

The Washington Post

Democracy Dies in Darkness

We still don't know how many people are infected with covid. We can find out.

Current statistics used to track viral spread are flawed.

By **Len Covello**, **Andrew Gelman** and **Yajuan Si**

Jan. 25, 2021 at 5:21 p.m. UTC

The current way we track the prevalence of coronavirus infections is deeply flawed.

Ideally, health officials would test random samples of citizens in each community in a systematic way. But throughout the pandemic, the United States has lacked the political will or funding to pursue it. Instead, testing tends to ramp up when there are symptomatic outbreaks, which prevents the data from being representative of the population as a whole. Other factors compound that flaw: Outbreaks tend to cluster geographically and ethnically, for instance. What's more, if a state tests more people, it will tally more cases, and if clinicians get better at identifying the disease, and test the people they think have covid-19, the test-positivity rate (another metric used to gauge viral spread) also increases.

Until we get to widespread random testing, we propose a second-best methodology that — our data shows — outperforms current practices as a predictor of future burdens on the health system. Our method involves testing asymptomatic people who are visiting hospitals for a broad range of outpatient procedures — diagnostic as well as surgical — and then adjusting the rates of positive testing to match the area's demographics. Using data from the Community Hospital network in northwest Indiana, we analyzed the relationship between rates of positive test results for this outpatient population and covid-related admissions at all hospitals in the five Indiana counties the hospital network serves from late April into this month.

Our method predicted rises (or falls) in hospitalizations seven to 10 days before they occurred. In contrast, official state statistics for positive cases in the area lagged our data by roughly a week. State test-positivity rates — the proportion of tests given that were positive — were even less useful as a predictor of trends. In other words, switching to this new methodology more broadly could give policymakers at least a week's jump on coronavirus trends, crucial time to prepare.

The idea for this new measure of coronavirus spread arose out of practices undertaken to reopen hospitals last spring after the initial shutdowns: Specifically, in April, Community Hospital in Munster, Ind. — where one of us works — was faced with the need to deliver surgical and diagnostic services to a patient population that obviously contained some infected but asymptomatic people. (Symptomatic people could be screened out more readily.) Only by testing everyone could the hospital be sure it wasn't exposing staff and other patients to a potentially fatal illness. As a result, patients scheduled for procedures were required to say whether they had symptoms of covid-19 and contact with anyone with the disease. They also had to test negative for viral RNA four days before surgery. That regimen has continued.

As a side benefit, this protocol gave us an outstanding chance to measure viral prevalence. Demographically, these patients have an age, gender and racial-ethnic distribution that is similar to (though not identical to) the community at large. That they were asymptomatic was valuable, too: Testing them was more like testing people out and about at a local mall than like testing people who were already experiencing a fever and a cough. Overall, we collected information on 23,400 asymptomatic outpatients over the period we studied. (We also looked at the test results for more than 9,000 symptomatic outpatients, to explore how well they matched the demographics of the surrounding community; they did not match it well at all.)

The differences between the hospital population and the surrounding community were small but important: These patients are somewhat older and whiter, for instance. But statistical techniques allowed us to appropriately adjust what we found in the hospital setting. Since the hospital outpatient population contains fewer young people than the population at large, and younger people who are infected tend to be asymptomatic, we know our sample has fewer asymptomatic people than it "should." We reweighted to account for that, and adjusted appropriately for other over- and underrepresented populations.

The method we used — called multilevel regression and poststratification — is simple, easily duplicated in any hospital system (they're already doing the testing), and inexpensive. (To make implementation elsewhere easy, we have made the statistical demographic adjustment [available](#) online.) Our current work would imply that the same methodology can also be used to keep track of antibody prevalence, both naturally- and vaccine-acquired, which will be important as the vaccines roll out.

At 3,500 deaths daily, a week's improvement in forecasting — the advantage our model provides — is an eternity, so this new metric promises to be useful as a viral surge unfolds. And from a policy standpoint, it's just as important to know when it is safe to open as when to shut down. Our model can help there, too. The new metric is clearly superior to population-wide case numbers and especially positivity rates in showing clinical decrease in the virus. In a recent six-week period, for instance, the area around Munster saw a dramatic downturn in covid-related hospitalizations and ER visits — a trend foreshadowed by our data. This occurred even though positivity rates recorded by the state remained fairly high. That pattern is almost certainly replicated in many areas around the country with persistently high positivity rates.

While Indiana's policies have kept our restaurants and bars largely open during this interval, other presumably well-meaning state and local governments mandated closures of businesses like these under parallel circumstances. Throughout our country, overreliance on positivity rates to track virus behavior may be economically and socially damaging.

We should still strive for widespread random testing of the general population. But until we have the resources, we can leverage already existing hospital testing to get more reliable information about our adversary — and win this war.

Updated January 22, 2021

Coronavirus: What you need to read

The Washington Post is providing some coronavirus coverage free, including:

Coronavirus maps: Cases and deaths in the U.S. | Cases and deaths worldwide

Vaccine tracker: See how many doses will be available in your state

What you need to know: What you need to know about the vaccines | Covid-19 symptoms guide |
Coronavirus etiquette | Your life at home | Personal finance guide | Make your own fabric mask | Follow

[Show More](#)
