Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

System Engineering and Overshoot Damping for Epidemics such as COVID-19

Robert L. Shuler, Theodore Koukouvitis, Dyske Suematsu

May 4, 2020 (preprint)

Affiliation: Shuler Research, retired NASA

Email: robert@shulerresearch.org, theodore@koukouvitis.com, dyske@dyske.com

Abstract

The goal of this paper is to contribute the perspective of a systems engineer to the effort to fight pandemics. The availability of low latency case data and effectiveness of social distancing suggest there is sufficient control for successful smoothing and targeting almost any desired level of low or high cases and immunity. This control proceeds from spontaneous public reaction to caseloads and news as well as government mediated recommendations and orders. We simulate multi-step and intermittent-with-feedback partial unlock of social distancing for rapidly-spreading moderate-mortality epidemics and pandemics similar to COVID-19. Optimized scenarios reduce total cases and therefore deaths typically 8% and up to 30% by controlling overshoot as groups cross the herd immunity threshold, or lower thresholds to manage medical resources and provide economic relief. We analyze overshoot and provide guidance on how to damp it. However, we find overshoot damping, whether from expert planning or natural public self-isolation, increases the likelihood of transition to an endemic disease. An SIR model is used to evaluate scenarios that are intended to function over a wide variety of parameters. The end result is not a case trajectory prediction, but a prediction of which strategies produce near-optimal results over a wide range of epidemiological and social parameters. Overshoot damping perversely increases the chance a pathogen will transition to an endemic disease, so we briefly describe the undershoot conditions that promote transition to endemic status.

Keywords

Coronavirus, COVID-19, pandemic, partial unlock, social distancing, economic impact, ventilator utilization, SARS-CoV-2, overshoot, SIR, model, simulation, caseload management, undershoot

1. Introduction

The goals of efforts against pandemics or epidemics are threefold: reduce the number of infected individuals and therefore deaths, avoid overtaxing the healthcare system, and reduce the social and economic impact [1]. Modelers have an important role to play in all three goals, but there is a difficulty not present in other publically popular forecasting such as storm tracks. Humans are part of the epidemic system and upon receiving results of the model they will change their behavior. Later it seems the model was wrong. This is a well known problem in epidemiological modeling. And while there are efforts to model human response [2], it is dependent on many factors, some time varying, and not nearly as mature a science as epidemiological modeling. To reduce overshoot and save lives, caseload *forecasts* must become *targets* and human behavior elicited to meet forecasts.

In this paper, we show that uncalculated government action, spontaneous premature easing by society and overly aggressive caseload reduction before a planed easing are all likely to result in an overshoot condition and unnecessary deaths. We analyze and support all types of scenarios. We provide time estimates for reaching full pathogen elimination, overshoot analysis for small step eco-

© 2020 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

nomic relief, and scenarios for optimum attainment of herd immunity. The choice of scenario is up to each country or region.

At the time of this writing, near the end of April, 2020, the US is adding 25,000 new cases and 1200-1800 deaths per day as the effective replication rate hovers between 1.05 and 0.95, meaning the number declines very slowly and might reverse at any moment (We will include trend continuation in our analyses). Improvements in mortality are in the offing, but improvements in social distancing such as more use of masks is likely to be offset by easing of lockdown for economic and social relief.

Four months into a global pandemic there had been no calculation of how long either the US or global economy could be shut down without permanent, lasting damage. There had been no calculation of personal costs in the US other than unemployment figures whose impact was masked by extending unemployment benefits to 9 months, covering only half the expected period. The UN had suggested those on the verge of starvation would double to 250 million. There were no credible numbers presented with options, only vague projections of models with exponential functions at their root which have sensitive dependence on external (degree of future unlock) and internal conditions (case ratio, for example). Only the state of New York and one California county had actually measured case ratio. No efficacy numbers or distribution percentages had been proposed for the unknown future vaccine. No vaccine existed for established human coronaviruses, including SARS and MERS. The only strategy articulated by the US task force was testing and case tracking, which had never been used on a scale of at least 10 million undocumented cases (see section 1.2) perhaps up to half of which were asymptomatic. The pathogen invades cells via an ACE2 surface protein present in many species of animal (see https://www.cnet.com/how-to/coronavirus-and-pets-how-does-covid-19-impact-cats-dogs/) and so far had infected both house cats and zoo cats, providing an additional reservoir for future emergence. The favored medical treatment for severe cases was still ventilation with a 12-20% chance of survival (ventilators are hazardous for healthy people). The human response seems from the viewpoint of a systems engineer to be hope, fear and frustration, with no systematic consideration of options and their costs and impacts.

1.1 Options and Costs

To illustrate the difficulty of the solution space to a problem like COVID-19, we present in Table 1 some results for which we will later explain methods.

Table 1. Estimated deaths and economic cost of lockdown for a COVID19-like pandemic
showing **simply managed (upper)** and *theoretically optimum (lower)* impacts (projection from April 23 world data)

Unlock%	**Economic Cost** % of 18mo[5] lock - need to save[7]	**R0=2.5** case ratio= **7**	**R0=2.5** case ratio= **20**	**R0=2.5** case ratio= **80**	**R0=3** case ratio= **7**	**R0=3** case ratio= **20**	**R0=3** case ratio= **80**
0%	100% - 32M	3.4M[6]	1.55M	0.64M	3.4M	1.55M	0.64M
5%	96% - 30M	13.4M *6.7M*	4.7M *2.4M*	1.19M *.6M*	16.62M *8.31M*	5.8M *2.9M*	1.4M *0.7M*
25%	77% - 23M	42.2M *25.3M*	14.7M *8.8M*	3.6M *2.2M*	49.5M *34.6M*	17.4M *12.2M*	4.2M *2.9M*
35%	68% - 20M	51.2M	17.8M	4.3M	59.1M[1]	20.6M[1]	5M[1]
40%	64% - 19M	54.9M[1] *32.9M*	18.1M[1] *2.9M*	4.6M[1] *2.9M*	62.6M[2] *37.6M*	21.8M[2] *13.1M*	5.4M[2] *3.24M*
100%	**9.6%[3] (default)** *29-15-11%* *(near optimal)*	**76.2M[4]** *50.3M*	**26.6M[4]** *17.6M*	**6.6M[4]** *4.4M*	**80.1M[4]** *55.3M*	**28M[4]** *19.3M*	**7M[4]** *4.8M*

[1] Approximate herd immunity at this level (blue, meaning "clear, immune") – NOTE: All figures below this number in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

the same column leave susceptible overhand below herd immunity if no vaccine, or proportional deaths if vaccine not 100% effective

[2] Includes modest overshoot of 2-4%

[3] Costs incurred mostly before pandemic seriousness is recognized, based on COVID-19 response

[4] Around 20-25% of total deaths are due to overshoot (bold)

[5] 18 months is assumed lead time for vaccine development. Economic cost is percentage of an 18-month total lock-down.

[6] Figures in green pass cost/benefit analysis, subtract from **bold** overshoot figure in same column and compare to “need to save”. Note this does not consider medical resource protection, which is accomplished in the optimal cases of the last line.

[7] Approximate number of lives that should be saved to justify the social cost that could be used to save or improve other lives is scaled from Wilkinson’s analysis of New Zealand [3] using the minimum number of lives, i.e. including his entire “gray area.”

This paper is intended to be useful either later in the response to COVID-19, or in response to some other epidemic or pandemic spread by social contact. If applying to a disease with a vector such as mosquitoes or rats, a significant translation of the concept of lockdown and unlock in terms of either avoidance of the vector or control or partial extermination of the vector is required. If people are the vector, it is not acceptable to exterminate people. They are who we are trying to protect. Yet an 18-month-long lockdown may double the number of starving people in the world to over a quarter billion (see “Coronavirus: World risks ‘biblical’ famines due to pandemic – UN”, https://www.bbc.com/news/world-52373888)

This paper is not intended to decide goals for the society, only to provide an efficient method of implementing them once decided. The authors prefer a voluntary compliance approach. Those who wish to incur risk of infection by serving others and bear the burden of immunity may express themselves in that way. Those who wish to remain isolated, either for their own protection, or to effectively remove themselves from the herd to increase immunity through social isolation, may do that. But to help organize the dizzying range of numbers in Table 1 we mention two methods of making trades that in theory do not involve moral equivalency, i.e. quality life-years are compared only to quality life-years, though they involve some hazy assumptions. In the table, the green numbers are options justified in a simple scaling to the frame of Wilkinson’s social value analysis for New Zealand [3].

The other method is to assume some impact on a year of life from lockdown, say 10%. It will be much more than that for people who lose homes, permanently lose jobs, or cannot find food. We just look for an average number. Then we multiply by 40% of the world in lockdown to produce the current level of control, the number of years of lockdown 1.5, the number of people in the world about 7.5 billion, and the percent of lockdown cost of the line from Table 1 we are considering. For example, the cost of the 25% unlock line is 1.5 x 7500M x 77% of lockdown cost x 40% people actually locked down x 10% = 346 million life-years. To compare to lives saved we need an estimate of how many life-years are gained with each death avoided. Many of the deaths, perhaps 75%, are among the older who have pre-existing conditions. Few affect young adults or children. For the sake of example we use 10 life-years per death avoided. Then at R_0=2.5 and case ratio = 7 we have 42.2M deaths for the simple 25% unlock strategy, a saving of 76.2 - 42.2 = 34M deaths avoided, or 340M life-years gained. The 76.2M deaths were associated with lockdown costs already incurred of 9.6% of lockdown or 43.1M life-years, giving our 25% option a net cost relative to the immediate unlock option of 346 - 43.1 = 302.9M life-years against the 340M life-years gained by the 25% option in deaths avoided. Further calculation will show that this method approximately agrees with Wilkinson as long as the 10% impact assumption holds. That is what we meant by “hazy assumptions.”

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

1.2 Parameters for Simulation Examples (background on COVID-19 early data)

If the reader is familiar with SIR models and the history of the COVID-19 infection and various discoveries of the range of the replication factor (also called reproduction number), currently thought to be between 2.5 and 3.3 without social distancing, and the case ratio, with measurements finally coming in but varying from 7.5 in New York state exclusive of New York City, to the 50-85 range in Santa Clara county California, then we suggest skipping this section. Basically we conclude that a method must be demonstrated over a wide range of replication rate and case ratio, amounting to two orders of magnitude in epidemic impact (see Table 1 impacts).

Using COVID-19 as an example, just one month ago a case ratio of 7 to 1 total to documented cases was considered "substantial undocumented infection [that] facilitates rapid dissemination of novel coronavirus" [4]. The replication rate R_0 was estimated between 2.2 and 3.58 [5]. Those figures put us in the worst region of the chart in Table 1, the left most column of either R_0=2.5 or R_0=3. Death figures with rapid spread overshooting herd immunity in the range of 70-100 million were to be expected in the absence of any action. However when people are dying in the tens of thousands (now hundreds) with projections in the millions, the expectation of a 100-year pandemic event produced plenty of action. Most air travel and many borders were shut, businesses were shut, schools were closed and store shelves were evacuated of essentials by the end of the third week of March. Although this "lockdown" was not officially labeled as indefinite at the time, it was quickly apparent to anyone with an epidemiological model that it was a possibility. Vaccines, not certain to work, were reported 18 months away, and any unlock would resume the pandemic.

We use a standard SIR epidemiology model [6] implemented by the authors [7] to have certain features that track an ongoing epidemic, allow input of social contact behavior at a high level, and simulate a decision maker that looks at data and makes a simple on/off decision regarding a daily schedule of social contact level. The problem we are addressing is that an SIR model is very good for estimating the eventual outcome, but not particularly useful for either advising the public of how bad the epidemic is going to get or when, and thus not particularly useful to policy makers wanting to know how much modification of social behavior to put in place and when and how to remove it. The reason is important to keep in mind. The number of people who will be infected if no one has immunity is at least $1\text{-}1/R_0$. The replication rate R_0 might be known to within a factor of two or less pretty quickly. Suppose it ranges between 2 and 3.5. Then the minimum number of infected, lacking a way to stop the virus, is 50% to 71%. One can overshoot those numbers, but there are no sneaky variances producing wildly different numbers.

On the other hand, the rate of growth of the infection is determined by an exponential. If it spreads at 41% per day as COVID-19 was doing in the US in mid March, then in 30 days one person has infected 30,000. But if the rate changes just slightly to 31%, the number infected is ten times less, only 3,000. That is disturbing to a public who now expects hurricane tracks to be accurate, rain to start within half an hour of the predicted time, and spacecraft to hit an asteroid millions of miles away, and they are justified in believing the speaker doesn't know what he is talking about.

There is a second issue particularly acute with COVID-19. One doesn't know how many people are infected, especially if some of them express no symptoms. Do we have that 30,000 today, or will we have them in a month or ten months? There is no way to achieve credibility with the open loop aspect of a SIR model. However many other disciplines cope with exponential functions in a much more controlled and predictable way – your cell phone is full of them, your car, your computer, and these things all perform predictably. But it has not been apparent until the large scale public compliance with COVID-19 social distancing measures that a suitable control mechanism might exist for an epidemic, and the shortage of medical resources made the need for control particularly acute.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

Figure 1 shows that for a case ratio of 7 and a replication factor of 2.5, starting from the world case data through April 22, 2020, even after one year a partial 60% unlock produces a rebound pandemic of nearly the same magnitude as the original. The spike in cases is merely pushed to the right. If it is done suddenly as shown (either deliberately or as a result of social upheaval) then cases overshoot the herd immunity threshold causing 16 million completely unnecessary deaths.

Uncertainty over the replication rate of a new pathogen is to be expected. Initially there will not be tests for it, and the particular region in which it initially spreads may not be representative of broader society. In the first month of global COVID-19 awareness the WHO did not even realize it spread by social contact. Various self-protective factions in the country of origin hoped they could contain the epidemic without revealing its true danger. However, the case ratio completely spoiled those hopes.

Large numbers of COVID-19 infected individuals showed no symptoms. This was not known until tests were developed. In March a hospital in New York City (NYC) found 33 of 215 women in the hospital to give delivery tested positive, 15.3% of the sample [8]. Twenty-nine had no symptoms. During the sample period the number of documented cases in NYC varied from hundreds to around 40,000 at the end of the sample period. Taking the most conservative approach and using the 40,000 cases near the end of the period, or 0.5% of NYC population, the presumed case ratio would be at least 15.3% ÷ 0.5% ≈ 30. This would put us to the right of the middle column of either side of Table 1, with reductions of over 50-75% in eventual cases resulting in death. If we take 10,000 NYC known cases from March 18 as a basis, still in the latter part of the sample period, the case ratio becomes 120, less than any of the columns in Table 1, and 1.5 orders of magnitude less than the overshoot cases with no lockdown, or sudden release of lockdown.

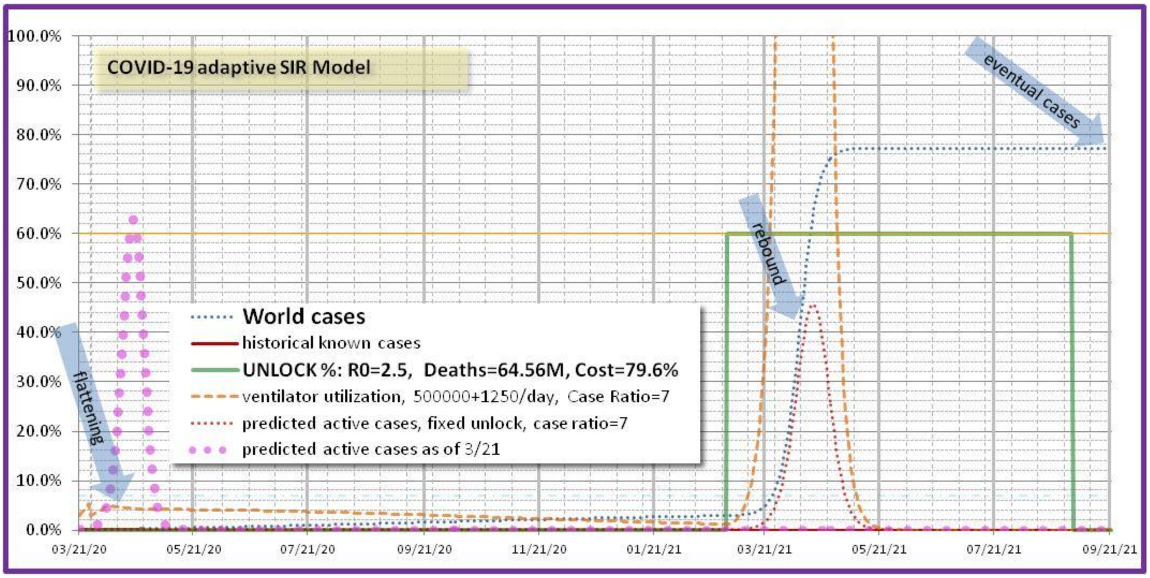

Figure 1. Projected COVID-19 rebound if lockdown maintained one year and abruptly eased 60%

By April, it was by inference from US data on non-fluenza doctor visits for Influenza-Like-Illness (ILI) that an upper bound on the COVID-19 case ratio in the US was estimated at 230 [9]. Infection rates in a homeless shelter in Boston were found to be 36% [10]. The sample is not random. If it were, since Boston case rates were lower than NYC, the implied case ratio would be in triple digits. A random sample study of case ratio in Santa Clara county, California showed values of 50 to 85 [11].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

On April 23, the New York governor announced antibody test results of 3000 shoppers, with antibody rates of 14% for the state of New York and 21% for New York City, and numerous caveats about the randomness and reliability of the testing. Nevertheless it was the most comprehensive testing to date. Using case data from about the time of the testing, case ratios were obtained of 12.4 for New York City, 10.5 for the aggregate state of New York, and 7.5 for the state excluding the city. Between the New York and Santa Clara results, we have the basis for the wide range of case ratios tabulated in Table 1. Case ratios much lower than 7 with high mortality rates will hopefully be handled by containment as were the earlier SARS and MERS outbreaks. Case ratios much higher than 80 with high replication rates can easily approach their maximum infection rates before they are fully understood. Thus we expect the table, and this paper, to cover a middle ground requiring broad social action.

Knowing the case ratio does not change any physical fact. It does not change the trajectory of the disease. It changes our knowledge of two things:

- How many individuals remain susceptible, and therefore at what level of known cases the new cases rate might decline, both with and without social distancing (two different numbers).
- The mortality rate, what fraction of those who become ill will die.

Observed mortality rates for COVID=19 vary from 0.1% (Qatar) to 14.9% (Belgium) (see https://coronavirus.jhu.edu/data/mortality). Variation may be due to case handling, population demographics (risk factors), or case ratio. Case ratio can be lowered by testing, though one should be cautious as some 90 different tests are available and the effectiveness of some is as low as 20% (see "Antibody Test, Seen as Key to Reopening Country, Does Not Yet Deliver" https://www.nytimes.com/2020/04/19/us/coronavirus-antibody-tests.html). If a new testing regime is implemented and the case ratio in a region changed from 20 to 7, that does not mean suddenly very many more people are going to die. It does not move the region from one column to another in Table 1. The table was constructed from the presently observed average world mortality rate. Gaining case knowledge in a region changes the mortality rate for that region exactly the complement of the amount it changes the case ratio, such that no net change occurs in expectations as a result of testing.

The entry of COVID-19 into the US was revealed to have been a month earlier than previously thought by posthumous test of a 57 year old woman who died in her home on February 6th, and had no travel history explaining her exposure to the virus (see "2 Californians died of coronavirus weeks before previously known 1st US death" https://www.cnn.com/2020/04/22/us/california-deaths-earliest-in-us/index.html). Even isolated indigenous people have died (see "First coronavirus deaths reported in indigenous communities in the Amazon" https://www.nationalgeographic.com/history/2020/04/first-coronavirus-deaths-indigenous-communities-amazon/).

1.3 Background on Herd Immunity Thresholds and Overshoot

In this paper herd immunity (or population immunity) [12] is a mathematical concept involving all factors that affect replication rate, from social distancing to vaccines. It refers to a threshold, a percentage of infection, at which the number of cases decreases rather than increases. Human immune response is only one of the factors. Behavior and changes in the makeup of the target population also affect the threshold. Herd immunity occurs when the combined factors drive the replication rate below 1.0. It is possible to have herd immunity even if individual immunity is weak or short lived. It is not desirable, but possible. If a vaccine is available, it has a percentage of effectiveness, and is given to a percentage of the population, both factoring into herd immunity [13]. Even tracking infected persons with cell phone apps contributes to herd immunity [14].

Herd immunity occurs nominally when $1\text{-}1/R_0$ of the population have been infected, or 60% in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

the case of R_0=2.5. Many of the would-be transmissions fall on immune members of the population. The susceptible population S that has not been infected is approximately (ignoring mortality) $1/R_0$ when herd immunity occurs and will continue decreasing as more are infected. If the total population is P, the effective replication rate will be $R_t = R_0 * S/P$, where S/P we call the herd immunity factor. The point where R_t=1 we call the herd immunity threshold H_t=1-1/R_0. If a vaccine is given to V fraction of P and has E fractional effectiveness, and social distancing results in a fraction D of the previous level of transmissions, then $R_t = R_0 * SVED/P$. And if T is the fraction of P that are long term tolerant of the pathogen in the case of incomplete immunity then H_t=1-T/R_0.

So why do the cumulative cases in Figure 1 shoot so far above the herd immunity threshold, shown as a yellow line? The infected fraction of the population is about 45% as it goes through 60% and Rt=1 in Figure 1. The cases at that time will on average only be half over and since Rt=1 at the threshold will create half as many new cases. Overshoot thus depends on the number of active cases as the herd immunity threshold is crossed, along with R_0. Mentioned in the literature by Handel et. al. [15, 16, 17], little guidance is given other than running a simulation. Using a spreadsheet a table is shown in Figure 2. In general cases must slow down when approaching the threshold, which can be managed by controlling social distancing or any other of the components of R_t.

The overshoot in Figure 1 is much worse than it looks because we have only unlocked 60%. The replication rate is reduced 40% from 2.5 to 1.9 and the immunity level at 60% unlock should be 47%. Thus in Table 1 the theoretical minimum deaths are barely more than half of the unmanaged step-function unlocks. These are not achievable, but we can get much closer. However, the gradual unlock required increases economic costs. These are estimated only for the last line of Table 1 and vary with R_0.

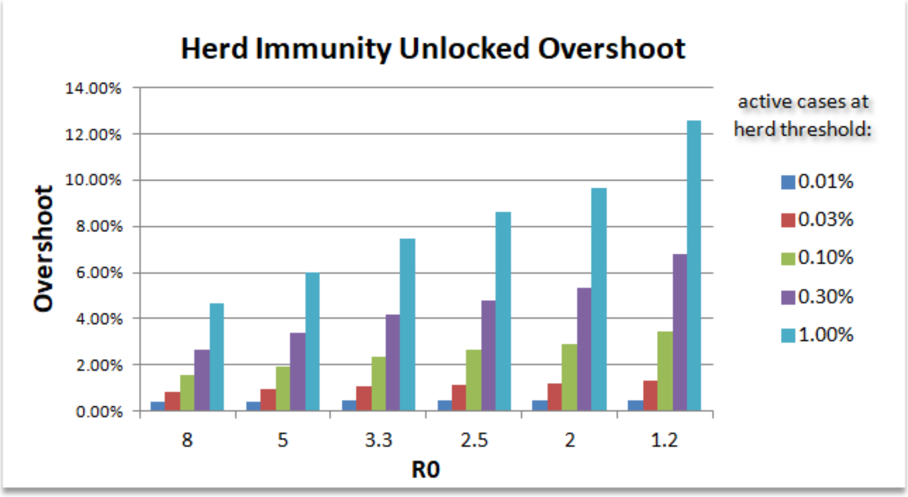

Figure 2. Overshoot as a function of R_0 and active cases at herd immunity threshold

If setting an unlock percentage to limit caseload on medical resources, then it is subject to the same overshoot. If policy is to allow people to get sick to build herd immunity, the target cannot be implemented in a single step without overshoot.

Do we imply the authors think human populations overshoot to such high percentages of illness? We do not if populations are informed about the methods of disease transmission. They may if mortality rates are low and the illness is mild, like a cold. However overshoot of low caseload targets will not trigger social reaction, and so overshoot will reach its full potential there. At a 10% unlock target, COVID-19's monthly death rate might be similar to the flu though extended over more

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

months.

There is a drawback to overshoot optimization. Coming too close to the threshold with a small caseload allows the disease to exist until population increases and/or immunity wanes, creating an endemic disease that remains in the population, especially regions where R_0 is a little higher or immunity or vaccine coverage a little lower. We discuss this in section 4.

1.4 Lockdown goals (background on containment, suppression, immunity)

As a heuristic guide to future response planners, about 5 weeks is all the US remained on full lockdown for a disease with around or less than a 1% mortality rate. Even the governor of hard-hit New York said on April 26 it would "be impractical to force people to remain in their homes all throughout the summer with nothing to do. 'There is a sanity equation here,' he [said] pointing to reports that domestic violence, drug and alcohol abuse and mental health problems have already increased" (see https://www.bbc.com/news/world-us-canada-52435648). Management of full lockdown is in any case not a problem for modelers, but for social planners and leaders, so it is not addressed by our method. All we can do is monitor R_t and if it is too close to 1.0 or above it, point out the full lockdown has failed and cases are accumulating toward the full death total, just more slowly. A plot of empirical replication rate of COVID-19 for the US and the world excluding the US is shown in Figure 3, tenuously clinging to 1.0 and not enduring long below it.

When the replication factor hovers near 1.0 the number of active cases stays relatively constant, and the cumulative cases mount linearly rather than exponentially. Depending on how long before some treatment, vaccine or end of lockdown changes the situation, death total may be the same or less than an unlocked situation. It might seem that a short period of super lockdown to reduce the number of active cases, and reduce the linear accumulation of deaths, would be advisable. But with a dramatically lower number of active cases the motivation to maintain the severe conditions necessary for 1.0 or below replication factor, already wavering, is not maintained. The hardest thing for modelers, leaders and the public to face is that not everything is possible right now. We might not be limited in what we can solve eventually but an epidemic held at 1.0 is a ticking bomb, and if society is stressed to breaking the lockdown will not be maintained.

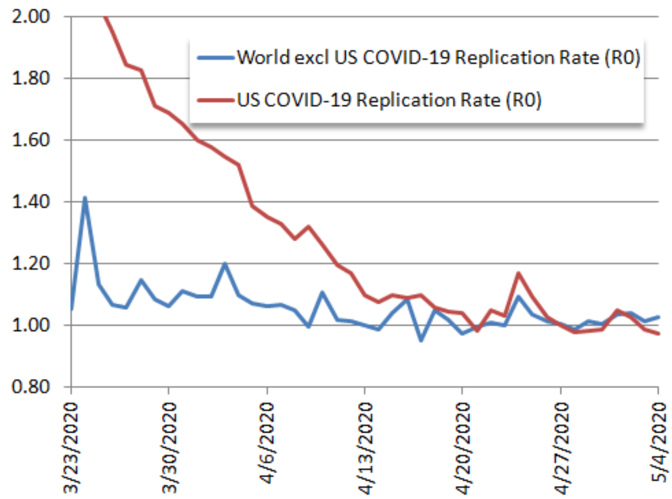

Figure 3. Effective replication rates for first full month of COVID-19 lockdown, 4-day moving average.
NOTE: There appears to be a weekly cycle, unclear if data collection issue or some type of bi-stability

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

With the understanding that vaccines are imperfect and not 100% distributed, and that social distancing contributes to herd immunity as long as it is in force, then containment and immunity become points on a spectrum of strategies with similar properties, not different strategies. Containment can have "profound adverse consequences for civil liberties and economic status" [18]. Consider the following scenarios:

[1] The first patient (called patient zero, patient one or the index case) or patients can be identified and everyone they have had contact with. Isolate only those people. That small group develops herd immunity. The pathogen ceases to circulate among humans until it again jumps from wildlife.

[2] The first patients are not identified but migration in or out of the affected region can be identified. That region is isolated to prevent spread elsewhere. Those with signs of infection are isolated until they develop herd immunity. The outcome is the same as before. Humans remain susceptible.

[3] Migration out of region of origin. Many cases not symptomatic. Large number of cases, millions to billions (many undocumented). People self-isolate or develop immunity by vaccine or contracting disease. Replication rate drops below 1.0 and infection declines until immunity wears off. Natural H1N1 immunity might endure practically indefinitely (100+ years without decline, including antibodies that can protect against a variety of flu strains) [19], while SARS CoV-1 immunity lasts for 2 years [20] as does vaccine immunity for H1N1 [21].

[4] Similar to [3] but herd immunity is defined by an effective replication factor R_0'<R_0. Cases are either held low until a vaccine is developed (waiting) or until the infections are reduced to zero (suppression). Note that waiting may or may not lead to suppression, and suppression is a different strategy than containment. Suppression works faster with a small number of cases, but unlike containment it requires broad social measures because all the cases aren't known. All these are herd immunity strategies. With waiting the infection level peaks much lower due to the lower effective replication factor. With suppression social measures are lifted when suppression is complete, but the mechanism of eliminating the disease is the same as herd immunity, not track and selectively isolate like containment.

Short hypothetical examples using COVID-19 and Ebola parameters will show how epidemics wind up in one category of the other. It is important to do a few calculations, if only on the back of an envelope, before basing public policy on hope. Ebola had R_0 of about 2, and a 50%-90% mortality rate [22]. The manner of death is a hemorrhagic fever, which is striking. It can fester in remote areas for weeks, but if a case comes to the attention of a clinician they have "good reasons to suspect Ebola if a mysterious disease occurs" (see https://www.who.int/csr/disease/ebola/one-year-report/factors/en/). So the case ratio will be nearly 1.0, and most cases not in remote areas will be noticed. A case close to patient one is likely to be identified. Geography aids in its containment. Virtually every case is critical and comes to the attention of a clinician if one is available. So Ebola and other Ebola-like diseases can often be dealt with by the first and second methods above.

Let's take R_0=3 for a new COVID-19-like disease, just to have a whole number. We conservatively use the New York estimate of a case ratio around 10. Of the known cases 10% are critical, but initially were not thought to be different than ordinary pneumonia. And that is only 1% of the total cases. Patients can recover in about a week if they are not on vents (official recovery is longer because of waiting for tests and required additional waiting times). If we approximate spreading as generations of cases instead of continuous then it is easy to see how this develops. The number at left in Table 2 is the generation, or week. The middle number is likely total number of cases. The right number is likely critical cases, rounded off. We assume it might take ten total critical cases to attract attention, partly because not all of them will be in the same hospital.

It could easily take 7 weeks to get noticed in a heavily populated area. The following timeline is from https://www.axios.com/timeline-the-early-days-of-chinas-coronavirus-outbreak-and-cover-up-ee65211a-afb6-4641-97b8-353718a5faab.html . On December 10 the earliest known patients began feeling ill. This is likely week 5 since there is more than one. Hospital admissions began on December 16, and there was a post from a hospital official on December 30 and an announcement of 27 cases on December 31. This agrees with the table: noticed and announced in week 7. The first case would have been in early November and could have been late October. An estimated (now, they didn't know the case ratio then) 2,165 people, many with no symptoms, have the illness. Wuhan health commission insists there are no new cases and on January 14 WHO says Chinese authorities have no evidence of human to human transmission. By then, week 9, there are likely 19,683 cases. The next day the patient thought to become the first US case leaves Wuhan carrying the virus. But within another two weeks patients with no travel history are dying in their homes in Santa Clara. There are likely 177,147 cases in the world, mostly undocumented and untraced. Within a month there will be 14 million before social distancing is seriously considered outside China, and with the lowest credible case ratios there are 25-50 million by late April, only 2.5 million of which are documented.

Table 2. Progression from patient 1 for a COVID-like illness (hypothetical).

Generation (week)	Total cases	Critical cases
1	1	0
2	3	0
3	9	0
4	81	1
5	243	2
6	729	7
7	2187	22 – noticed!

With tens of millions of cases in the world in late April, many of them asymptomatic and in 210 different countries, any one undetected case could recreate the dilemma at its present state in four months. A small group of cases suppressed to a replication rate around 1.0 could "hide" almost indefinitely. Eradication has been less often suggested, but only two diseases have ever been eradicated and COVID-19 is suspected to exist in an as yet unknown animal population. The goal of lockdown for most regions in the short term was to prevent entry of the disease which failed in 209 countries, and to prevent overload of medical facilities which mostly worked in combination with hospital rapid-construction (as of late April, with possibly a long way to go).

The World Health Organization (WHO) faces a difficult dilemma when a new disease is identified. If it alarms world governments with the possibility of a high replication rate each time, most of the time it will be wrong and governments will cease to listen. It was critical of China in the 2003-4 SARS outbreak and that one was contained. That appears to have been a better strategy.

The purpose and end of the lockdown must be articulated to retain public support. Public compliance with lockdown has dramatically postponed a surging epidemic peak. Now the question is how to exploit this "control" for public benefit without losing public compliance.

The articulation of a clear, achievable and worthwhile goal alone might maintain compliance.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

But if unrealistic goals are hidden behind near term projections, compliance may plummet. For that reason our model is available to anyone and is not hard to use. It may be found at http://shulerresearch.org/covid19.htm . There is even an online interactive JavaScript version with data presently from all 50 US states, most countries, and many regions, courtesy of one of the authors.

The period of lockdown will either end with a highly effective vaccine, or it will end by being lifted so the disease can run in an orderly fashion to herd immunity without overshoot, and that is our primary objective. But simple (one step) lockdown strategies that involve a partial lifting can overshoot their targets, especially low caseload targets that may not elicit public fear. Our discussion of avoiding overshoot will apply to managing any target unlock rate or case rate, not only to the final herd immunity threshold.

2. Approach

We will define multi-step unlock scenarios to meet a variety of objectives, including (1) small step unlock to provide economic relief with minimum case build, (2) caseload management to the capacity of the medical system, and (3) achievement of herd immunity in minimum time with minimum overshoot. We simulate those over scenarios over the range of R_0 and case ratio parameters presented in Table 1 for the US. We also vary the unlock degrees by +/-50% to evaluate sensitivity to degree of unlock.

2.1 Control of Unlock

The means of obtaining influence over social distancing are not a subject of this paper, only the use of it. Our most predictable scenarios require that social distancing be turned up and down in some quantity according to a signal. This might be a schedule of what hours or what days certain businesses or activities may operate. It might even divide workers or customers into separate non-interacting groups. Certain industries (factories, shops) are amenable to this type of control and certain ones are not (generally anything scheduled in advance, like airlines).

The more predictable approach uses feedback. We assume a manager or a committee of them looks at current new case data and decides whether to maintain or change the existing unlock schedule – not every few weeks as policymakers are currently doing but daily. Their recommendations would be implemented within one to two days. From Figure 3 we see from late March when serious lockdown advice promulgated through most of the US, replication factor went from 2.5 to about 1.0 in two weeks. Since we only need a move of about a quarter to a third of that, a week should be sufficient.

Figure 3 does not show such control authority over the world. The world is a big place and not homogeneous with respect to mutual contact. The epidemic follows a separate course in regions with a lot of self-contact. SIR models are in fact only valid within regions in which anyone could possibly contact anyone else. If the infection spreads linearly along a peninsula or travel route, it will not follow an SIR model. Figure 4 shows qualitatively how more or less independent regions might sum up, each one of them governed by an SIR spreading model. The erratic up and down of the sum suggests regions be managed independently.

The authors of this paper believe that unlock measures should be voluntary. Such a measure can be viewed as allowing people their freedom, or as requesting them to participate in the economy for the benefit of other people at risk of getting sick. Below the herd immunity threshold, officials cannot truthfully claim unlock is safe. As our strategy is not a one-shot affair, but requires repeated manipulation of the lock level, honesty is likely the best policy. Those healthy and at low risk and in economic need are likely to be willing to expose themselves to the environment, especially if gov-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

ernments maintain the integrity of the healthcare system and people are not dying from neglect. We take it as an assumption that fear could be overcome to some degree and a partial unlock implemented.

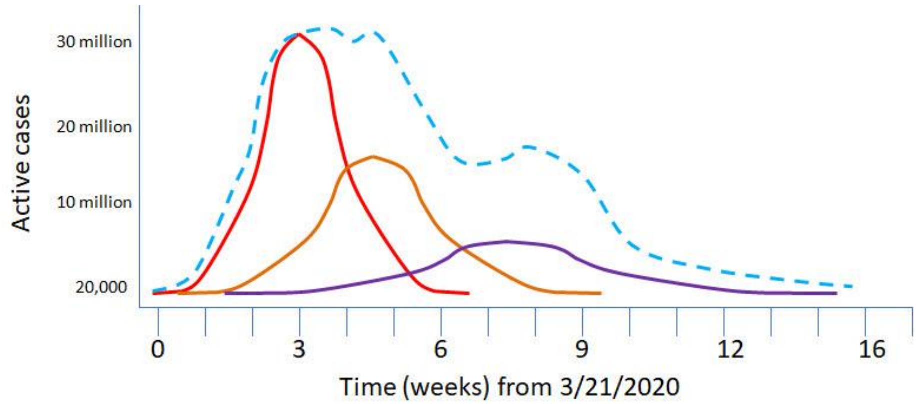

Figure 4. Several regions peaking at different times do not make a smooth total peak

We do NOT require a simulation to *implement* the scenarios. We only use a simulation to validate the scenarios and test them over a range of parameters. Therefore we do not have to be right about the parameters, and no specific prediction is ever of importance. The course of the epidemic comes under the control of humans.

2.2 Description of model

The model extrapolates based on a current R_0 based on the ratio of new known cases to known active cases (not the public number, but using a 6-day average spreading window). What is actually measured by the data is R_t. Reversing the equation for R_t we compute $R_0 = R_t * P / S$. The S (susceptible percentage) is inferred from the population, the total known cases and the case ratio. A 4-day moving average filters this number. We establish a percentage of R_0 recovery called "partial unlock" to the initial level of March 21. For the intermittent-daily scenarios we simulate a decision maker who would decide what days of the week would be partially unlocked, and the simulation responds to his selection 6 days later to allow latency in case development. During a partly unlocked period, the model extrapolates based on the unlock percentage difference between the last measured R_0 and the original R_0, reducing this by the percentage of susceptible individuals remaining.

For the fraction of cases that require resources such as a ventilator we used 5.0% of known cases, or about half of critical cases. This is arbitrary. In another epidemic one may not even care about ventilators. There may be another real resource, or this may really be only a control variable, or a scale factor of the number of cases. We run the simulation for 18 months and take data such as fatalities at that point, assuming a vaccine is available even though this is uncertain, or of uncertain distribution and effectiveness. Mortality rates are a parameter to the simulation and were established from public data, as a percentage of ventilator cases. For the US this is 105% (some deceased patients were never on ventilators, or died at home).

Total ventilators is set to 100,000 for the US, about half of all total ventilators. This is only for reporting utilization in most simulations, but is used for control in the intermittent-daily-unlock simulations. Some ventilators are old, some are in use, and that number was sufficient at the 5% of known cases utilization rate under our optimized scenarios. Unlock level for most of the period of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

intermittent unlock is set to 3 times the static step unlock for the same period, because unlock is intermittent and the net effective value is set by the intermittent duty cycle. Other unlock step levels are the same for comparing intermittent and multi-step scenarios. In addition, loop gain is lowered prior to July 15 because when the cumulative cases are far below any kind of herd immunity thresholds they build more rapidly, increasing loop gain. The method of lowering is to provide unlock on only 3 days per week prior to July 15. In most scenarios, the first peak has been passed by this time. If re-using the simulation the user must adjust this.

Threshold for the feedback loop will be reported with simulations that use it. In addition, threshold match levels of each of the days of the week are adjusted slightly, by about 2% of the threshold, to avoid all three days turning on or off simultaneously if the control parameter is moving slowly near the threshold.

And finally, the simulation adds 300 ventilators per day after April 12 to simulate both manufacturing and distribution of ventilators from stockpiles.

3. Results

3.1 Suppression strategies

The model checks for total cases (dependent on case ratio) falling below 0.5, which can be interpreted as a probability. On this condition it assumes cases have ended. It runs until December 2022 looking for an end date, and cases are not zero by then it declares “NOT ended.”

End dates may vary considerably as they are based on the validity of the last empirical locked down R_0 for the last collected case data. The dates also assume NO cases migrate from other regions. Migration/travel was how the disease spread out of China in the first place. Case-free regions may well be a magnet for those wishing to escape hard hit regions. However model estimates can provide a rough first cut at the viability of a suppression strategy, and insight into how to structure such a strategy.

Based on data through April 7, 2020 for the COVID-19 epidemic in the US, if lockdown at that level were perfectly maintained, cases would be expected to disappear in January of 2021 at a cost of 1.16 million lives. Using data through April 27 the lockdown has reduced R_t considerably. This has pros and cons. The bad news is cases now last 18 months longer until July of 2022. So it is not a viable strategy. However if somehow one could maintain lockdown for two and a half years, the death toll from the disease is one tenth as great. There likely would be social unrest, even in so stable a country as the US.

Dividing a population into separate groups which do not intermix, regions with travel bans for example, reduces the time to suppression. At the current US R_0 this amounts to only one month for each halving of region size, so 64 equal-sized regions would only reduce the time to suppression by 6 months. Not a productive use of resources. Reducing R_0, perhaps through masks and interaction rules as the US has had about all of lockdown it can stand, is much more effective. If it could be reduced from its current value of around 1.0 to around 0.9, the suppression date would be September 2021. That is too close to a vaccination availability date to be of great interest.

Small and relatively isolated communities with small numbers of cases may be able to reach suppression within a few months, or sooner, if they can control travel in and out. They will have to maintain travel restrictions until surrounding communities achieve herd immunity and suppression (at whatever their target level of lockdown, not necessarily full herd immunity) and suppression, or until a vaccine is available.

If New York (state) could maintain its lockdown it could suppress by December 31, 2020. But

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

that seems unlikely, especially given travel. Island nations like New Zealand, which claims to have [nearly] achieved suppression as of the end of April 2020 may have a viable strategy. It is fairly easy to control travel to New Zealand. Argentina has declared its intention to end air travel until September, at a cost of 300,000 jobs. The model suggests this is possible, at least in theory. But the rest of the world will not be clear.

Higher degrees of unlock produce suppression quicker, at a cost in fatalities, as evidenced by the projection from US April 7 data. Full unlock is the fastest. Overshoot on any of these is large. Eliminating overshoot requires some slowing of caseload buildup.

It is apparent from sorting through these alternatives that international travel restrictions will remain in place until most countries are cleared, and that suppression is not generally faster or slower than partial unlock scenarios. Therefore it is not necessary for all countries to follow the same strategy. In fact any one strategy would work very poorly somewhere, so it is best to regionally optimize. Sometimes media articles seem astonished that some country is going its own way, but this is exactly the right approach. Spending a few minutes with our model of the entire world in which cases are "NOT ended" under any fatality-conserving scenarios should be enough to convince most people that regional is best. It is hard to find a region which cannot be handled faster. The problem is that effective solutions depend on local conditions.

3.2 Waiting strategies

If a suppression strategy does not succeed by the time a vaccine is available, and in many regions it won't, then the suppression and waiting strategies are the same. And if suppression does succeed, there is no need to keep waiting. Therefore a separate discussion of waiting strategies is not needed.

3.3 Minimal economic relief strategies

We examined two cases, a constant 5% unlock and a constant 25% unlock. These assume waiting at current levels is unsustainable. By unlocking, they may incur overshoot. For the 5% case the US shows no overshoot. It is nearly at the right level. For the world there is significant overshoot at 5%, but as stated we do not consider it practical to control at the world level. Doing so for a 5% unlock would by extremely sub-optimal for most regions.

Using the US as an example, R_0=2.5 and case ratio of 20 for illustration purposes (the other cases are similar), Figure 5 shows about 80% overshoot for the 25% scenario and deaths of 0.34 million. Keep in mind deaths are inversely proportional to case ratio. For case ratio of 10 the overshoot is the same percentage but deaths are 0.72 million.

Note the economic cost is reduced by about the percentage of unlock, but cases are 2/3 of the way to the 60% level which would be natural herd immunity for R_0=2.5. The yellow line on this chart is effective herd immunity, based on the effective R_0 at 25% unlock. Figure 6 shows the same eventual degree of unlock with an early period at 15% to avoid running through this threshold with too many cases active.

For a 3% increase in economic costs deaths are cut by a third. This is an astonishingly worthwhile benefit. Who can argue? The authors feel such clear choices need to be made clear to the public as motivation. Current announcements are extremely vague as to the goal, probably because officials are afraid the population does not agree on goals. That is an important consideration, but in the authors' opinion not a reason to kill 220,000 people.

This scenario was simulated for +/- 50% of these unlock levels. The totals change because the effective herd threshold changes. But the overshoot damping is the same in all cases.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

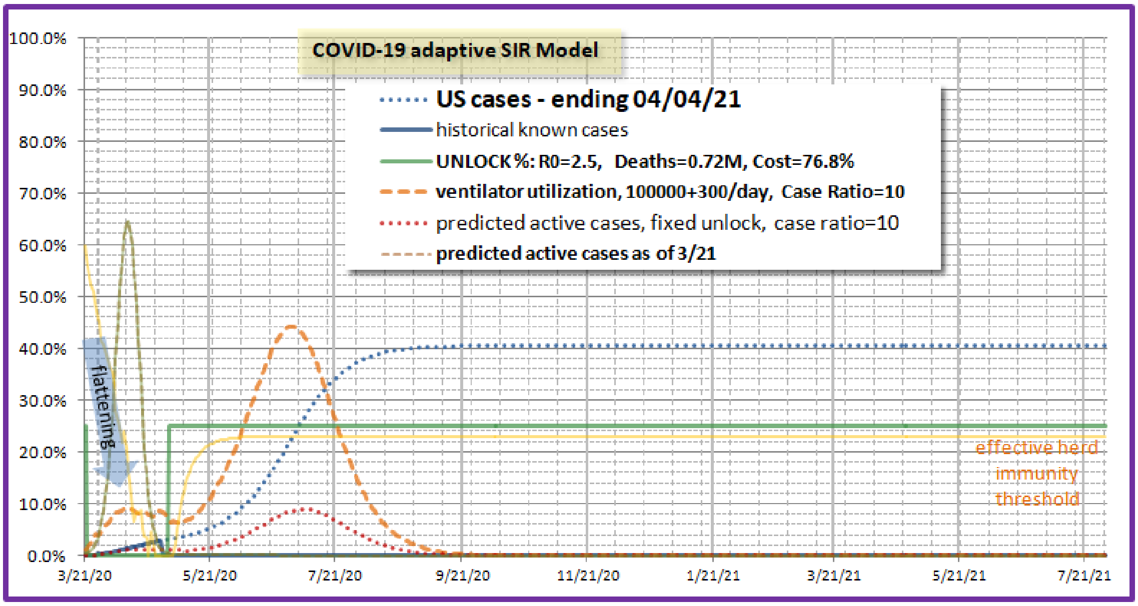

Figure 5. 25% economic relief unlock without overshoot suppression.

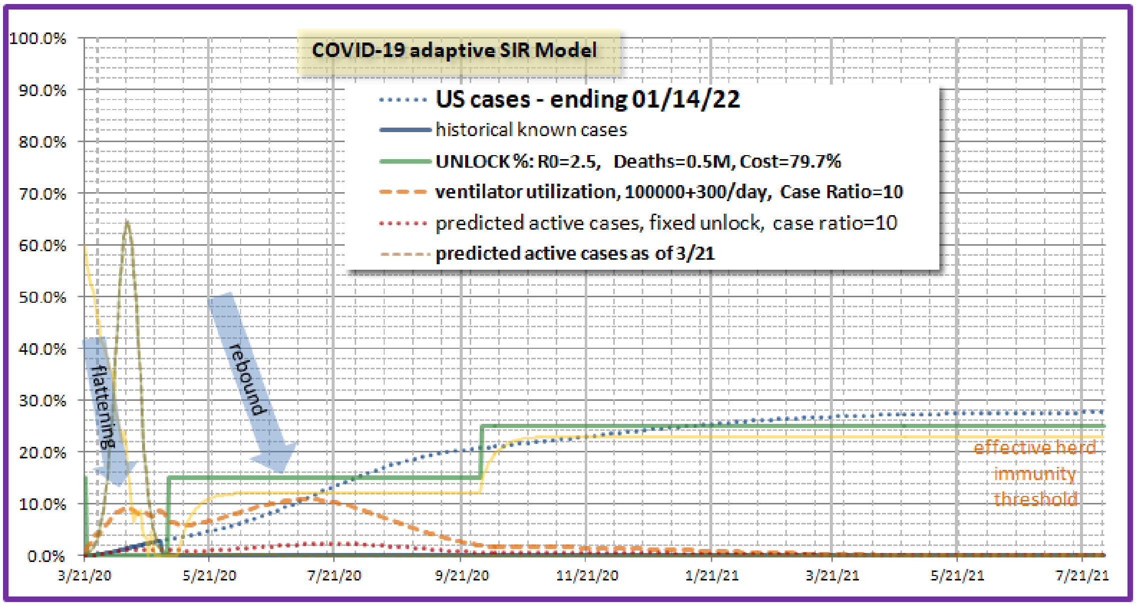

Figure 6. 25% economic relief unlock with overshoot suppression.

The optimal early unlock is actually 16% but it reduces over half the overshoot in the range 10% to 20%.

There is a modest increase in deaths if the 25% step is a month earlier. Later reduces the overshoot slowly toward the theoretical optimum (yellow line) but increases economic costs proportionately.

Varying R_0 and case ratio over the range in Table 1 has little effect on the scenario with the exception of R_0=2.5 and case ratio of 80, where overshoot is not substantially reduced. It is a reasonably robust strategy.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

It is the principle we are trying to convey. This is not a recommendation for a specific epidemic. The incident timing will have passed by the time the reader views this. The method to reduce an overshoot that results from a step unlock is to have a pre-step of less magnitude. The caseload should be falling toward zero but not quite there when the second step occurs. The difference in step magnitudes should not allow the cases to build momentum as they reach the second step H_t threshold. Although we did not have trouble with second step rebound in this example, if the cases are too small but not zero, and the second step is much higher than the first, a rebound can occur and cause overshoot. Giving a numerical recipe that covers an infinitude of non-linear cases is not productive either. A planner need only make a couple of attempts and simulate them across variations in parameters as we illustrated above to ascertain the scenario is robust. An overly optimized approach is often not robust.

3.3 Caseload targeting strategies

The 25% US unlock case already illustrates caseload management. In the single step case, cases increase over what the US has already seen by about five times, using about half of US ventilators. This could be alarming and result in spontaneous public re-lock, postponing the problem. In this case if the final unlock is maintained at 25% there is only a loss of time. But if the low case rates following a re-lock tempt a larger step on the second unlock, a large overshoot may occur as illustrated in Figure 7. The goal of reducing caseload peaks in the medical system is not even met by this scenario.

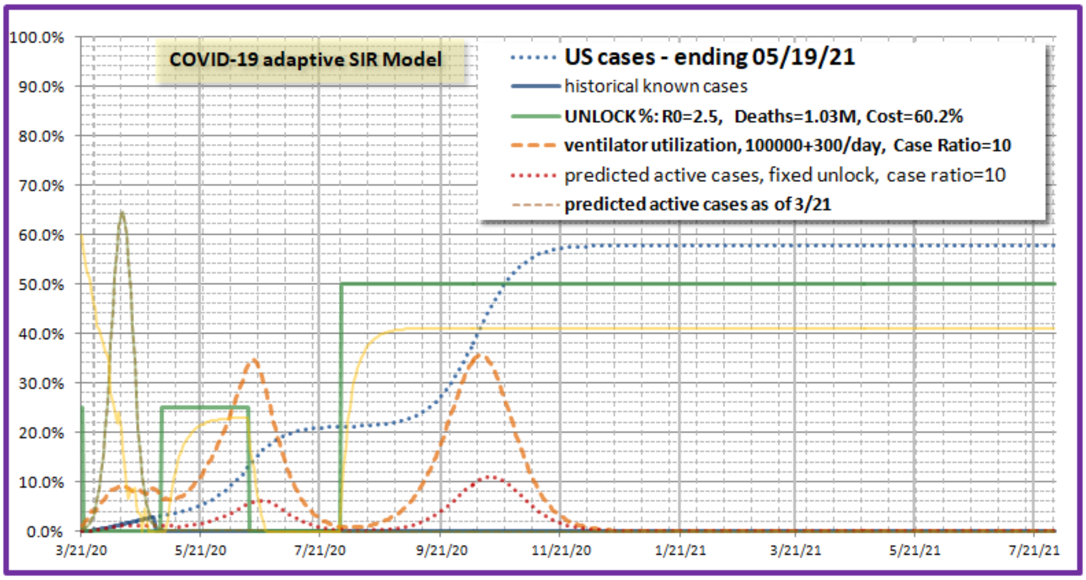

Figure 7. Initial 25% scenario followed by re-lock and unsafe 50% unlock.

3.4 Herd immunity strategies (background on second wave vs. rebound)

All the preceding strategies amount to targeting effective herd immunity based on R_t. What we consider here is targeting natural herd immunity based on R_0, which allows full unlock and also achieves virus suppression (though the virus may hide in animal populations and wait for immunity to weaken over time or in subsequent generations, much as the swine flu did from 1918 until 2003).

With overshoot avoidance, and if case ratio is at the high end of estimates (around 80) this strategy may produce good economic results while incurring fatalities only somewhat worse than a bad flu season. For low case ratios of 7 to 15 it is deadly. Our analysis does not consider a possible second

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

wave. A second wave is not a rebound in our terminology. We use it to mean a mutated virus with different characteristics. In 1918 the second wave was deadly for young adults, for example. Cities which failed to contain or suppress the first wave were little affected by the second wave, so from this rather unquantifiable prospective the herd immunity strategy is safer. Use of the terminology second wave is not consistent in the literature. We found no study of a mutated virus possibility, presumably because the difficulty of predicting mutations. One of the most high profile studies uses "second wave" to refer to what we term "rebound" [23]. Our terminology seems much clearer for communicating to the public. In academic literature an authors may define whatever terminology they desire. This qualification gets lost in the viralization of news.

As we have seen, overshoot may result from an initial peak that crosses a threshold with too high a caseload, or from a rebound after a second unlock threshold is in place. Therefore strategies for eliminating overshoot will suppress rebounds to low levels. If a second wave has the same immunity response, then suppressing overshoot will suppress a second wave as well. In an unanticipated change of the pathogen's immunity response, even a vaccine will not prevent a second wave, just as failure to anticipate the prevalent strains of seasonal flu cause a high number of cases.

The US scenario for herd immunity shown in Figure 8 has good overshoot suppression over the range of parameters in Table 1 and +/-50% unlock levels, except for R_0=2.5 and case ratio of 80 where moderate overshoot returns. The case levels are low at case ratio of 80, so in an overall composite risk picture this contributes little. As more measurements of case ratio are made, the strategy can be adjusted if case ratio is high.

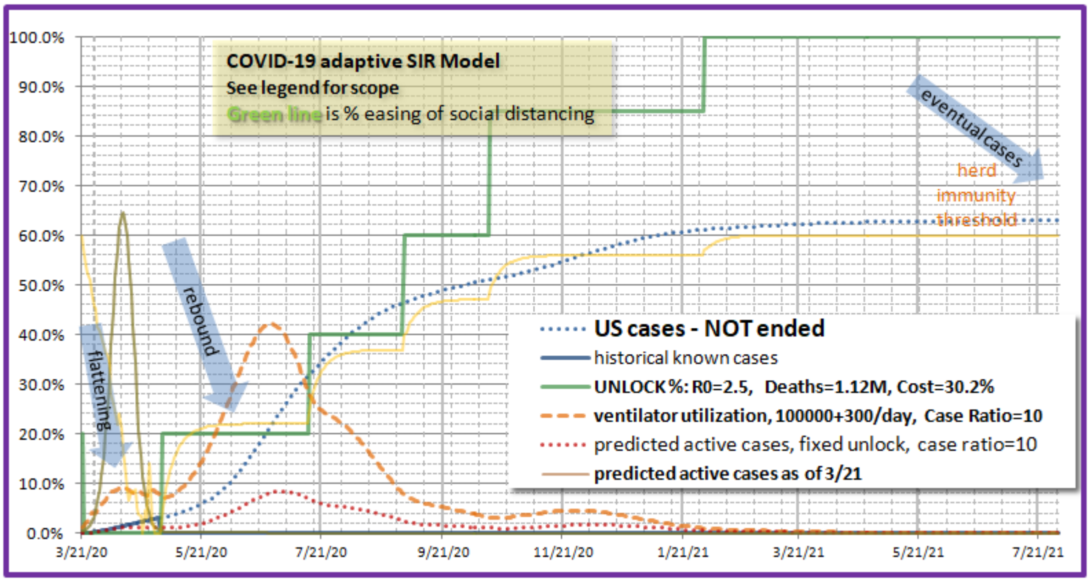

Figure 8. Herd immunity from multi-step unlock for US with overshoot suppression.

The economic cost is less than a third of full 18 month lockdown. The deaths are comparable to the poorly managed 25%-50% unlock case, only somewhat more than the 25% single step case, and about twice the 25% optimized case. A full academic school year is provided, essentially full unlock within the US by October, and full previous activities including travel by February with essentially no risk to the US. Whether other countries will welcome our travelers is another question, but our cases will be vanishingly small.

We continue to use a case ratio of 10 for our nominal cases because it is confirmed in New York state and much lower than Santa Clara. We consider that a likely lower bound for the US, and that 20 or 30 is quite possible, with correspondingly lower deaths and more incentive to go with herd im-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

munity. For the case ratio of 80 the model estimates deaths at 160,000, two or three time a flu season. It is still very early in the study of treatment protocols for COVID-19 at this writing and improvements might be rapid. For another epidemic with a third of the lethality of COVID-19 at the 80 case ratio, herd immunity is the obvious choice. For something similar to COVID-19 at low case ratios, it is not very attractive. There might be ways to open the economy and maintain social distancing. It may be possible to start with something like the 15%-25% unlock and wait scenario, and then transition to a herd immunity strategy in July if mortality comes down, or increase economic activity more than 25% while essentially maintaining the low effective replication rate of the 25% case.

We now demonstrate the feedback concept, simulating a decision manager who provides a daily on/off schedule. In Figure 9 it provides the front end step on a 25% unlock strategy, with moderate overshoot suppression, about half way between our manual handling and a sudden 25% unlock. The unlock degree during the daily schedule is greater than 25% on the unlocked days, but is designed to have an average 25% effect. This simulates adequately over the same range of parameters and unlock-variation as before.

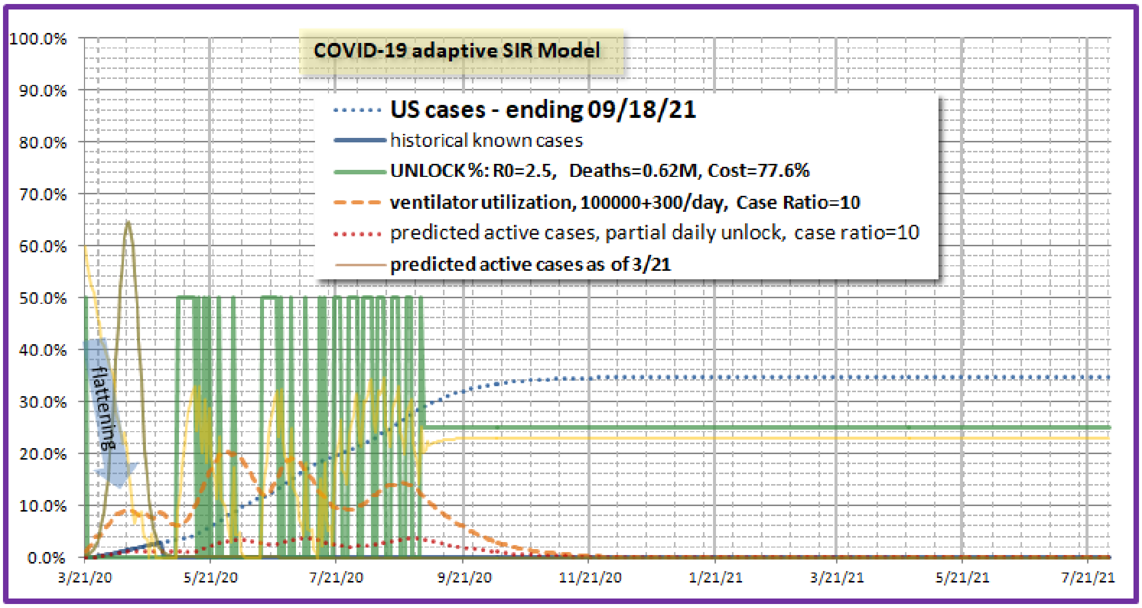

Figure 9. Intermittent daily unlock as front end to 25% unlock.

The advantage is that up until September 1 no decision has to be made as to whether to hold the 25% economic mercy unlock or go for herd immunity. Figure 10 shows that perfect overshoot efficiency is achieved, and once again this was simulated across the range of parameters.

A ventilator caseload target of 20% was set since the scenario was using the feedback for overshoot control rather than simply working off cases at the maxim rate of which the medical facilities are capable. Working off cases at the maximum rate in the next few months does not provide much acceleration of the time table because the larger number of cases has to die away before approaching the herd threshold. Pushing the medical system to its maximum limit for a particular set of parameters undermines the ability to run our gauntlet of variations and causes the R_0=7 and low case ratio scenarios to fail. In the authors' opinion, an "optimum" scenario maximizes results under uncertainty and does not gamble on a long shot.

Notice for the scenario shown in Figure 10 we had to extend our zero-cases check into 2024 to prove herd immunity had been achieved. This is the result of hitting the target *too* closely. Intermediate targets do not particularly matter, but natural herd immunity needs a slight margin of error to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

avoid the pathogen becoming endemic as new susceptible targets are born allowing a low infection to persist, then bloom again when population immunity fades.

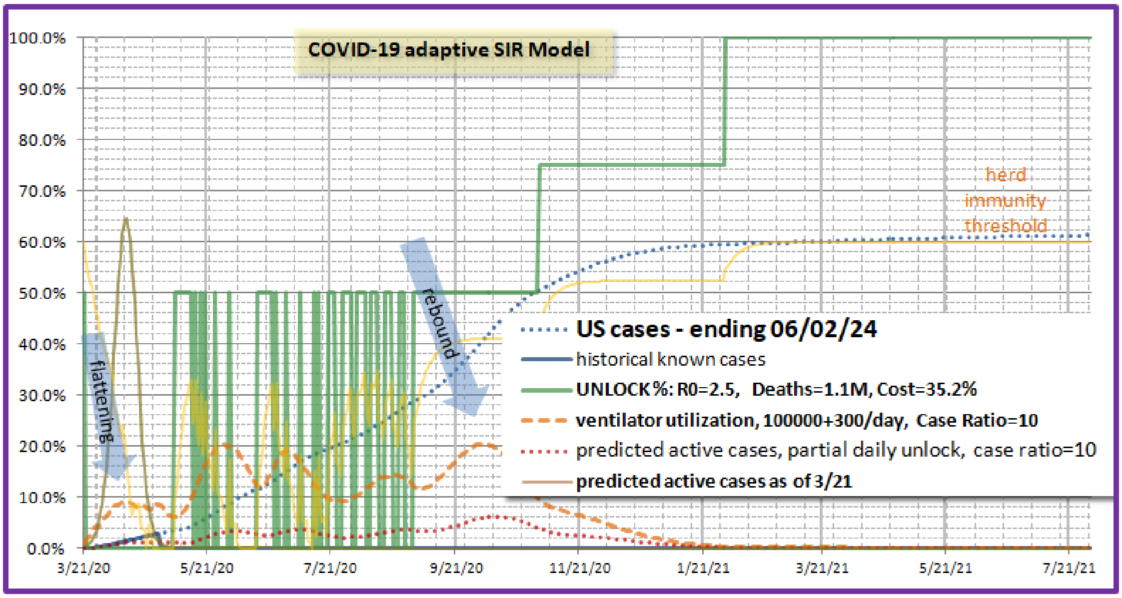

Figure 10. Intermittent daily unlock as a front end to herd immunity strategy.

3.5 Self-regulation (trend continuation) analysis

A different kind of feedback is to assume the empirical R_0 derived from data has settled near or just under 1.0 is because populations are tolerant of the current caseloads and risk. It might be worth an experiment in several regions around the world to try and establish a different lockdown/unlock factor to see if it is stable, to inform future pandemic management efforts. Figure 11 continues the current caseload with its uncertain but apparently slight decline.

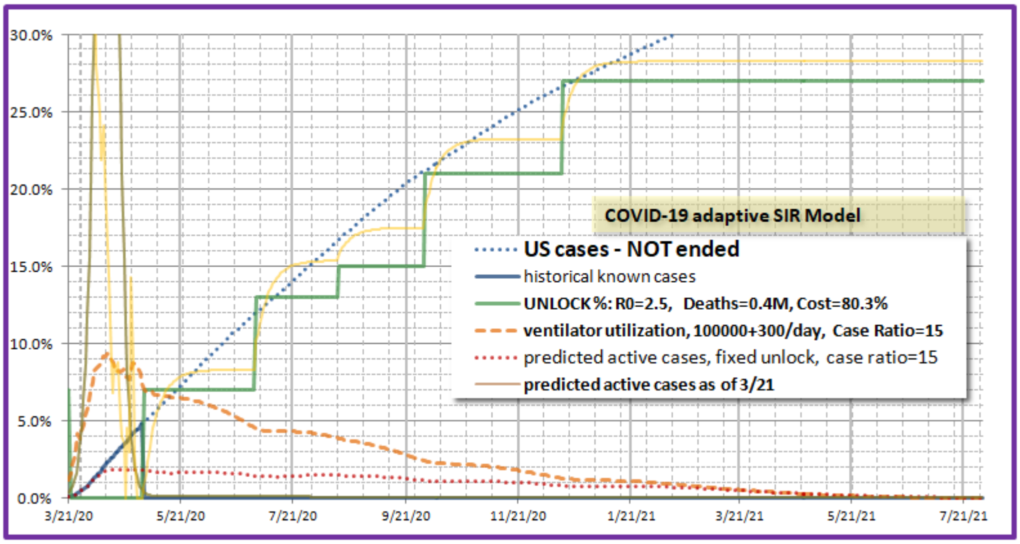

Figure 11. Self-regulation or trend continuation forecast results for US (note scale is different).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

The particular percentages and case ratios are not important in this example. It is possible to establish case ratio over a wide range and still match this steady decline by controlling unlock percentages. We also believe the economic cost might be somewhat less than the 80% shown due to improvements in efficiency of social distancing as more Americans wear masks and each culture makes the adaptations it needs, but we don't know how much. The end result is similar to the 25% overshoot optimized scenario. The total cost in deaths is insensitive to case ratio and R_0 because it is basically a continuation with decline of the current daily death rate. This scenario effectively postulates that the combined actions of the public and officials might well implicitly be using a feedback approach which controls to tolerable death rates.

The only role of the pandemic parameters then is the magnitude of costs that must be suffered to hold death rates down, and whether or not any herd immunity takes hold. If there is a vaccine then only the economic costs will be pandemic-determined. If the cases turn up in the months following April, it probably indicates the economic costs of maintaining this level was too high. And if cases turn down, either summer quieted the virus or case ratio was at the high end of the expected range and costs will be lower than expected.

4. Undershoot and possible transition to an endemic disease

Notice that Figures 8 and 11 have the annotation "US cases – NOT ENDED." Figure 10 claims an ending only in June of 2024. We don't even yet know if natural immunity to COVID-19 lasts that long. We saw earlier with the H1N1 that vaccine immunity lasted only two years. We presume, then, that with population growth and declining immunity, cases would not in fact have ended in real life. Also R_0 is not an exact quantity and varies from place to place. If we conclude from modeling that world cases should have ended, there will likely be many regions where R_0 is higher, making the herd threshold H_t higher, and it has not ended. In examples with substantial overshoot such as Figure 1, it is much more likely cases will actually end and no endemic condition will be established outside the source animal population. Modifying our model to include 0.6% population growth for the US the possible endemic transition scenarios of Figure 12 were produced.

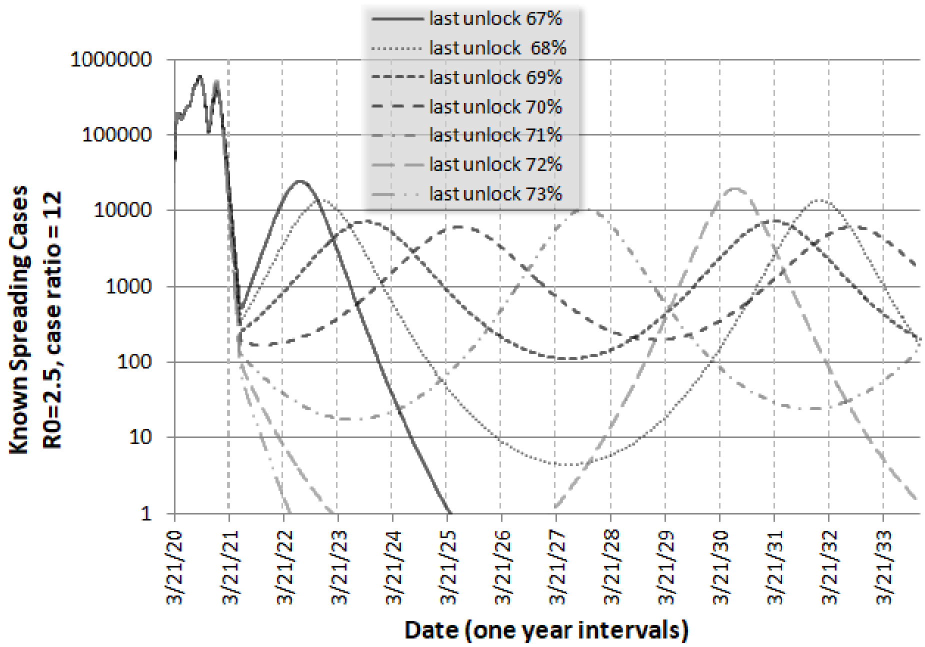

Figure 12. Possible endemic scenarios from a COVID-19-like pandemic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

The possibilities in Figure 12 were derived using US COVID-19 known cases through May 4, 2020, R_0=2.5 and a case ratio of 12. Similar possibilities occur for all plausible R_0 and case ratio values. The partial unlock values used were 10% May 1, 20% July 1, 30% August 15, “last unlock” from figure on Nov. 1, and full unlock at any time from early March 2021 through mid-November. None of these parameters are particularly critical. The endemic transition is caused by minimizing overshoot, arriving H_t with too few cases to assure the threshold is firmly passed, and the number of cases changing very slowly with time since R_t≈1 at that point.

For COVID-19, regardless of which unlock level is being used to target the threshold, about a 5% range which causes trouble is typical. The authors plan to publish a more extensive analysis separately. The problem is, the range is typically right around the most desirable target. In addition to the possibility planners may manage too well and create an endemic disease, the natural reaction of humans to avoid getting sick will reduce the number of cases. But it can take years for the number of cases to approach zero, including asymptomatic and undetected cases, and people are likely to relax their social distancing before that.

Notice that “known” cases can disappear for years at a time and then reappear. The numbers of deaths from the cycles are not insignificant, about 60,000 typically. With immunity wearing off whether natural or from vaccination the cycles will progress higher and faster. Vaccination coverage for existing diseases ranges from 35% to 85% (see https://www.who.int/news-room/fact-sheets/detail/immunization-coverage) and in some cases has taken nearly a century to reach those levels (tetanus).

5. Discussion (training and simulation)

In the fall of 2016 a SARS-like virus germinating among pig farms in Brazil spread to every country in the world. Nicknamed CAPS, it crippled trade and travel, sending the global economy into freefall. Social media was rampant with rumors and misinformation. The death toll was 65 million. Governments were collapsing and citizens revolting.

That sounds strangely familiar. It was an Event 201 simulation (see https://hub.jhu.edu/2019/11/06/event-201-health-security/). By the time this article reaches readers perhaps we will know whether the last ominous prognostication about governments collapsing was as accurate as the rest, and perhaps whether it would be attributable to plague deaths or economic disruption. It was held in New York City, hardest hit of the early COVID-19 hotspots in the US. There was another simulation held in New York in October of 2019 one month before COVID-19 actually developed. There was a hospital simulation in 2016 that concluded an Ebola patient whose diagnosis was known when transferred through an intermediary hospital would result in contagion [24]. Something is wrong with transferring simulation experience into practice. The authors are not talking about physical preparations, about whether there were stockpiles of protective gear and ventilators, but skill in assessing situations, determining what information to gather, and making decisions (or recommendations).

The success of flying large commercial airliners and manned space expeditions is entirely dependent on many hours of extremely realistic flying in simulators. The persons who are going to fly the craft in actual reality must be the ones who fly the simulators. The procedures are much like we have followed in this paper, but divided between operational personnel and simulation personnel who devise every conceivable scenario to throw at the operational personnel.

Operational personnel are not allowed to pick their favorite strategy or to imagine they have enormous stockpiles of resources. Every possible strategy is practiced in every possible condition, including inadequate resources. Event 201 just picks one scenario to practice every few years. It is a demonstration, perhaps a research event for academics and scholars, but not training. Training

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

would need to be part of public administration curricula, and then part of periodic drills conducted by governments like other disaster response drills. Epidemics have a much wider range of behavior than storms, at least as wide as chemical spills and require training on every possible strategy, not just benign ones. Non-government agencies like the WHO simply cannot adequately consider options that require substantial tradeoffs, such as loss from economic and civil disruption vs. direct loss from illness, because they aren't the elected representatives of the people affected.

Reactions must be automatic because panic will set in among the most seasoned professionals. Add to that the complexity that in a pandemic the final decision makers will be politicians who never heard of the simulations, who were perhaps just elected the month before, who may be under unrelated pressure, and operational personnel must be absolutely convincing. If a professional is not the decision maker, the professional should present all strategies even-handedly, not just the ones politically in vogue or supported by the current group think in public health circles. Public health organizations, even with part of government, are not chartered to consider the full picture. The decision maker is.

The economic and social responses should be part of the simulation training. At the US level, this means professionals from departments of Commerce, Education, the Treasury and Defense must be required to participate in the framework of something like a war game. Perhaps loans, interest and employment should be frozen. Ordering industries to remain open or produce certain products can be simulated, even partly carried out to produce equipment for stockpiles. The public can be involved as they were in air raid and nuclear attack and weather preparedness drills. It is likely COVID-19 has already done more damage than a small nuclear terrorist attack. New York has already experienced 6 times the deaths of 9/11/2001 with more to come.

From Table 1 the world fatalities could be from under 1 million to 80 million. One thing we see missing from the Event 201 and similar simulations, besides repetition and breadth of cases, is uncertainty over 2 orders of magnitude such as we have currently. Decisions under uncertainty are much more difficult. People have a natural tendency to guess at the underlying situation to simplify their decision, which produces narrowly optimized strategies prone to failure. What we have shown is that it is possible to devise strategies that do work over the entire range of Table 1. Inclusion of such methods in pandemic simulations would provide the opportunity to practice making such decisions. In such practice also comes tolerance for other people's choices, and the understanding that a diversity of approach may be best.

6. Conclusion

It has been suggested that COVID-19 and its related cousins (SARS, MERS) do not kill directly, as the virus count has already passed the peak when death occurs. The *coup de grâce* is delivered by the immune system fighting back too hard [5, 25]. Uncoordinated strategies that result in overshoot and up to double the number of deaths to achieve a targeted caseload amount to the same thing, dying from a poorly organized fight.

The goals of this paper were to demonstrate scenarios to accomplish most of the variety of public goals that have been articulated, from suppression to caseload management, from waiting in isolation to achieving herd immunity. We have demonstrated how to calculate the costs and how to save lives lost to overshoot, up to 50% in some cases.

It appears that different choices may be suitable for different regions and countries. Some with few cases can achieve suppression rather quickly. They will have to remain isolated until other countries achieve suppression or a vaccine is available. Herd immunity can be achieved at any target level of social distancing, not only at the natural unlocked level. Once it has been maintained long

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

enough for suppression, assuming there are not frequent transfers from animal reservoirs, then travel can resume between countries that used different approaches, possibly well before a vaccine is available.

In any epidemic with non-trivial mortality natural herd immunity is expensive in lives. If it lasts for some time, which has not at this writing been determined, it supplements any vaccine program. It also makes it easy to re-suppress new cases that escape containment measures. Countries with high compliance, strong social safety nets, and low numbers of cases may logically opt for suppression. Countries with high numbers of cases, poor safety nets, and populations at risk of dying from economic disruption and social unrest may opt for a minimum impact path to herd immunity, or a targeted level to stay within medical resources. Work on improved treatment and lower mortality will expand options for future pandemics more favorably than simply manufacturing medical equipment whose use only reduces mortality a few percent.

The success of pandemic management may lead to complacency, which may lead to slackening of efforts to go beyond the herd immunity threshold (whether by vaccines or actual cases). The result can be an annoying and deadly endemic disease. The authors suggest characterization of the conditions for this as a topic for future research.

Future pandemics might well have higher replication rates. A period of asymptomatic transmission would allow even a high mortality pathogen to spread like measles. As there appears no possibility the process of drug and vaccine development can operate at the short time scales of such a pathogen, it is important to learn how to implement less destructive social response, while being careful not to create new endemic diseases.

References

[1] Uddin, M., Mustafa, F., Rizvi, T..A., Loney, T., et.al. (2020). SARS-CoV-2/COVID-19: Viral Genomics, Epidemiology, Vaccines, and Therapeutic Interventions. *Preprints 2020*, 2020040005 https://doi.org/10.20944/preprints202004.0005.v1

[2] Fenichel, E.P., Castillo-Chavez, C., Ceddia, M.G., Chowell, G., et. al. (2011). Adaptive human behavior in epidemiological models, PNAS, 108 (15), 6306-11, March, https://dx.doi.org/10.1073%2Fpnas.1011250108

[3] Wilkinson, B. (2020). Quantifying the wellbeing costs of COVID-19, *The New Zealand Initiative*, Research Note, April 9, https://nzinitiative.org.nz/reports-and-media/reports/research-note-quantifying-the-wellbeing-costs-of-covid-19/

[4] Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W. and Shaman, J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2), *Science*, eabb3221, March, https://doi.org/10.1126/science.abb3221

[5] He, F., Deng, Y., Li, W. (2020) Coronavirus disease 2019: What we know? *Journal of Medical Virology*, no volume specified, 1-7, March 12. https://doi.org/10.1002/jmv.25766

[6] Marathe, M., Vullikanti, A.K.S. (2013). Computational epidemiology, *Communications of the ACM*, 56, 7, 88-96, July, http://doi.acm.org/10.1145/2483852.2483871

[7] Shuler, R.L. and Koukouvitis, T. (2020). Partial unlock caseload management for COVID-19 can save 1-2 million lives worldwide, *medRxiv*, April, https://doi.org/10.1101/2020.04.13.20064139

[8] Sutton, D., Fuch, K., D'Alton, M., Goffman, D., (2020). Universal Screening for SARS-CoV-2 in Women Admitted for Delivery, NEJM correspondence, *NEJM*, April, https://www.nejm.org/doi/full/10.1056/NEJMc2009316

[9] Silverman, J.D., Hupert, N., Washburne, A.D. (2020). Using ILI surveillance to estimate state-specific case detection rates and forecast SARS-CoV-2 spread in the United States, *medRxiv*, April, https://doi.org/10.1101/2020.04.01.20050542

[10] Baggett, T.P., Keyes, H., Sporn, N., Gaeta, J.M. (2020). COVID-19 outbreak at a large homeless shelter in Boston: Implications for universal testing, *medRxiv*, April, https://doi.org/10.1101/2020.04.12.20059618

[11] Bendavid, E., Mulaney, B., Sood, N., Shah, S., et. al. (2020). COVID-19 Antibody Seroprevalence in Santa Clara County, California, *medRxiv*, April, https://doi.org/10.1101/2020.04.14.20062463

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2020

[12] Fine, P., Eames, K. and Heymann, D.L. (2011). “Herd Immunity”: A Rough Guide, *Clinical Infectious Diseases*, 52 (7), 911-916, April, https://doi.org/10.1093/cid%2Fcir007

[13] Rashid, H., Khandaker, G., Booy, R. (2012). Vaccination and Herd Immunity: What More Do We Know? Curr Opin *Infect Dis*., 25 (3), 243-9, June, https://doi.org/10.1097/qco.0b013e328352f727

[14] Bulchandani, V.B., Shivan, S., Moudgalya, S., Sondhi, S.L. (2030). Digital Herd Immunity and COVID-19, *medRxiv*, April, https://doi.org/10.1101/2020.04.15.20066720

[15] Handel, A., Miller, J.C., Ge, Y., Fung, I.C. (2020). If containment is not possible, how do we minimize mortality for COVID-19 and other emerging infectious disease outbreaks? *medRxiv*, April, https://doi.org/10.1101/2020.03.13.20034892

[16] Fung, I.C., Antia, R., Handel, A. (2012). How to Minimize the Attack Rate during Multiple Influenza Outbreaks in a Heterogeneous Population, *PLOS/ONE*, June, https://doi.org/10.1371/journal.pone.0036573

[17] Handel, A., Longini, I.M., Antia, R. (2007). What is the best control strategy for multiple infectious disease outbreaks? *Proc Biol Sci*, 274 (1611), 833-837, March, https://dx.doi.org/10.1098%2Frspb.2006.0015

[18] Institute of Medicine (US) Forum on Microbial Threats (2007). Ethical and Legal Considerations in Mitigating Pandemic Disease: Workshop Summary, *National Academies Press (US)*, ch. 3, Strategies for Disease Containment. Available from: https://www.ncbi.nlm.nih.gov/books/NBK54163/

[19] Yu, X., et. al. (2008). Neutralizing Antibodies Derived From the B Cells of 1918 Influenza Pandemic Survivors, *Nature*, 455 (7212), 532-6, Sept., https://doi.org/10.1038/nature07231

[20] Wu, L., et. al. (2007). Duration of Antibody Responses after Severe Acute Respiratory Syndrome, *Emerg Infec Dis*., 13 (10), 1562-4, Oct., https://dx.doi.org/10.3201%2Feid1310.070576

[21] Hong Kong Polytechnic U. (2015). Immunity from H1N1 flu vaccine wears off rapidly, *Science Daily*, 2 November, https://www.sciencedaily.com/releases/2015/11/151102084629.htm

[22] Bauch, C.T., Lloyd-Smith, J.O., Coffee, M.P. and Galvani, A.P. (2005). Dynamically Modeling SARS and Other Newly Emerging Respiratory Illnesses: Past, Present, and Future, *Epidemiology*, 16 (6), 791-801, Nov., https://doi.org/10.1097/01.ede.0000181633.80269.4c

[23] Xu, S., Li, Y. (2020). Beware of the second wave of COVID-19, *The Lancet*, **395** (10233), 1321-2, April, https://doi.org/10.1016/S0140-6736(20)30845-X

[24] Lum, L.H., Badaruddin, H., Salmon, S., Cutter, J., Lim, A.Y., Fisher, D. (2016). Pandemic Preparedness: Nationally-Led Simulation to Test Hospital Systems, *Ann Acad Med Singapore*, **45** (8) 332-337, https://pubmed.ncbi.nlm.nih.gov/27683737/

[25] Pearson H., Clarke T., Abbott A., Knight J., Cyranoski D. (2003). SARS: what have we learned? *Nature*, 424(6945):121-126, June 30. https://doi.org/10.1038/424121a