

Many of us anxiously scan coronavirus statistics, looking at trends and cross-country comparisons. Warning: We need to be cautious in interpreting those numbers. There's lots of noise in the data, meaning that it's not always an accurate measure of what we want to know about the disease. Even death counts are not always reliable — Florida has had about 1200 extra pneumonia deaths this year, which are [probably](#) undercounted coronavirus cases.

I want to make this as clear as I can to people who don't know or care about statistics. I'll avoid the usual statistical terminology, which is needlessly confusing until you get the hang of it. What we're interested in are testing errors. No test is perfect. It may identify someone as having the disease, even though they actually don't. Let's call that a false alarm. Or it may say someone is disease-free when they actually have the disease. Let's call that an alarm failure.

There are two types of tests for coronaviruses. The first kind of test looks for the RNA of the virus. (These are called RT-PCR tests, if you're interested.) That's the kind of test that's used to see if people are currently infected. There are a lot of different tests out there, and because of the urgency of the crisis, the FDA hasn't been screening them as tightly as they might otherwise. The big problem with these tests seems to be alarm failures. That is, the tests say someone *doesn't* have the coronavirus when they actually do.

A recent [study](#) in the *Annals of Internal Medicine* illustrates the problem. This study synthesized data from seven researchers to examine the rate of alarm failures. On the day symptoms began, the study found on average a 38% rate of alarm failures. (It was even higher for non-symptomatic patients.) The best results were achieved when tests were given three days after symptoms began, with an average 20% rate of alarm failures. The rate of alarm failures then started going back up again when tests were done after three days.

What this means is that the tests are significantly underestimating the number of infections. How much depends partly on when the test is given during the patient's illness. That's why statistics may include probable cases, where everything but the test result point toward coronavirus. If everybody used exactly the same protocol, using the same tests on the same population of patients on the same day after on-set of symptoms, we could be pretty comfortable that trends or cross-country comparisons were accurate. But there are differences between countries and maybe within the United States in the mix of tests and in testing protocols. So we can't be as confident about trends and comparisons in the number of active coronavirus cases.

The other kind of coronavirus tests look for antibodies to determine whether someone has

had the coronavirus. This is important for understanding how deadly the disease is, because otherwise we have no way of knowing about people who had mild cases or no symptoms at all. It also tells us whether we're anywhere close to the "herd immunity" you hear so much about. The problem with these tests is false alarms. That problem is especially severe because the percentage of people who have actually been infected is still quite low except in a few hot spots like New York City.

There are lots of nice explanations of this using numerical examples. But I'll try to explain it without numbers. Suppose you're a burglar alarm company and everyone in town has one of your alarms. The alarms never fail to detect a burglary: if there's a burglary, the alarm will go off. But the alarms may sometimes go off even if there's no burglary.

The odds that any given house's alarm system will fail on any given night might be quite small. On the other hand, there are a whole lot of alarm systems in town. That means that on any given night, the alarm company is likely to be chasing some false alarms.

If the burglary rate in the town is very low, you'll only get a few alarms for actual burglaries every night. A fair share of the alarm calls the company gets are going to be false alarms. In other words, the number of alarms going off may overestimate the number of actual burglaries pretty substantially.

That's the problem with the antibody tests. One of those [tests](#) is 95% accurate, which means there's only a 5% chance that it says someone has had coronavirus when they really haven't. That sounds great. But think back to the burglar alarms. If 5% of the alarms go off every night even when there isn't a burglary, the alarm company is probably going to be seeing a huge number of false alarms in proportion to the actual number of burglaries. So counting the alarms going off is going to badly overestimate the number of actual burglaries in town.

Similarly, at a time when not many people have actually had the coronavirus virus in some areas, the antibody tests are going to really inflate that number. It's hard to correct for this unless we know how many people have actually been infected from some other source. And again, we can't be sure whether the testing in different locations or times is using the same tests and the same procedures for picking people to test.

And there's another problem, which is data reporting. We know that in some places, the authorities have tried to conceal coronavirus statistics to make things look better. It also turns out that in several states, and at the CDC level, results from the two kinds of tests, the RNA tests on sick people and the antibody tests on the general population, have been

combined in [reports](#) on the outbreak. That makes them really hard to interpret. Apparently that is being straightened out now that it has become public. But we don't know what other problems lurk in the reports from states and the CDC.

I certainly don't mean to suggest that we should ignore these numbers. Most of the problems I've discussed suggest that the situation is actually worse than the numbers suggest. If nothing else, the numbers give us a benchmark. We should treat the numbers as important information, but not gospel.

The experts who analyze these statistics are aware of these issues, and no doubt they're doing their best to control for all this stuff. Presumably this is one reason that model forecasts have had large uncertainty ranges. But making these adjustments is necessarily an inexact science, and surely not something that we lay people can do.

In short, we should all try to resist our urge to play amateur epidemiologist, or at least to remember how shaky our amateur conclusions may be. We should also keep in mind that there are limits to what even the experts can deduce from the noisy data. This really is like the "fog of war," only here it's disease spread that's hard to nail down.