The Coronavirus, by the Numbers

The New York Times Science Section, Tuesday, March 10, 2020 carried an important article on the Coronavirus. It speaks to various factions such as length of infectious stage, transmission, and serious cases.

Children are much less susceptible than older people, author Adam Kucharski, mathematician who studies the spread of disease. Interview by James Gorman.

The Interview

We hear a lot about the percentage of sick people who are dying. Is that the case fatality rate?

The case fatality rate measures the risk that someone who develops symptoms will eventually die from the infection.

And how is that rate calculated?

Ideally, we would monitor a large group of people from the point at which they develop symptoms until they later die or recover, then calculate the proportion of all these cases who had died.

So can we just look at the total number of deaths and the current number of cases?

The problem with just dividing the total number of deaths and total number of cases is that it doesn't account for unreported cases or the delay from illness to death. The delay is crucial: If 100 people arrive at hospital with Covid-19 on a given day, and all are currently still alive, it obviously doesn't mean that the fatality rate is 0 percent. We need to wait until we know what happens to them eventually.

Any deaths will be people who got sick two to three weeks ago, so it's not simply deaths at the moment divided by cases at the moment. If you have two deaths from two cases, as happened in Iran last month, that most likely means you've missed a bunch of cases.

We've seen all sorts of numbers for fatality rates. Does the latest estimate of 3.4 percent globally make sense?

Early on, people looked at total current cases and deaths, which, as I said, is a flawed calculation, and concluded that the case fatality rate must be 2 percent based on China data. If you run the same calculation on yesterday's totals for China, you get an apparent CFR (case fatality rate) of near 4 percent. People are speculating that something is happening with the virus, where it actually is just this statistical illusion that we've known about from Day 1. I'd say on best available data, when we adjust for unreported cases and the various delays involved, we're probably looking at a fatality risk of probably between maybe 0.5 and 2 percent for people with symptoms.

What about R, the reproductive number, or how many people a given patient is likely to infect. Why is it important?

At its simplest, R is the answer to the question: How worried should we be about infection? If R is above one, each case, on average, is giving it to at least one other person. You're going to see growth.

If it's less than one, then a group of infected people are generating less infection. From a policy-planning point of view, it gives you a very clear objective. For example, in the Ebola response in 2014, it was a really prominent part of the response. The aim was to get R below one.

That seems very simple and straightforward, but you write that it's more complicated than it seems. In your book you say that to calculate R you've got to know duration, opportunity, transmission probability and susceptibility (the "DOTS"). Let's take them one by one. What is duration?

How long someone is infectious. If someone is infectious twice as long, that's twice as long they are around to spread infection.

Do we know what the duration is for this coronavirus?

On average, we'd probably be looking at a week or two. Of course, if people get hospitalized, then they're not in the community spreading infection in the same way.

The second component is opportunity. How do you determine that?

That's a measure of how many people you come into contact with for every day you're infectious. With something like flu, you're not infectious very long but a lot of your interactions could potentially spread it. With something like H.I.V., the duration is much longer but the number of sexual partners you have relative to the number of conversations you have is obviously much lower.

And transmission probability?

This is a measure of the chance the infection will get across during an interaction. For example, during a sexual encounter, the virus won't necessarily get across.

Finally there's susceptibility.

Susceptibility measures the chance the person at the other end of the interaction will pick up the infection and become infectious.

What's the equation to come up with R?

If you multiply them together, you get the reproduction number. So if you scale up or scale down any one of these things, it directly affects the value of R.

If you were the average person, what would you pay attention to — in terms of the news and the numbers?

One signal to watch out for is if the first case in an area is a death or a severe case, because that suggests you had a lot of community transmission already. As a back of the envelope calculation, suppose the fatality rate for cases is about 1 percent, which is plausible. If you've got a death, then that person probably became ill about three weeks ago. That means you probably had about 100 cases three weeks ago, in reality. In that subsequent three weeks, that number could well have doubled, then doubled, then doubled again. So you're currently looking at 500 cases, maybe a thousand.

I think the other thing that people do need to pay attention to is the risk of severe disease and fatality, particularly in older groups, in the over-70s, over-80s. Over all we're seeing maybe 1 percent of symptomatic cases are fatal across all ages. There's still some uncertainty on that, but what's also important is that 1 percent isn't evenly distributed. In younger groups, we're talking perhaps 0.1 percent, which means that when you get into the older groups, you're potentially talking about 5 percent, 10 percent of cases being fatal.

In thinking about social behavior and thinking about your interactions, the question should be, "How do we stop transmission getting into those groups where the impact could be really severe?"