Appendix:
Estimates of Required COVID-19 Testing for the US

April 8, 2020

1 Introduction

We are interested in how many people the US will need to test per day in order to control the COVID-19 pandemic. We present three estimates. The first is a simple extension to the standard susceptible-infected-recovered model used in the epidemiological literature. The second is an equilibrium analysis that derives the number of tests required to keep the number of new hospitalizations from exceeding the number of people recovering from hospitalization.The final approach uses the experience of Taiwan and South Korea to back-out a best-case number of tests per day for the US. Even assuming targeted testing, our main conclusion is that the US will need on the order of millions of tests per day to control the spread of COVID-19. Table 1 provides a breakdown of testing implications for each model.

An important caveat: none of the authors is an epidemiologist and we believe a more complex and realistic model developed by experts in the field will yield a more accurate estimate. Our goal is provide a simple baseline that we hope captures the first order testing requirements. We also hope that once testing becomes widespread, we will get a better handle of the true infection rate and so will be able to calibrate the model more accurately.

Model	**Required Tests per Day under Targeted Testing**
Targeted SIR Model	$1 - 10$ million
Equilibrium Analysis	~ 4 million
Taiwan and South Korea Case Study	~ 3 million

Table 1: Implied testing requirements by model

2 SIR Model with Targeted Testing

2.1 Model Setup

We build on the standard susceptible-infected-recovered model used in the epidemiological literature. Specifically, we allow for testing and a quarantined population. Our model is governed by the following set of differential equations:

$$\dot{S} = -\beta \times S \times \frac{I}{N} \tag{1}$$

$$\dot{I} = \beta \times S \times \frac{I}{N} - N_{test} \times f \times \texttt{posrate}\frac{I}{N} - \gamma \times I \tag{2}$$

$$\dot{Q} = N_{test} \times f \times \texttt{posrate} \times \frac{I}{N} - \gamma \times Q \tag{3}$$

$$\dot{R} = \gamma \times (I + Q) \tag{4}$$

$$H = \text{hosprate} \times (I + Q) \tag{5}$$

where:

- $I(t)$ is the number of infected people who are not in quarantine at a given point in time
- $R(t)$ is the number recovered or deceased people
- $S(t)$ is the number of individuals susceptible to being infected
- $Q(t)$ is the number of individuals in quarantine
- $H(t)$ is the number of hospitalized individuals
- N is the total population
- β is the number of people an infected person infects per day
- γ is the fraction of infected people that recover per day
- f is targeting efficiency defined as the ratio of the probability that the population given the test have the virus vs. general population
- `posrate` is the probability test gives positive result, given person has virus
- `hosprate` is the fraction of infected who require hospitalization
- N_{test} is the number of people to be tested per day

To simulate the model, we begin with a set of initial conditions and then integrate the differential equations above.

2.2 Calibration

In order to get an order of magnitude sense for how many tests the US will need to control COVID-19, we plug in reasonable parameter values and simulate the model. Our parameter choices are:

- $I(0) = N/1000$ is the number of infected people who are not in quarantine at the beginning of the simulation
- $R(0) = 0$ is the number recovered at the beginning of the simulation
- $Q(0) = 0$ is the number of individuals in quarantine at the beginning of the simulation
- $S(0) = N - I(0)$ is the number of individuals susceptible to being infected
- $N = 330$ million is the total population
- $\beta = 1/6$ number of people an infected person infects per day
- $\gamma = 1/12$ mean recovery rate of infected
- $f = 1, 10, 30$ ratio of fraction of people administered a test who have the virus with respect to the fraction of people in the general population who have the virus
- $\texttt{posrate} = 0.80$ probability that the test gives a positive result, given a person who has the virus
- $\texttt{hosprate} = 0.20$ fraction of infected who require hospitalization

One of the more important parameters is the extent to which the testing is targeted (f). Recall that f is the ratio of probability that people given a test have the virus vs. general population. A value of $f = 1$ corresponds to the case where we simply randomly sample from the population as a whole. We refer to this as "naive testing." A value of $f = 30$ corresponds to the case where people who are tested have 30 times higher probability of being infected than general population.

2.3 Simulation

We solve the model under different testing assumptions and show the result in Figure 1. We see that the size of the hospitalized population falls off with increased testing, as expected. Further, the number of tests required under naive testing strategies (low f) is dramatically higher than for more targeted testing (high f). Precise values of f are likely to vary by geography and over time, but this plot shows that keeping them as high as possible is important. For approximate values of f (between 3 and 30), we see that between 1 and 10 million tests/day will be required in order to keep the peak hospital population below

1 million. Although high f is desired, tests should also be spread out as much as practical while keeping a high f value — oversampling a single cluster will result in unchecked growth in other regions and degrade the advantages provided by the testing.

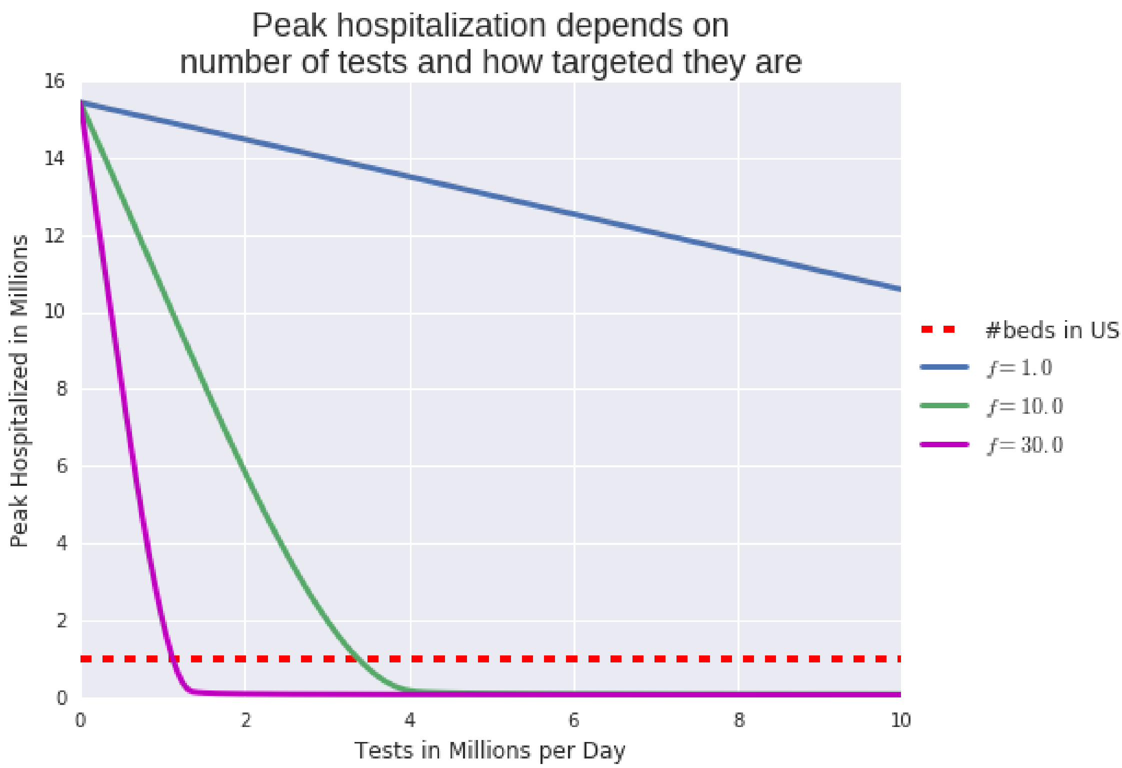

Figure 1: Maximum fraction hospitalized as a function of number of tests

In this figure, we summarize the outcome of our model. The various lines show the number of hospital beds required to treat COVID patients under various different targeting efficiencies f (larger is more targeted) as we increase the number of tests that can be performed per day. The total number of hospital beds is included as a dotted line; of course, the number of beds available for COVID patients is a subset of all beds. From this plot, we can read off our headline numbers: for reasonably efficient targeting, f around 10 to 30, a few millions of tests per day will be needed to keep total hospitalizations below the maximum.

3 Testing Requirement Based on Equilibrium Analysis

3.1 Summary

We derive the minimum number of daily tests required to ensure that hospitals will be able to stay within capacity throughout the COVID-19 crisis. We consider two equilibria:

1. the largest possible size of infections within the population at which hospitals can still function

2. the required number of daily tests to ensure that the number of infections does not grow beyond this limit.

The main takeaway is that the number of tests estimated by this method is **at least 41 million per day with random testing** or **at least 4.1 million per day with targeted testing** (for test efficiency $f = 10$). The number of required daily tests can be expressed as:

$$\text{Number of Daily Tests needed} \geq \text{Size Of Population} \times \frac{1}{f} \times \frac{1}{\text{posrate}} \times \tag{6}$$

$$(\text{rate of spread} - \text{rate of recovery/removal}) \tag{7}$$

where:

- Size Of Population = size of the US population, assumed to be 330 million

- rate of spread = how many people a spreader (= infected and not quarantined person) infects per day, assumed to be 1/6 because every 6 days one new person gets infected

- rate of recovery/removal = fraction of spreaders that recover or die per day, assumed to be 1/15 as it takes 15 days to recover/die

- f = target efficiency = ratio of probability that people given test have virus vs. general population

- posrate = probability test gives positive result, given person has virus, assumed to be 80%.

Dependence on targeting efficiency:

Targeting efficiency	1	10	30
Daily tests needed	41.3Mn	4.1Mn	1.4Mn

3.2 Estimation

We need to hold the total number of new infections per day below a certain threshold or we won't have enough beds in hospitals. The number of new beds needed, N_{needed}, is the number of newly infected $N_i \cdot r_s$ times the hospitalization rate, h. Here N_i is the number of currently infected people, and r_s is the number of people who would be infected for every person currently infected. The number of beds freed is the total number of beds B_{total} times the rate of recovery/removal r_r (we include deaths here to keep notation simple). Therefore, we have

$$B_{\text{needed}} = B_{\text{free}} \quad \Rightarrow \quad N_i \times r_s \times h = B_{\text{total}} \times r_r \tag{8}$$

This relation can be used to give a bound on the number of currently infected people, at which hospitals run at maximum capacity

$$N_i \leq \frac{B_{\text{total}} \times r_r}{r_s \times h} \tag{9}$$

When equality holds in equation 9 we call the above condition the **critical equilibrium of maximum hospital capacity**.

Now we study the condition under which the number of spreaders will stay at the critical equilibrium (equation 9). Crucial here is a **large enough number of tests**. Recall that $N_s(t)$ is the number of *spreaders* on day t. The number of spreaders $N_s(t+1)$ on the next day is given by

$$\begin{aligned} N_s(t+1) =& (1 - \texttt{rate of recovery}) \times N_s(t) + (\texttt{rate of spread}) \times N_s(t) \\ & - (\texttt{number Of Daily Tests}) \times \frac{1}{f} \times \frac{1}{\texttt{posrate}} \times \frac{N_s(t)}{\texttt{Size Of Population}} \end{aligned} \tag{10}$$

It is composed of 3 parts (corresponding to the 3 summands in the above expression):

1. Number of spreaders that are *still* infected: $(1 - \text{rate of recovery})N_s(t)$
2. Number of newly infected people: $(\texttt{rate of spread})N_s(t)$
3. Number of infected people that are are removed from the population of spreaders because they have been diagnosed by testing: f = test efficiency, assumed 1 for random testing, 10 for targeted testing, `posrate` = probability test gives positive result, given person has virus, assumed to be 80%.

In order for the number of spreaders to not grow (and in particular not grow at the critical threshold (equation 9), we need $N_s(t+1) \leq N_s(t)$. Using equation 10, the inequality becomes the **minimum required testing condition**

$$\text{Number of Daily Tests needed} \geq \text{Size Of Population} \times \frac{1}{f} \times \frac{1}{\text{posrate}} \times (\text{rate of spread} - \text{rate of recovery}) \tag{11}$$

4 Minimum Tests for Partial Social Distancing: the Case of Asian Countries

4.1 Introduction and Summary

Several Asian nations have managed to control the spread of COVID-19 without fully going to stay-at-home orders. We take Taiwan and South Korea as models of this approach. We try to extrapolate some lower bound on the level of testing needed in the US from the level of testing used in South Korea and Taiwan.

To complement the structural modeling approaches outlined elsewhere in this white paper, we can take the South Korean/Taiwan (hereafter, SK/T) numbers and scale them to the scale of the problem in the United States. This approach can give us an optimistic **minimum** number of tests needed for a country the size of the US with cases on the scale of the US. This approach does not try to account for all the differences between the US and SK/T and assumes that such details will tend to average out in the final estimate. We expect some of the particularities of the US to push the estimate up, while other particularities (lower density) to push the estimate down. Finally, we emphasize that this is an estimate **assuming that there is still partial social distancing**. Return to "normalcy" will necessarily require more tests than this, perhaps by an order of magnitude or more.

In the sections below we estimate that based on the level of testing in SK/T the US would need several million tests per day,

$$R_{\rm USA} \simeq \texttt{3Mn tests /day} \tag{12}$$

This number should be understood in the context of the current level of testing in the USA of about 100,000 per day.

4.2 Basic Data Used

As of April 6th, the number of confirmed cases in SK is 10,248, while the number of tests conducted are 443,000 to date. The testing was done over a 40-day period with essentially 10,000 tests a day.

`https://www.statista.com/statistics/1102777/south-korea-covid-19-daily-new-cases/`

`https://ourworldindata.org/covid-testing`

The situation in TW is a bit less clear, but appears to be about 400 cases and 20,000 tests. The time scale over which these tests were conducted is not clear.

`https://focustaiwan.tw/society/202003240013`

Currently there are 300,000 known cases in the USA, and about 1.27 million people tested.

4.3 A Naive Scaling Argument

It seems plausible that for effective mitigation the number of tests needed is roughly proportional to the number of known cases:

$$N_{\text{test}} = \alpha N_{\text{known}}$$

Based on South Korea (10,000 cases with roughly 400,000 tests) and Taiwan (400 cases with roughly 20,000 tests) we estimate that $\alpha \sim 50$. This seems reasonable and can be interpreted as the average number of people any of the N_{known} cases could have interacted with. Note that the total size of the population does not play a role here as we assume that some sort of localization of the problem is still feasible ("contact tracking").

For the US:

$$N_{known} = 300,000$$

Which would suggest that a total of,

$$N_{test} = 15\texttt{Mn}\ \ (\alpha \sim 50)$$

This is the total number of tests we would need to perform assuming that we start today with 300,000 known cases and a total of 1 million tests already taken (which we neglect as it is much smaller than the number needed).

To translate this number to the number of tests needed per day, we need some notion of time-scale in the problem. One natural time-scale is the doubling time, namely how long would it take before the number of known cases doubles (go from 300k to 600k in the US). This number is currently estimated at $\tau = 5$ days for the US. A testing scheme that would take much longer than this would hardly have any relevance since the number of cases would be exponentially larger by the time it is concluded. So an order of magnitude estimate for the number of tests per day to come anywhere close to dealing with the problem is:

$$R_{\text{USA}} = \frac{N_{test}}{\tau} = \frac{\texttt{15,000,000}}{5} \simeq \texttt{3Mn / day}$$