ISSUES IN SCIENCE AND TECHNOLOGY

JANUARY 30, 2020

Clarity, Please, on the Coronavirus Statistics

BY MAIMUNA S. MAJUMDER

Sorting through the different ways of describing the "deadliness" of infectious diseases.

While the ongoing novel coronavirus (2019–nCoV) outbreak continues to grow, one of the questions on the minds of both researchers and the general public alike has been, How deadly *is* this thing?

With 5,974 confirmed cases and 132 deaths reported so far out of China alone (as of the end of the day on Tuesday, January 28), it would seem as though making an assessment would be easy. It turns out, however, that questions such as this one are misleading in their simplicity. Here's why:

(1) **There's more than one way to describe the "deadliness" of a given disease.**

Two of the most common ways to describe the deadliness of a disease are its *population mortality rate* and its *case fatality rate* (CFR). Both are ratios that consider the number of

deaths that have occurred as the numerator—but their denominators are different.

For the population mortality rate, the denominator is the entire population at risk. So in an infectious disease outbreak, this statistic approximates an answer to the question, "For any individual in a given population, what is the chance of dying of the infectious disease?"

By contrast, CFR approximates an answer to the question, "For any individual diagnosed as infected, what is the chance of dying of the infectious disease?" When calculating CFR, the denominator includes only diagnosed cases. Cases that have not been diagnosed—that is, mild and asymptomatic individuals who don't seek care—are not typically included in this calculation until after seroprevalence surveys can be conducted. Seroprevalence surveys allow us to discover previously unidentified mild and asymptomatic cases—at which point, CFR estimates for the outbreak may be revised downward.

Furthermore, when calculating CFR, the denominator includes only cases that have resolved—that is, cases where the patient has either died or recovered. This is why dividing total deaths by the total number of cases—a simple calculation that many media outlets have been wont to do with 2019–nCoV—isn't an accurate indication of CFR until the *end* of an outbreak. Before the end of an outbreak, dividing the total deaths by the total number of cases may be better defined as the *proportion of fatal cases* (PFC) rather than as the CFR.

(2) **Outbreak dynamics inherently affect CFR estimates over time.**

At the beginning of an outbreak, severe cases are more likely to be identified than less severe ones. Less severe cases are more likely to be identified later in the outbreak due to improved screening and awareness. Because severe cases are at a higher risk of death than less severe ones, CFR estimates made at any point prior to the end of the outbreak—that is, before all cases have resolved—are likely to change with the passage of time.

In the context of the ongoing 2019–nCoV outbreak, and focusing just on China, only 235 cases have resolved as of the end of January 28, 2020—132 (56%) of which have died and 103 (44%) of which have been discharged after recovery. However, a whopping 5,739 remain hospitalized—which is in large part why we should expect these numbers to fluctuate considerably as the outbreak progresses. Notably, only 21% of the confirmed cases that are currently hospitalized are considered "severe," which suggests that many more may recover in the days and weeks ahead.

Given these numbers, we can at least calculate the proportion of fatal cases, which, as mentioned above, is a somewhat cruder statistic that divides the total deaths by the total number of cases. As of now, the PFC can be calculated as about 2%—although it, like the CFR, will continue to fluctuate until the end of the outbreak, when the two figures will ultimately converge.

(3) Case fatality rates aren't uniform across a given population.

Though public health scientists often calculate CFR based on statistics from all affected patients, the reality is that the risk of death from a given infection is rarely uniform. Certain subsets of the patient population are usually more likely to die than others are—which means that how deadly 2019–nCoV is may very well vary across demographics. We've actually seen this phenomenon before with the related SARS coronavirus in 2003, where the case fatality rate exhibited age-dependence.

As more information becomes available, we'll be able to make stronger scientific assessments regarding the fatality (and mortality) rates associated with this novel pathogen. For now, though, let's do our best to be thoughtful, careful, and vigilant in our discussions about 2019–nCoV—both regarding its lethality and otherwise.

Maimuna S. Majumder *is a faculty member in the Computational Health Informatic Program based out of Boston Children's Hospital and Harvard Medical School. Twitter:* @maiamajumder.

YOUR PARTICIPATION ENRICHES THE CONVERSATION

Respond to the ideas raised in this essay by writing to forum@issues.org.

CITE THIS ARTICLE

Majumder, Maimuna S. "Clarity, Please, on the Coronavirus Statistics." *Issues in Science and Technology* (January 30, 2020).

555 N. CENTRAL AVE., SUITE 302, PHOENIX, AZ 85004-1248