Putting it all together
We can estimate the SIR values by using the relationships among them and information from clinical studies. The *New York Times* has an interactive online tool (nyti.ms/3CfqpNf) which allows a user to see immediately the effect of changing any of these parameters.

Covid-19 can be understood as a generic epidemic which has its own values of these parameters. Of course, policy and human behavior influence the parameters too, by reducing transmission through social distancing, reducing fatalities through better treatment, and ultimately, reducing susceptibility through a vaccine. The processes over time can be seen in Figure 1 (page 13).

The top graph shows the number of deaths each day, which is the only measure we can really observe. The middle graph shows the number of new infections each day. Note that it leads the deaths by a couple of weeks, and the length of this lead is another variable in the model which can only be observed through limited clinical studies. The bottom graph shows the SIR values: the susceptible population (green line) starts with everyone and declines over time. The removed group (orange line) rises for a while then slowly declines. The removed population (blue line) rises over time, eventually including everyone – the dead and the survivors.

Different modelling projects approach this framework differently. Which parts are assumed, measured, or modelled vary among different studies. Some models let the interactive user guess different values of $R_0$ (and thereby $\beta$) or $p$ (and thereby $\gamma$), while other models incorporate measures from small clinical studies. Some models include an intervening term, exposure, between the susceptible population and infection (not all susceptible people are exposed, and the unexposed people cannot become infected). The mathematics that connects all the pieces is also different in different models.

In the long term, we will learn which models were best. However, time is too short for more than a tiny number of these models to be subject to formal peer review in time to be relevant. That means it is more important than ever that engaged laypeople (especially journalists) have at least a minimum sense of how to read these essential studies.

**Reference**

1. Johndrow, J., Lum, K. and Ball, P. (2020) Estimating the number of SARS-CoV-2 infections and the impact of social distancing in the United States. Preprint, arXiv:2004.02605v1. bit.ly/2Y1zLl

**Further reading on the SIR model**

- Kermack, W. and McKendrick, A. (1927) A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London Series A: Mathematical and Physical Sciences, 115, 700–721.
- Anderson, R. M. and May, R. M. (1991) *Infectious Diseases of Humans*. Oxford: Oxford University Press.

**Acknowledgements**

I am interpreting these ideas from a paper by James Johndrow, Kristian Lum, and me. I am grateful for their comments, as well as for suggestions from Maria Gargiulo, Megan Price, Tarak Shah, Stella Pierce, Danielle Fugere, Hope Howard, and Jacob Nelson.

---

**Why we need more coronavirus tests than we think we need**

**James J. Cochran** on the importance of testing a random sample

In the United States (as of 9 April 2020), President Donald Trump has said that testing for novel coronavirus infection will be limited to people who believe they may be infected. But if we only test people who believe they may be infected, we cannot understand how deep the virus has reached into the population. The only way this could work is if those who believe they may be infected are representative of the population with respect to novel coronavirus infection. Does anyone believe this is so?

The common characteristic of those who believe they may be infected is that they all show some outward symptoms of infection by the virus. In other words, people who are being tested for the novel coronavirus are disproportionately showing severe symptoms. This would not be a problem if someone who is infected by the novel coronavirus immediately shows symptoms, but this is not the case. We have strong evidence that some people develop mild cases, show no symptoms, and carry the virus without knowing it because they are asymptomatic. Thus, efforts to understand the virus’s penetration into the population must include observation of the asymptomatic.

The estimate of the proportion of the population who are infected can be calculated as:

$$\hat{\beta} = \frac{\text{number of symptomatic infections} + \text{number of asymptomatic infections}}{\text{number not infected}}$$

So, we need data from a random sample of the entire population in order to gather data from infected people who are showing symptoms, infected people who are asymptomatic, and people who are not infected. All have some probability of being included in a true random sample of the population.

As of 23 April, leaders in Germany and New York State (see bit.ly/2Kp2ixd and dailym.ai/3bxZ5Au) had moved to implement random testing to assess how widespread the virus is, but there has been resistance from leaders elsewhere. This could be due to ignorance, disregard, or lack of appreciation of...
Tests in short supply? Try group testing

Christopher R. Bilder, Peter C. Iwen, Baha Abdalhamid, Joshua M. Tebbs and Christopher S. McMahan explain how, by pooling specimens, testing capacity for SARS-CoV-2 can be increased.

Cough, fever, and difficulty breathing are all symptoms of Covid-19, but symptoms alone may not qualify someone to be tested for SARS-CoV-2. In the USA (as of 9 April 2020), only those individuals who meet stringent requirements, such as hospitalisation with no other pathogen detected or exposure to an individual with Covid-19, are tested due to the shortage of testing resources. Shortages like this around the world have hampered efforts to understand Covid-19 and to prevent its spread, especially by asymptomatic individuals.

In similar situations where resources are in short supply, a procedure known as group testing (also known as pooled testing and specimen pooling) is often employed. Its most basic implementation begins by obtaining specimens from a set number of individuals and then pooling parts of each specimen into a “group” for a single test. If the group tests negative, all of its members are declared negative. If the group tests positive, each member has the remainder of their original specimen tested separately to determine the positive/negative outcome. This process of forming groups and testing is repeated over all individuals that need to be tested. Robert Dorfman originally proposed this particular testing algorithm to screen US soldiers for syphilis during World War II, and the process is often referred to now as “Dorfman testing”.

When disease prevalence in a population is small (less than 15%), group testing results in a significant reduction in the number of tests performed. Group testing is used in a vast number of applications, including blood donation screening for infectious diseases (rcblood.org/3eQ2NHV), sexually transmitted disease detection, bacteria in food detection, and chemical compound discovery for use in new pharmaceuticals. Group testing was even investigated for the influenza pandemic (H1N1) of 2009.

There are two important considerations for applying group testing. First, pooling specimens leads to each individual specimen being a smaller portion of the whole. This dilution could make it more difficult to detect the pathogen, thus potentially producing false negatives. Second, a group size needs to be selected. Too large a group size results in too many groups testing positive, leading to an exorbitant number of separate tests. Too small a group size results in many more group tests than needed. To prevent this from occurring, the expected number of tests is used to determine the most efficient group size. In other words, we want the group size that gives the smallest average number of tests if group testing was applied to a continuous stream of specimens coming into a laboratory. Our box (page 16) presents the technical details for selecting this group size.

The Nebraska Public Health Laboratory has been a leader among laboratories with regard to applying group testing to detect SARS-CoV-2. Their positivity rate was estimated to be 5% from initially testing specimens separately. The most efficient group size for Dorfman testing is 5 with this rate, resulting in 57% fewer tests on average than when testing specimens separately.

Validation of the group testing process showed...