of sensitivity analyses, which we have incorporated.

Contents

Background and assumptions

After emerging in Wuhan, China in December 2019, COVID-19 has confirmed spread to more than 100 counties, with more than 100,000 confirmed cases worldwide as of March 15, 2020, a number rapidly growing every day.

In response, countries have taken a range of dramatic countermeasures, including quarantining cities, closing schools, and generally, reducing social interactions to slow and stem the spread of the virus.

Broadly, it seems that these measures have been effective at significantly slowing the spread of infection. For example, while China experienced thousands of new cases daily for several weeks, under lockdown it (as of March 15) now reports dozens of new cases daily. In Italy, Iran, and South Korea, the growth in the rate of daily new cases has also slowed in recent days. The simple takeaway from this is that countermeasures can reduce the number of transmitted cases. In particular, it seems clear that the R_0 value—the number of new infections caused by each infected individual—has been successfully decreased in several different countries independently—even below 1 in some cases. However, we will argue that approaches which use all available tools to minimize the infection rate as much as possible will actually result in **greater** mortality—more deaths from COVID-19— than selective measures which allow greater transmission rates among low-risk populations. Our conclusion, which we will support with both simple arguments and simple epidemiological models, rests on the following assumptions:

Assumption 1: A modifiable transmission rate

As discussed above, the experience of the hardest hit countries indicates that the R_0 rate of infection can be brought below 1 with responses undertaken by a range of governments and economic systems. Our modeling conclusions will not even require that governments can bring the rate of infection below 1, just that they can decrease it significantly—to below around 1.8 with a 1-year mitigation window, or below 2.1 with 6-month mitigation window, for the high-risk population.

Note that when governments' initial response is to implement drastic measures across the board, this initial period gives an opportunity to veryify the validity of this assumption. If this assumption is false, and COVID-19 has little response even to dramatic isolation measures, than our conclusion about selective measures would not hold; every reduction of infection—in both low-risk and high-risk populations—would likely save lives.

Assumption 2: Complete worldwide containment of COVID-19 is not possible in the short-term; instead, mitigation strategies must be able to survive reintroduction of infection

Despite the success of individual countries in controlling the infection rate of COVID-19, the worldwide spread of COVID-19 and the fact that it is highly-infections in the absence of dramatic containment motivates our second assumption: COVID-19 will not be made extinct, with zero human cases worldwide, in the short-term (on the scale of months to a year), unless it is the case that a significant fraction of individuals have been infected.[1]

In particular, we assume therefore that society may face periodic reinfection events in the short-term; society's mitigation strategies must be able to survive reintroduction of the virus.

Assumption 3: Society will eventually return to near-normal transmission levels

Our third assumption is simply that drastic measures to curb the transmission of infectious diseases cannot continue indefinitely. For our model, we will simply assume that relatively normal economic and social activity must be resumed after 1 year. More precisely, we assume that after a 6 month period, interventions will gradually be relaxed, so that transmission rates of respiratory illnesses will (linearly) increase from their 6-month level to their normal levels at the 1 year mark.[2]

The most important thing about this assumption is the inevitable return of normal transmission levels, rather than the precise time scale of the return. (As we discuss below, if a vaccine with high effectiveness is likely to be developed very soon, that would challenge this assumption.) In our sensitivity analyses below we show that our conclusions also hold if we assume that normal transmission rates are resumed after 2 years (with a linear return from the 18 month point).

Note that our Assumptions 2 and 3 together imply that the epidemic will not end without some meaningful decrease in the susceptible fraction of the population. Under these assumptions, some degree of herd immunity is required for the epidemic to end.

Assumption 4: Mortality from COVID-19 is dramatically higher in older populations

This assumption is consistent with all investigations we know of which have examined the issue. In particular, the mortality rate among individuals over 70 appears to be roughly 50 times the mortality rate for individuals under 50. Moreover, while there is great uncertainty about the absolute mortality rate of COVID-19 due to the limits of current testing regimes, these age-relative mortality rates are remarkably consistent across studies.

Assumption 5: Infection confers immunity

Despite some early reports about some people testing positive, then negative, then positive for COVID-19, current consensus is that these test results were more likely artifacts than true reinfection events. Snopes has a summary of this issue with expert input. A very recent study on rhesus macaques is consistent with the assumption that infection confers immunity.

Are our assumptions reasonable?

Our assumptions depend on social, economic, and geopolitical factors, and they may or may not hold in the various COVID-19 fights faced by governments around the world.

But our assumptions are clear and simple. And, we will argue, they imply that blanket, untargeted measures to reduce transmission of COVID-19 will kill more people—quite possibly, many more people— than efforts which target transmission rates among specific age groups. As such, our assumptions need to be reckoned with. **In the absence of evidence, rejecting these assumptions does not constitute a conservative, "safety first" approach to saving lives**; mistakes about assumptions will have consequences for widespread mortality. Most crucially, **if nations undertake dramatic, universal containment efforts in the belief that containment can be achieved in the short-term and that assumption proves incorrect, millions of people may needlessly die.**

The basic idea

In this section we explain intuition for our thesis, which we will also support below with simple mathematical epidemic models.

The starting point for our explanation is that Assumption 2 suggests that we have limited control of the final number of individuals worldwide who will be affected by COVID-19. This is **not** to say that the number of final infections is not modifiable; for example, it may be reasonable to expect that measures to modify transmission rates could modify the final number of infections by a factor of 2, but not by a factor of 100.

The key insight, then, is that the mortality rate by age group for COVID-19 varies by nearly two orders of magnitude more than our control over the final number of infected individuals. In particular, many more lives will be saved if we **shift the age profile** of the eventually infected population (even at the expense of allowing a greater number of total infections).

The power to shift the age-profile of the eventually-infected population comes from applying containment measures which affect transmission rates differently for different age groups. For example, closing schools affects transmission rates among student- and teacher-aged populations, while curtailing nursing home visits affects transmission rates among older populations. Sharply curtailing all social and economic activity affects transmission rates among all populations, while selectively reducing social and economic activity for targeted age groups can affect the final balance of which age groups are among the eventually affected population.

In short, the counter-intuitive message of our argument is that given our assumptions as laid out above, **overall deaths from COVID-19 might be minimized if transmission rates among younger (e.g., less than 65) populations are kept at near-normal levels,** while transmission rates among older populations are reduced. As we demonstrate with our model in the next section, this is because **keeping transmission rates at normal levels for younger people will dramatically reduce the fraction of COVID-19 cases which result in hospitalization and death;** in particular, the reduction in this ratio will be more consequential than the the corresponding small increase in the number of eventually infected individuals, and a result in a lower **number** of hospitalizations and deaths overall. Indeed, because of the plausible role of hospital overcrowding and resource-shortage on death-rates, we will show in our model that **a greater transmission rate among younger populations should reduce COVID-19 deaths for both younger and older populations**.

The mathematical model

We demonstrate our thesis with a simple dynamic epidemic model, which is a variant of the SIR model. This is a simple dynamic model which tracks three groups in a population over time, the *Susceptible* population, the *Infected* population, and the *Recovered/Removed* population. Transmission, and mortality/recovery rates affect the rates at which individuals move from Susceptible to Infected to Recovered/Removed.

In our variant of this model, we track Susceptible, Infected, Recovered, and Mortalities for each of two age groups (e.g., under/over 65) for a total of 8 population groups. The evolution of the 8 populations are governed by the rates of transmission among and between each population, as well as the recovery and mortality rates for each population.

The R code for our model is available here. A mathematical description of the model can be found here.

Model parameters

Our model involves the following parameters:

- **transmission rates between/within under- and over-65 populations:** These determine the rate at which encounters between infected and susceptible individuals lead to new infections. Due to Assumption 3, we assume that after 1 year, these transmission rates will return to levels that imply an R_0-value of 2.8 in a completely susceptible population.[3]
- **recovery and mortality rates:** We assume a rate of recovery based on a 14-day average recovery time. We assume the mortality rate for our older population is 50 times greater than for our younger population.[4] While the relative mortality rates are much more important for the validity of our conclusions than the absolute rates, we have chosen mortality rates for the two groups which correspond to an overall mortality rate of 0.5%.[5]
- **medical system capacity:** To model the effects of overburdened medical systems, we assume that above a threshold of 500,000 infected cases, mortality increases by a factor of 2.[6]

Note that we have set these parameters roughly based on the U.S. medical system. However, our intention is to illustrate a basic phenomenon which can occur in the presence of a pandemic meeting the basic assumptions we have laid out above. Below, we show that are general conclusions are quite robust to the variations these parameter selections, with the exception that our conclusions **do** depend on Assumption 4: that mortality rates are significantly higher among older populations than younger ones.

Model scenarios

We will present 5 different model scenarios, each with different management of the transmission rate within age groups. In each model, we assume that after 9 months, transmissibility for both groups begins to return linearly over 3 months to a level which would be equivalent to an R_0-value of 2.8 in a completely susceptible population. We allow ourselves to choose the level of transmissibility within age groups before the 9 month point for each model. In particular, for each model, at any time point, there are two transmissibility levels; α is the transmissibility level for within the under-65 population, and β is the level for within the over-65 population and between the two populations. In the discussion here, we discuss values for α and β in terms of R_0 equivalence; that is, α and β set to 2.8 correspond to normal transmission rates, under our assumptions.

In Scenario 0 we will let α and β both be 2.8 for the entire simulation. This corresponds to no mitigation strategies being taken. **Among all scenarios we consider, this results in the greatest number of fatalities.**

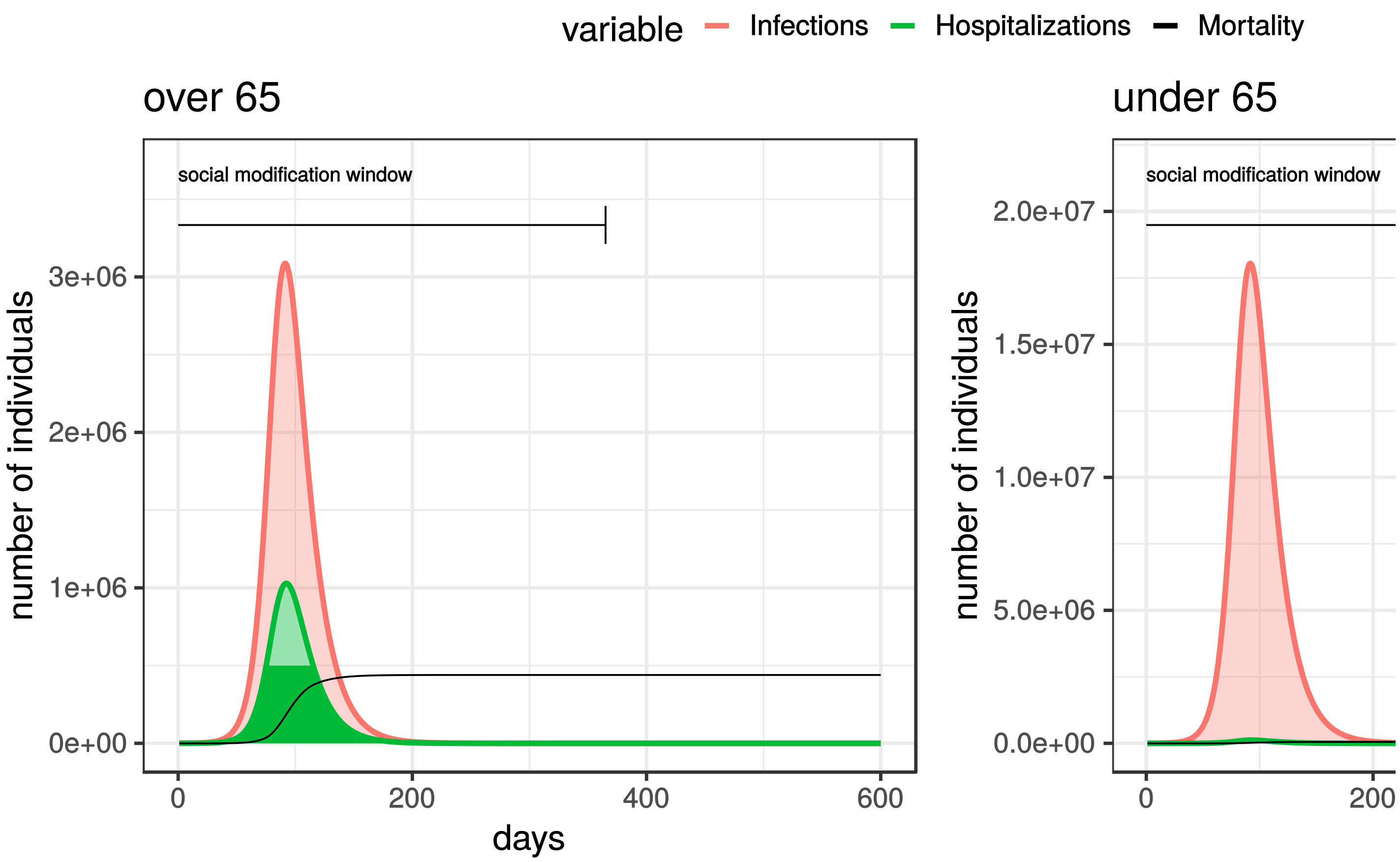
In Scenario 1 we will let α and β begin very low (.9) for 9 months, before returning linearly to normal levels. This corresponds to extreme homogeneous measures being taken on the 9-month time scale. This also results in a very large number of fatalities.

In Scenario 2 we let α and β be 1.8 for 9 months, before increasing to normal levels. 1.8 is chosen because for our parameter regime, this minimizes mortalities among all homogeneous mitigation strategies. Many lives are saved in this scenario, with mortalities dropping by roughly a third.

In Scenario 3 we consider extreme heterogeneous measures. β is controlled .6, while the younger population retains completely normal transmission rates with α at 2.8. Again, transmission rates return to 2.8 linearly between 9 months and a year. Among all scenarios we consider this minimizes the total number of fatalities, with more than a 60% drop from Scenario 0.

Finally, in Scenario 4, we consider a (possibly more realistic) case of heterogeneous measures, where β begins at 1, while α is slightly depressed to 2.4. In this scenario we still see far fewer mortalities than can be achieved by homogeneous measures, with roughly a 50% reduction from Scenario 0.

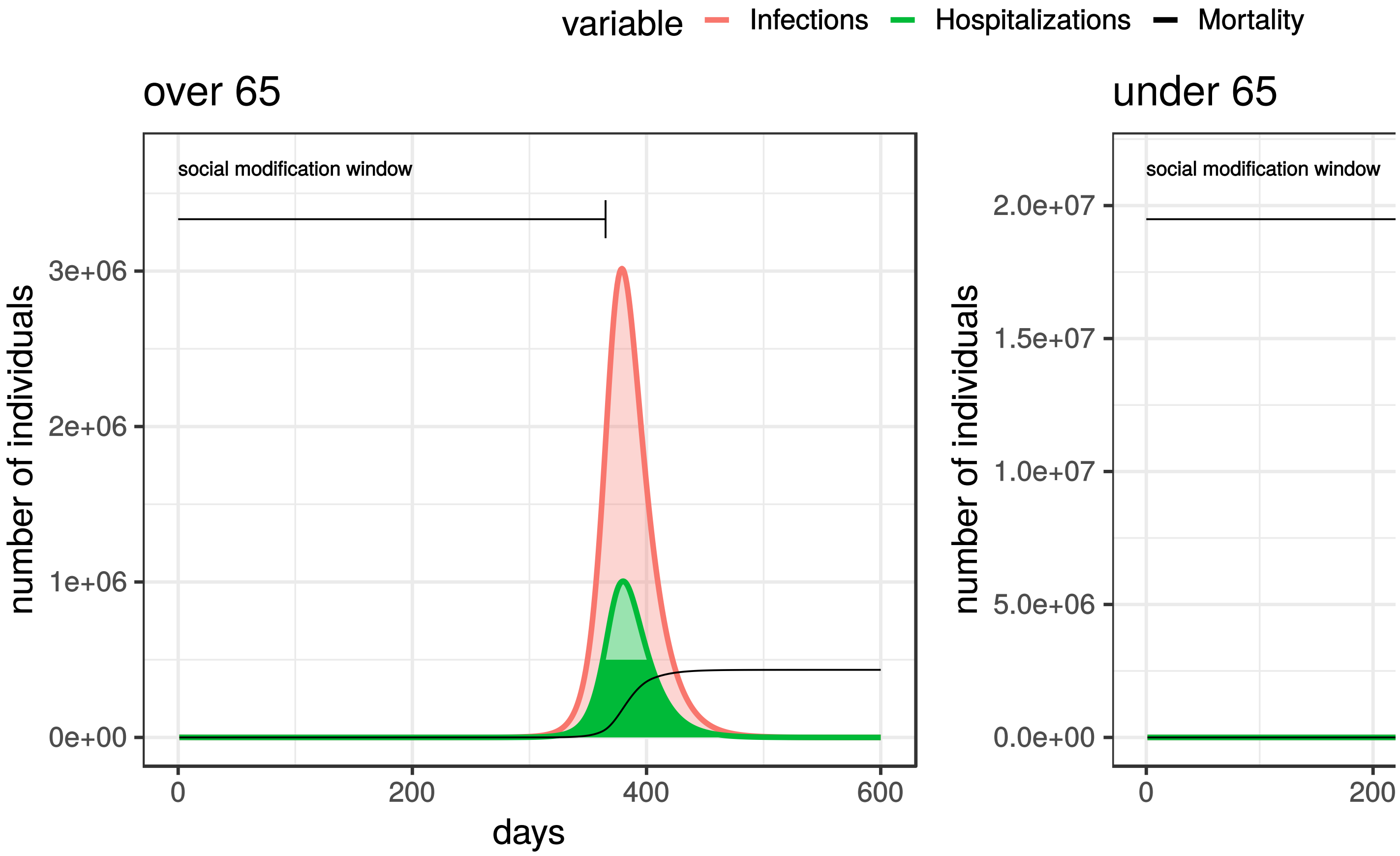
Scenario 0: no interventions



For this scenario, we assume no interventions, so that α are β are left at 2.8. Here we see the greatest number of total hospitalizations and mortalities; overall we see **491,000 mortalities**. The light green region under hospitalization curve shows overcrowding.

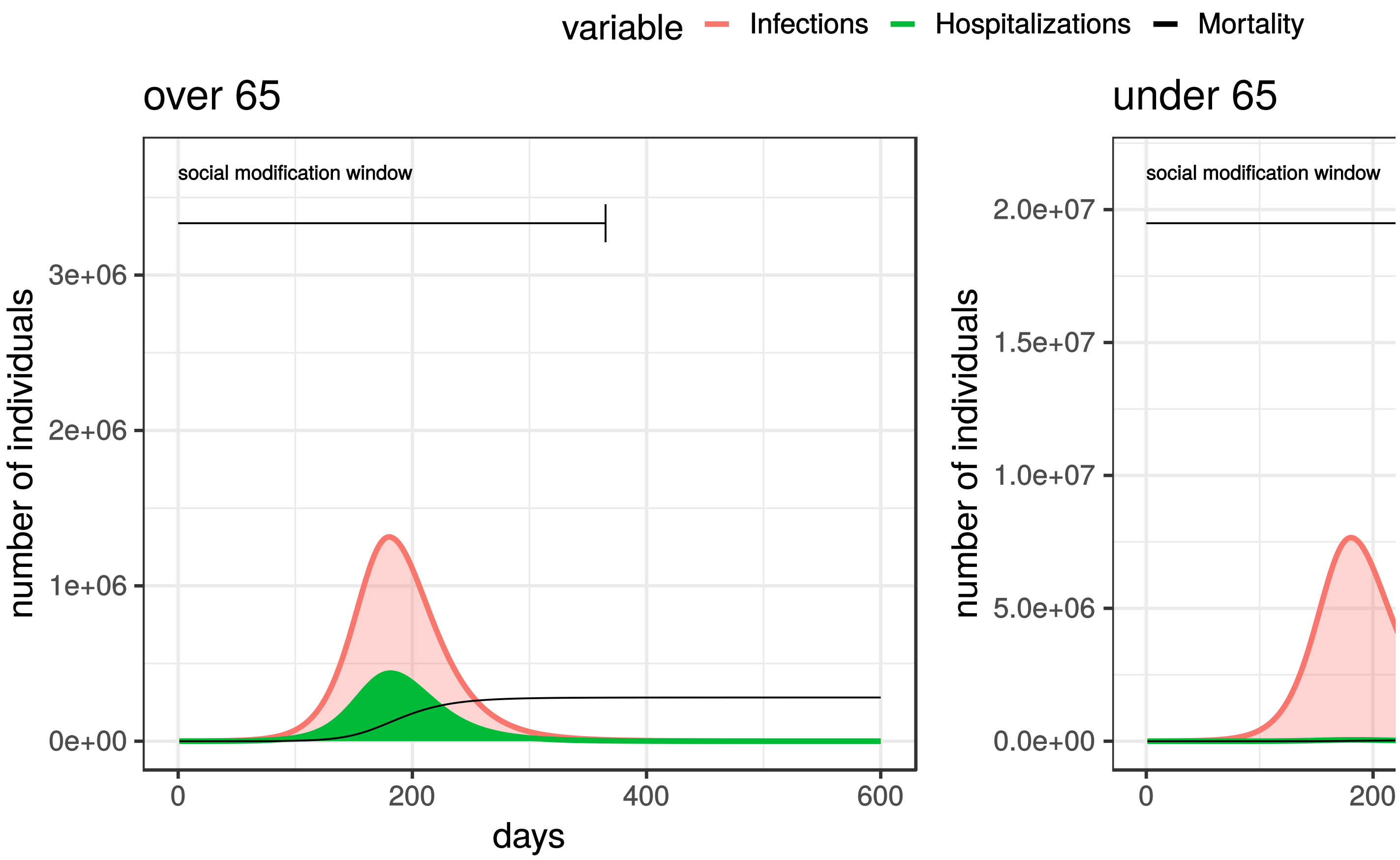
Note that in the under-65 plot in this scenario and those that follow below, it appears that there are very few hospitalizations are deaths. In fact there are more than 50,000 deaths and 500,000 total hospitalizations among the under-65 population in this scenario, but this is nevertheless small relative to the millions of infections which occur in this age group, and thus hospitalizations and deaths simply are dwarfed by the scale of this plot.

Scenario 1: extreme homogeneous measures



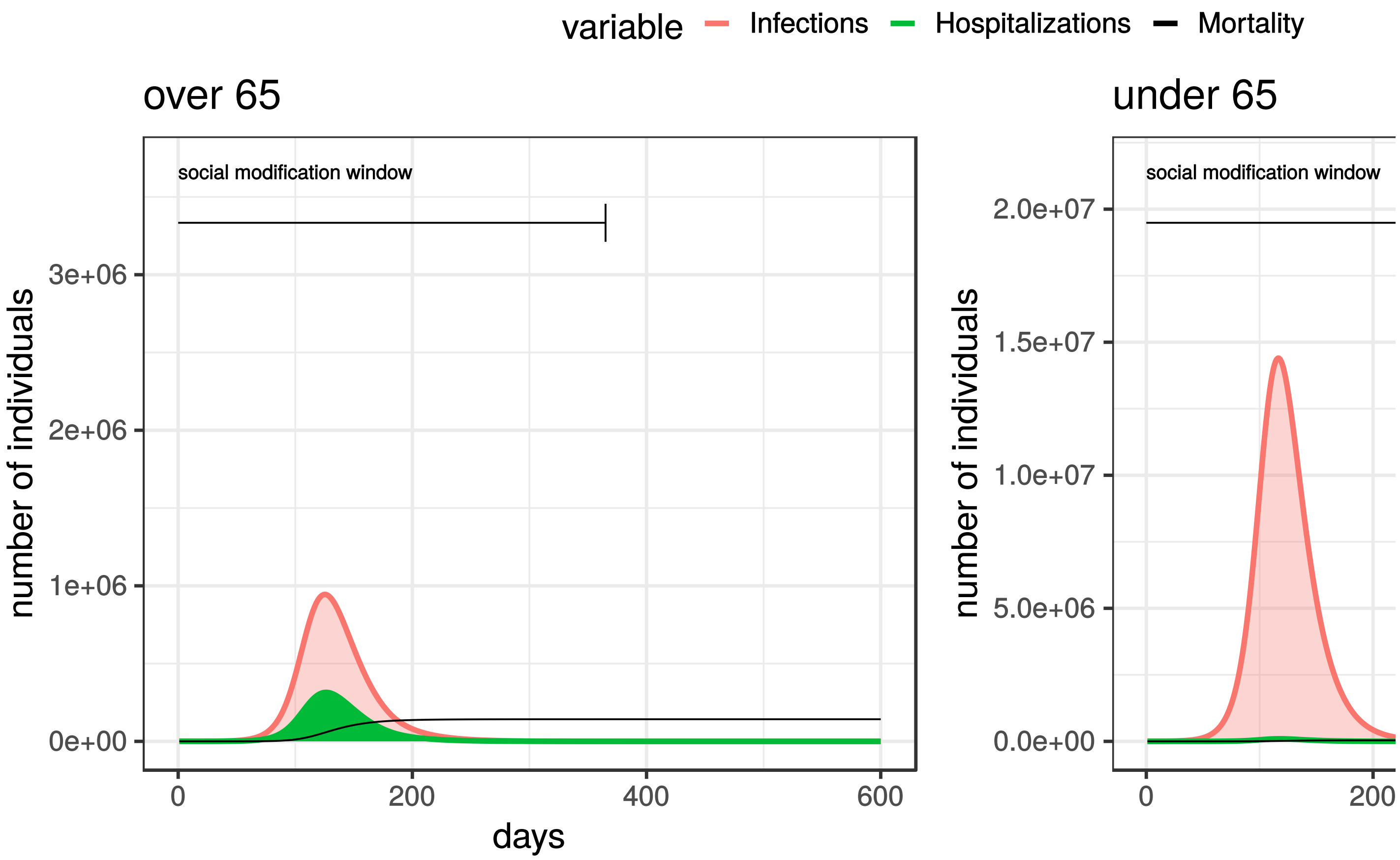
Here α and β are both .9 for 9 months before linearly returning to 2.8 at the 1-year mark. This scenario keeps the infection at bay until transmission rates return to normal, but essentially only succeeds at delaying an epidemic almost as large as would be seen with no interventions. We see **486,000 mortalities**. The light green region under hospitalization curve shows overcrowding.

Scenario 2: intermediate homogeneous measures



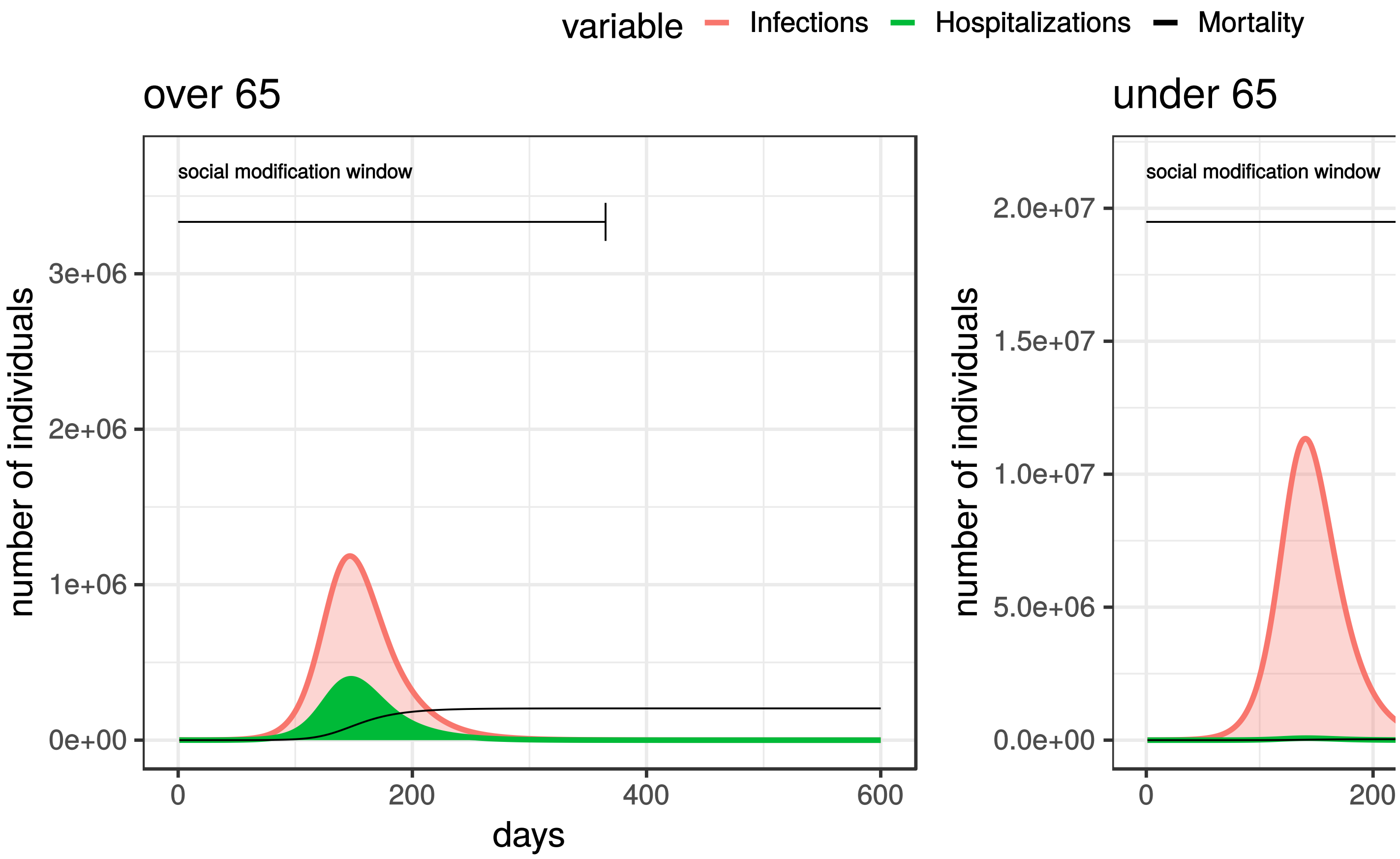
In this scenario α and β both begin at 1.8. After 9 months, they linearly return 2.8 at the 1 year mark. 1.8 is chosen for this scenario because it minimizes mortalities among all homogeneous interventions in our model. This scenario saves many lives compared with no interventions: We see **314,000 mortalities,** roughly a third less than Scenario 0.

Scenario 3: extreme heterogeneous measures



In this scenario α and β begin at 2.8 and .6, respectively. In particular, transmission rates have been sharply reduced among the high-risk population, but not curtailed among the low risk population. After 9 months, they linearly return 2.8 at the 1 year mark. This scenario cuts mortalities by more than 60% from Scenario 0; we see **180,000 mortalities.**

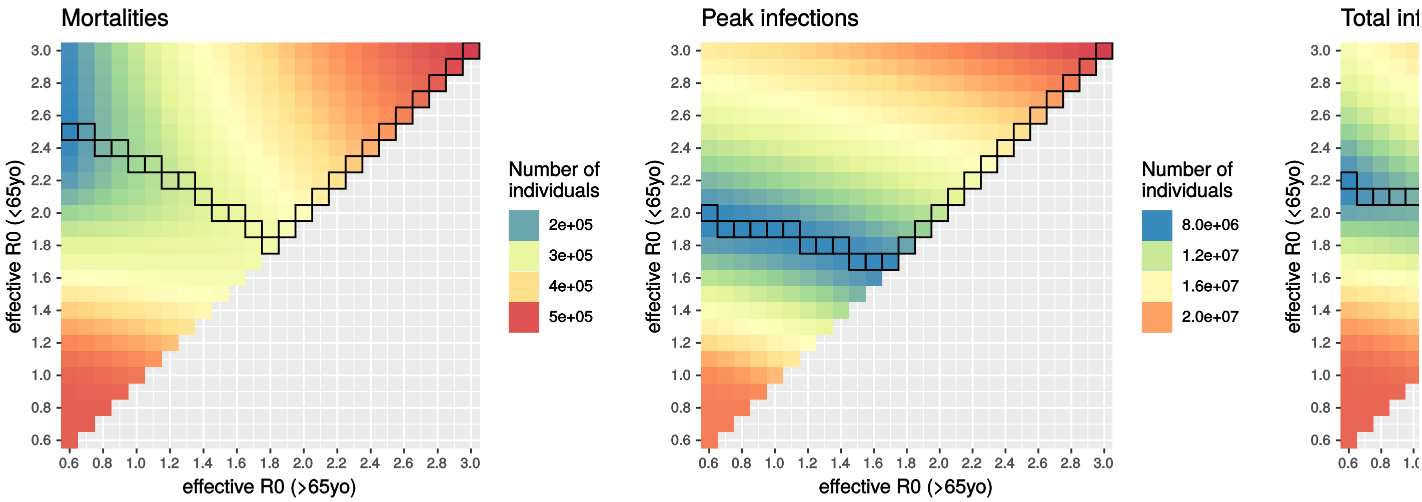
Scenario 4: intermediate heterogeneous measures



In this scenario α and β begin at 2.4 and 1, respectively. In particular, transmission rates have been sharply reduced among the high-risk population, but modestly curtailed among the low risk population. After 9 months, they linearly return 2.8 at the 1 year mark. This scenario cuts still cuts mortalities by roughly 50% from Scenario 0; we see **240,000 mortalities.**

Scenario heatmaps

Below we show heatmaps showing the response of mortalities, total infections, and the infection peak to a range of selections for α and β. (We do not consider cases in which β is set to be greater than α, which result in a large number of fatalities.)



The black boxes highlight the minimizer for each value of β. Note that columns correspond to β, and rows to α. Some simple takeaways from this figure are:

- Minimizing the number of infections and the number of mortalities are not the same goal. In particular, the optimum choice for α is significantly higher when minimizing mortalities vs. infections.
- While some diagonal strategies (α=β) do reasonably well from the perspective of minimizing the number of peak infections or total infections, they do much worse when the goal is to minimize the number of mortalities.
- If mitigation strategies are much less effective than we expect, so that it is not even feasible to bring R_0 below 2, than heterogeneous measures are worse than homogeneous measures. (That is, for β>2, diagonal strategies with α=β are best.)

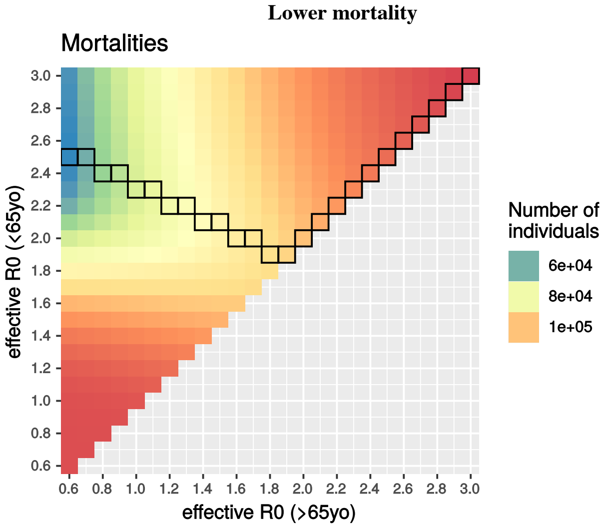
Robustness of our findings

In the following sections we discuss how robust our broad findings are to the various parameters we have to choose when running our model. In each case we show the heatmaps of the response of mortalities to selections of α and β.

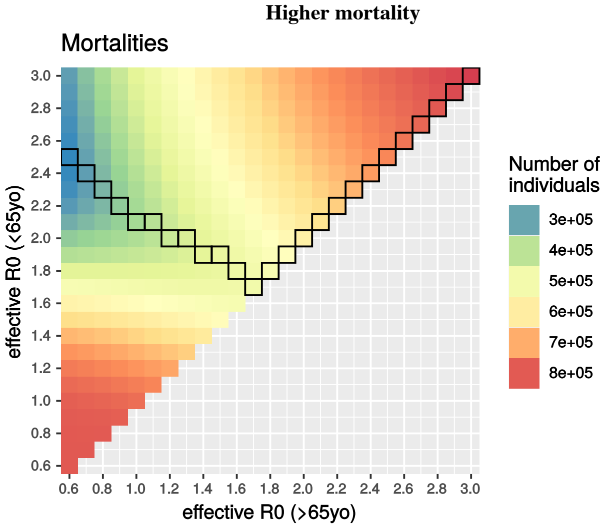
Sensitivity to overall mortality rate

Our model parameters correspond to an overall mortality rate of roughly 0.5%. In the two figures below, we see the response of our model if instead the overall mortality rate is set at 0.15% or .75%, respectively. While the *number* of mortalities is very different between the two heatmaps, the response of mortalities to the choices of α and β is very consistent.

Lower mortality

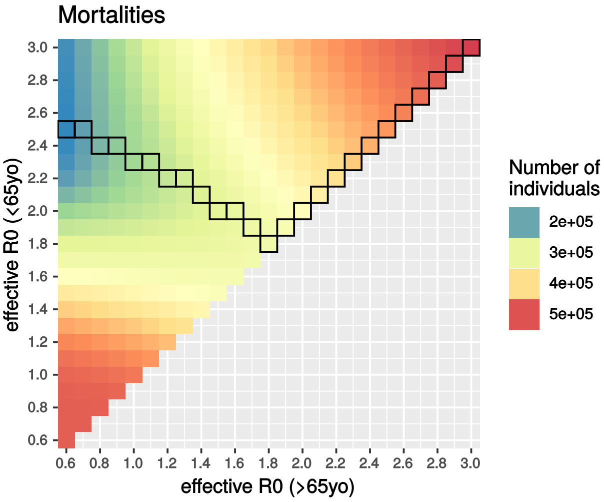


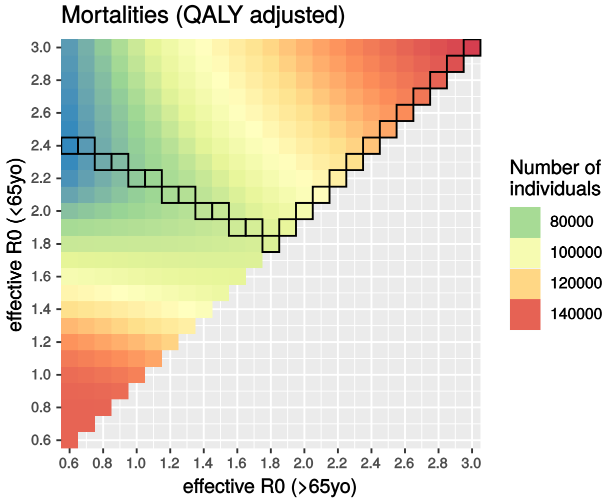
Higher mortality



Sensitivity to target

In our model we have examined strategies to minimize total mortalities, but we might care to optimize a different target, such as quality-adjusted life years lost to COVID-19. Naturally, there are complexities involved in accounting for this precisely, since, for example, it is likely that greatest mortality for COVID-19 (even within age groups) occurs for individuals with the fewest expected quality-adjusted life years remaining. But we can crudely examine the effects of these kinds of considerations by simply applying a multiplier, assuming that the number of quality-adjusted life years remaining is, on average, 5 times higher in the under-65 population than in the older population. As seen in the following comparison, this has little affect on our conclusions about the optimum choices for α and β.

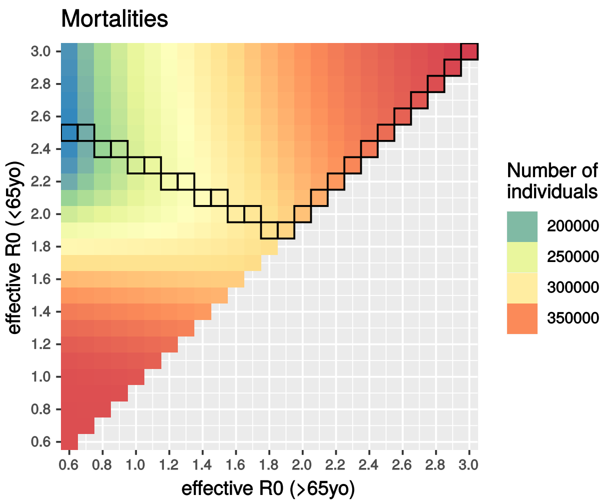




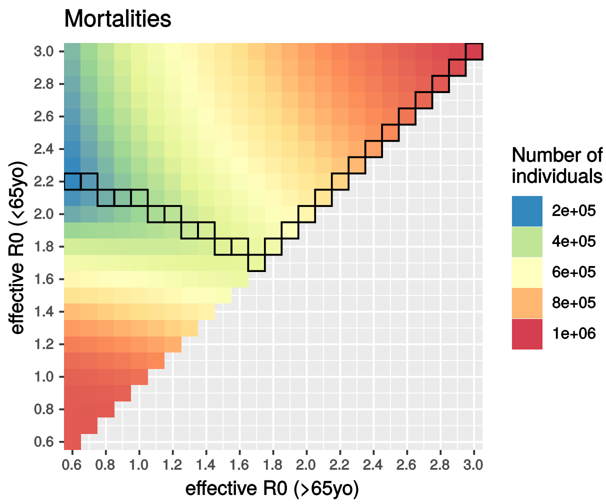
Sensitivity to medical system capacity and overcrowding cost

Our model of overcrowding cost is relatively simplistic. We simply assumed that above a hospital capacity of 500,000 (available for COVID-19 treatment), overcrowding doubles the mortality rate. Our model is not sensitive to these choices, however. In the first heatmap below, we see the response of our model if we assume unlimited hospital capacity, while in the second, we see the result if a threshold of 250,000 is used, and we assume that overcrowding quadruples hospital mortality rates.

Unlimited capacity



Reduced capacity, greater cost



One effect that overcrowding does have on our model is on the direction of the (small) impact of heterogeneous strategies on mortalities in the younger population. For example, in the reduced capacity, greater cost regime, total mortalities are decreased in absolute terms for the younger population with heterogeneous measures, even though the younger population has a higher infection rate and slightly more total infections under heterogeneous transmission strategies.

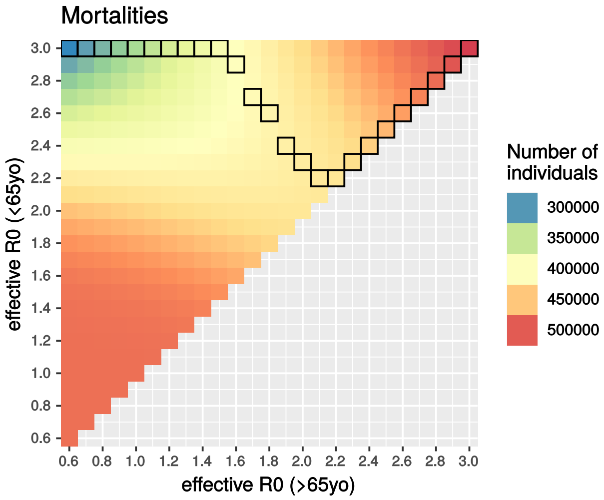
Sensitivity to the hospitalization rates in each age group

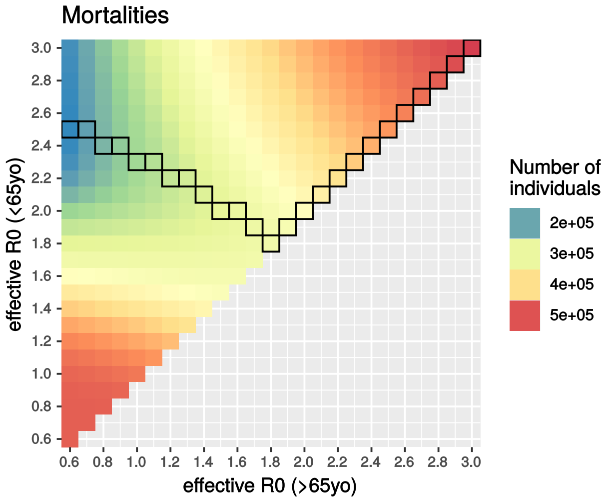
In our model we assume that each death corresponds to roughly 10 hospitalizations. However, this model parameter is only affects our model output when we employ an overcrowding cost on mortality as a result of exceeded hospitalization capacity. As seen in the previous sensitivity analysis, our conclusions hold even if the hospitalization overcrowding cost is completely dropped from our model. **In particular, even in a simplified model in which hospitalization rates are completely ignored, our conclusions hold.** Moreover, as seen in the section below on the sensitivity to relative mortality rates, our conclusions hold even if mortality rates differ much less between age groups than is currently suspected. This also captures the effect of varying the relative hospitalization rates between age groups in our model, since our model assumes that each mortality is associated with roughly 10 hospitalizations.

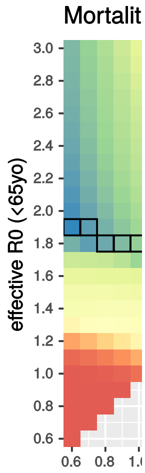
Sensitivity to duration of mitigations

Our Assumption 3 dictates that eventually, society will return to near-normal interaction levels. In our model we have assumed that after 9 months, society will begin to linearly return to normal interactions, which will be reached at the 1 year mark. Here we compare the results of our model with this standard assumption (middle) with cases where we assume that mitigations begin at 3 months to return at normal at the 6 month mark (left), or remain in full force for 18 months, before linearly returning to normal levels a the the 2-year mark.

6 month modification period **1 year modification period**





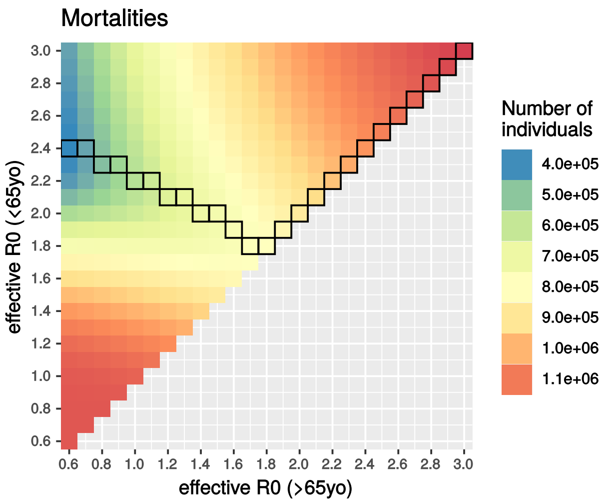


We see that the shorter the social modification window, the more advantage heterogeneous mitigations have over homogeneous mitigations. But even with a 2 year window, heterogeneous mitigations dominate if they can achieve that β<1.5.

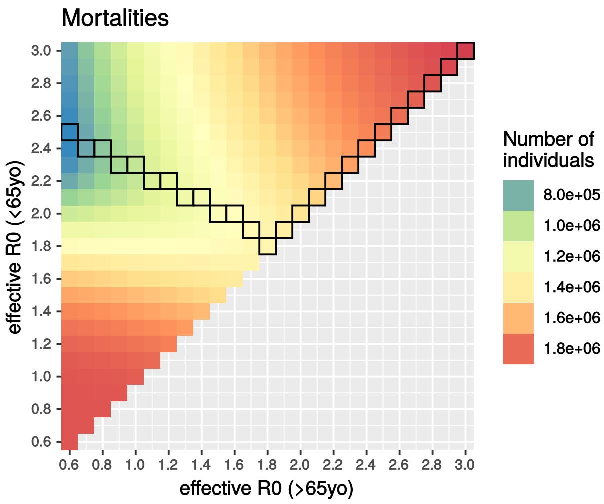
Sensitivity to the size of the initially susceptible population

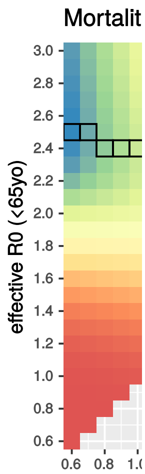
In our model we assume that initially, 25% of the population is susceptible to infection by COVID-19.[7] This assumption primarily affects the overall estimates for mortalities in our model. As seen below, setting the initially susceptible fraction instead to 50%, 75%, or 100%, respectively, has little effect on the response of mortalities to choices of α and β.

50% susceptible



75% susceptible



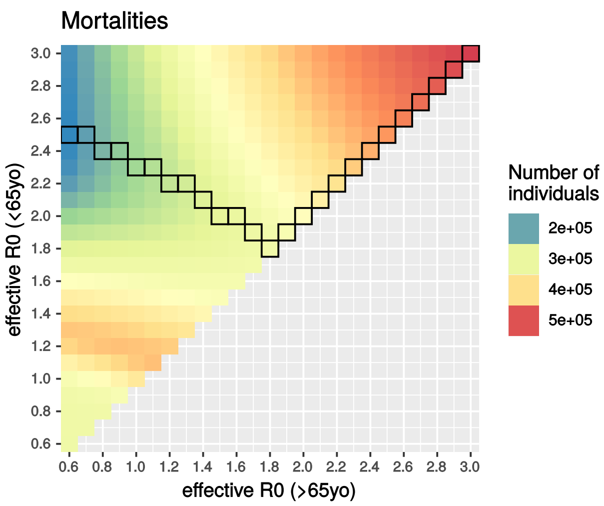


Sensitivity to future changes in the overall mortality rate

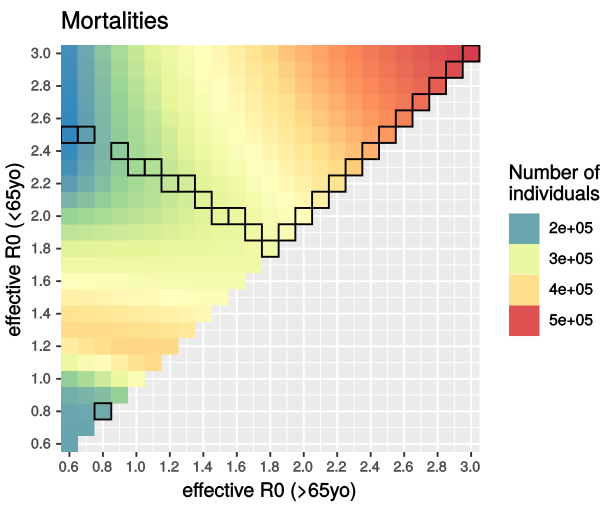
Like our ability to affect the final number of infected individuals, future mortality decreases, if modest, would be dwarfed by the multiple-orders-of-magnitude effect of age on COVID-19 mortality.

Below we show the effect on our model if mortality rates are reduced by 30% and 50%, respectively, at the 1 year mark.

30% reduction



50% reduction

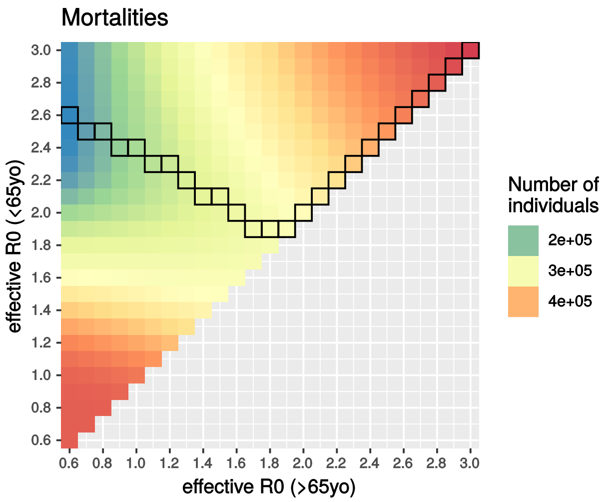


With a 30% reduction in mortality, heterogeneous strategies still strongly dominate all homogeneous strategies. A 50% mortality reduction is around the tipping point where some very extreme homogeneous strategies become competitive with heterogeneous strategies. Note that even at a 50% mortality reduction, the best strategies are heterogeneous. And the homogeneous strategies which are competitive in this regime presumably involve much greater social and economic costs.

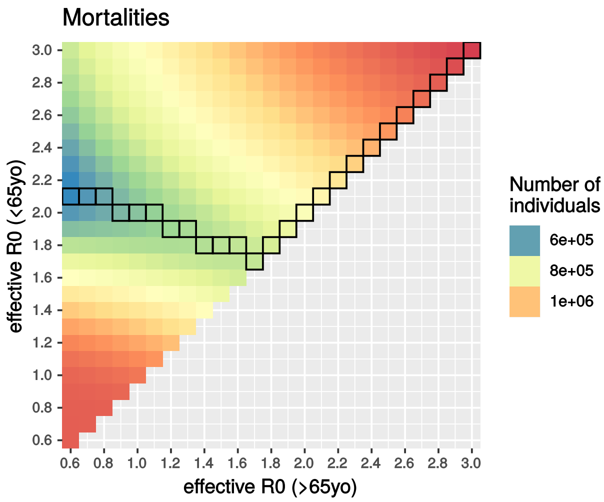
Sensitivity to the relative mortality rates of the age groups

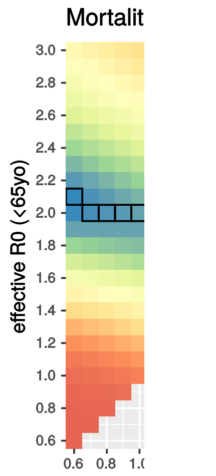
In our model we assumed that the older population has a mortality rate 50 times greater than the younger population. In the following plots we see the response of our model if this ratio is instead set to 100, 5, or 1, respectively. We see that the heterogeneous transmission thesis holds most strongly when the ratio is large, and essentially disappears when the ratio is 1. (In the third heatmap, mortalities can nearly be minimized with a strategy where α and β are equal.)

100:1



5:1





In particular, our findings **are sensitive** to the age-relative mortality rates; they only hold under the assumption that mortality rates are significantly higher in older populations than younger ones.

Counterarguments

Previously we discussed possible objections to our assumptions. Here again we survey some of the simplest arguments against the heterogeneous transmission thesis, in particular as it applies to fighting COVID-19 mortality.

It is not feasible to selectively modify transmission rates among age groups

One objection could be that what we are suggesting—selectively modifying transmission rates by age group— is simply not feasible. While we have clear evidence that overall transmission rates have been controlled in some cases, it is not clear whether heterogeneous transmission has been tried, so we are not equipped to evaluate its effectiveness.

However, we would argue that that the effectiveness of overall transmission reduction suggests that heterogeneous transmission is also achievable. Many of the strategies employed (having employees work from home, restricting social and economic activity) can be applied in an age-sensitive fashion. Moreover, it seems quite plausible that with more of society functioning normally, it is easier to effectively mount reduction of transmission in specific groups. For example, health organizations have widely discouraged public use of face masks due to shortages and the need for training for proper use. But if the task is to protect a smaller, at-risk section of society, a smaller number of masks and instruction in their use could be heterogeneously deployed.

A very effective vaccine or treatment for COVID-19 may be around the corner

If a vaccine with very high effectiveness for COVID-19 could be developed and deployed worldwide very quickly, then it is quite possible that overall transmission reduction would result in fewer fatalities than heterogeneous transmission. Note, however, that the vaccine would have to be significantly more effective than flu vaccines, and have to be developed on a short timescale. And, as with all other assumptions at play, mistakes will have consequences for fatalities. **If mitigation efforts are currently only justified by assumptions about the feasibility of discovering a very effective vaccine in the near term, and those assumptions prove incorrect, more lives will be lost.** Assumptions behind mitigation strategies must be made explicit, so that their plausibility can be considered by the relevant experts.

Even if an effective vaccine cannot be developed, one could believe that if future treatments could reduce the mortality of COVID-19, then minimizing transmission rates until that point could save many lives. The problem for this perspective is that there does not appear to be precedent for a mortality reduction of the magnitude which would be required in a disease with similar characteristics to COVID-19. For example, Tamiflu therapy in high-risk patients results in a 30% mortality reduction for some strains of flu, and no reduction in others.[8] In particular, it may be dangerous to trade off our ability to shift the age distribution of the eventually infected population in the hopes of buying time for a reduction of mortality which may be more modest than what can be achieved with heterogeneous transmission, or may never materialize at all. In the Sensitivity analysis we show above, we show heterogeneous strategies still save more lives if we assume modest future decreases in mortality.

COVID-19 can be made extinct without infecting a significant fraction of the world's population.

Like the previous counterargument, this is a challenge to Assumption 2. The principal problem for this counterargument is that there appears to be no historical precedent for an infection to initially spread as widely and rapidly as COVID-19, and then dissipate without a significant overall infection rate.

Our parameter selection may not capture the COVID-19 dynamics well enough

In this document we have argued that under the assumptions we have laid out, achieving heterogeneous transmission can result in more lives saved than minimizing transmission over the same initial timescale. Of course, the model we have used is both simplistic and involves the selection of parameters, and while we have tried to choose reasonable parameters and explain our choices, they could nevertheless prove incorrect. However, we have demonstrated above that our main conclusions are generally not sensitive to the parameters in our model, with the exception of the relative mortality rates of COVID-19 in different age groups. However, the large mortality difference between age groups is one of the statistical features of COVID-19 about which there is the least uncertainty given current knowledge.

The world can sustain significant mitigation efforts indefinitely.

If society can weather the economic and social costs of significant mitigation efforts indefinitely, without intolerable harms (e.g., in the long term, excess worldwide fatalities caused by a global economic collapse which outnumber COVID-19 fatalities), then our conclusions do not hold.

Related work

- MS Majumder. Modeling transmission heterogeneity for infectious disease outbreaks, Ph.D thesis, MIT (2018).
- Luca Bolzoni, Leslie Real, Giulio De Leo. Transmission Heterogeneity and Control Strategies for Infectious Disease Emergence, in *PLOS ONE* (2007).
- Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, etc. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, London: Imperial College COVID-19 Response Team, March 16 (2020).
- Istvan Szapudi. Staged social distancing: a COVID-19 mitigation strategy.

Footnotes

[1] This assumption is motivated by several simple observations:

- COVID-19 has spread worldwide. The resources and medical infrastructure among the world's nations vary widely and will limit any effort to cause early global extinction of COVID-19.
- Complete containment of COVID-19 would require *worldwide* containment. So long as active transmission of COVID-19 takes place, all countries should expect new infections to be introduced to their population at some rate, unless the susceptible population has been significantly reduced.
- There is no recent precedent for a highly contagious respiratory illness infecting hundreds of thousands of people worldwide and then becoming extinct without infecting a significant fraction of the world's population. For example, in the 2009 pandemic, H1N1 is estimated to have infected 11-21% of the worlds population, and the annual flu (for which there is an established vaccine infrastructure) routinely infects 10% of the population annually, even in western countries with advanced medical infrastructure. In contrast, outbreaks of new infections diseases which were contained (e.g., SARS, MERS) infected relatively small numbers of people worldwide (less than 10,000) in a relatively small number of countries.

[2] It should be pointed out that even if the intent were not to resume near-normal social and economic activity by the 1 year period, it may be reasonable to expect *compliance* with mitigation efforts to decrease over time, making it particularly important to reckon with the eventual return of greater transmission rates of respiratory illnesses, as it relates to total COVID-19 mortality. On the other hand, if transmission rates could somehow be reliably and significantly decreased without significant economic cost, it might be reasonable for them to continue indefinitely.

[3] This is the median of COVID-19 R_0-values found by this survey

[4] This is roughly the observed ratio between the death rates for people under 50 vs those over 70. Of course, finer-grained models could be built which smoothly vary the mortality rate by age, but our goal here is to use a simple model to demonstrate the basic phenomenon at the heart of the heterogeneous transmission thesis.

[5] Significant uncertainty remains regarding the overall fatality rate of COVID-19, but this value is much less significant for our findings than the *relative mortality rates by age* about which there is much greater confidence. We chose the .5% overall level based on research which examined the relatively controlled case of infections on cruise ships. A recent report published by the CDC suggests a range of .25%-3.0% for the overall fatality rate, while positing that "lower estimates might be closest to the true value".

[6] The U.S. has roughly 900,000 hospital beds, and 100,000 intensive care beds. Of course, not all hospital beds will be available for COVID-19 patients. While it is difficult to estimate how much mortality increases when cases exceed the capacity of the medical care system, a factor of two difference in mortality for excess cases may well be an underestimate.

[7] The true susceptible fraction is not known for COVID-19 and may well be quite larger, quite possibly even close to 100%. The cruise ship infection rates show that it should be at least 25% or so. The point of this Sensitivity analysis is that the Susceptible fraction has no impact on the benefits on heterogeneity. Note also that the fraction we assume to be susceptible does not actually affect that rate at which the epidemic spreads in our models, since one normalizes the transmission constants to match the observed transmission rate R_0. (For example, an R_0-value of 2.8 in a completely susceptible population corresponds to an R_0 value of 5.6 in a 50% susceptible population.)

[8] Lytras, Mouratidou, Andreopoulou, Bonovas, Tsiodras. Effect of Early Oseltamivir Treatment on Mortality in Critically Ill Patients With Different Types of Influenza: A Multiseason Cohort Study, in *Clinical Infectious Diseases* **69** (2019).