


Accurate Statistics on COVID-19 Are Essential for Policy Guidance and Decisions

 See also Morabia, p. 923, Tarantola et al., p. 925, and the *AJPH* COVID-19 section, pp. 939–977.

Disease surveillance forms the basis for response to epidemics. COVID-19 provides a modern example of why the classic mantra of “person, place, and time” remains crucial: epidemic control requires knowing trends in disease frequency in different subgroups and locations. We review key epidemiological concepts and discuss some of the preventable methodologic errors that have arisen in reporting on the COVID-19 crisis.

NUMBERS VS RATES

By “frequency” we mean the attack rate over a defined period: the number of COVID-19 cases (numerator) divided by the population size (denominator). The size and source of the denominator is important; for example, a headline proclaiming “Italy Surpasses China” based on total case counts is misleading because, compared with Italy, China has about 24 times the population (the “surpassing” thus happened much earlier); it is also younger in age distribution and covers 32 times the area with far more extensive geographic and ethnographic diversity. Thus, even if the test is perfect, case count comparisons and their trends across populations and places should be replaced by rate comparisons when deciding

which countries are “in the lead,” if and when we should lockdown, what to do when the lockdown is over, and whether waiting for herd immunity is an option.

Counts can be useful to show when incidence is starting to recede as public health measures take effect in a particular population. The shape of the portrayed trends in case counts enables us to see that the United Kingdom, France, Italy, and Spain are currently on similar trajectories, whereas Korea and other Asian countries have been “flattening the curve.” Still, graphs of case numbers cannot be used to say that one country is ahead of another. For example, the headline “The United States Is Now the Epicenter” (i.e., of cases) does not reflect that, the US population is over five times that of Italy and is spread over a much larger area, with large differences between various American states.

SELECTION AND MISCLASSIFICATION

Reporting rates solves one methodological problem, by adjusting the count to the size of the population. However, the selection of those tested is critical

for accurate estimation. In a given population (P) at a particular time, some will have been infected (I), some will have tested positive, and some will have symptoms associated with COVID-19; then, among the infected, some will have died (D). If testing for SARS-CoV-2 is done randomly and the test has very high sensitivity and specificity, one can obtain reasonably valid estimates of the infected (I), the population attack rate ($AR = I/P$), and the infection fatality rate ($IFR = D/I$). But the current situation does not resemble this ideal condition.

Accurate estimation of the AR and the IFR depends on the testing strategy, the prevalence of infection, and the test sensitivity and specificity. Differences between countries or over time may merely reflect differences in selection for testing and in test performance, including the following:

1. With survey exceptions, those tested for SARS-CoV-2 have been people with symptoms. In countries overwhelmed by the epidemic, most of those tested have had severe symptoms; in other countries, people with less severe symptoms have also been tested. Unfortunately, the resulting test-positive rate and diagnosed-case fatality rate (COVID-19 deaths among those testing positive) are respectively inflated estimates of the population attack rate (I/P) and infection fatality rate (D/I). More generally, by testing only people with symptoms, the resulting estimated attack rate will be a serious underestimate if there are many who are infected but are nonsymptomatic. Similar problems will afflict the reported case fatality rate (i.e., for a certain period of time the rate of death among those testing positive). In Italy, this has been greater than 12%; this may largely reflect that severely symptomatic people (who tend to be older) were most often targeted for testing. Germany has employed more widespread testing and had an apparent case fatality rate of around 3%. Although some of

ABOUT THE AUTHORS

Neil Pearce is with the London School of Hygiene and Tropical Medicine (LSHTM), London, UK. Jan P. Vandenbroucke is with the Leiden University Medical Center, Leiden, Netherlands, Aarhus University, Aarhus, Denmark, and the LSHTM. Tyler J. VanderWeele is with the Harvard T. H. Chan School of Public Health, Boston, MA. Sander Greenland is with the Departments of Epidemiology and Statistics, University of California, Los Angeles.

Correspondence should be sent to Professor Neil Pearce, Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London WC1E 7HT, United Kingdom (e-mail: neil.pearce@lshtm.ac.uk). Reprints can be ordered at <http://www.ajph.org> by clicking the “Reprints” link.

This editorial was accepted April 7, 2020.
doi: 10.2105/AJPH.2020.305708

this difference might be attributable to differences in case management, the huge difference shows that death counts and fatality rates among symptomatic cases are grossly inadequate for determining infection fatality.

2. Test performance (sensitivity, specificity, predictive values) of the tests in the field are as yet largely unknown and will vary across place and time. For example, even with the near-perfect specificity of a molecular or antigen test, because of technical errors there may be occasional false positives and considerable false negatives (because of difficulties in getting a good swab, differences in virus shedding over the disease course, specimen cross-contamination, etc.). Consequently, the positive predictive value (PPV, the chance that a test-positive individual is actually infected) will be poor in very low-prevalence situations (e.g., when an epidemic is beginning or when it is phasing out). Thus if the test had 70% sensitivity and 99% specificity but the prevalence of infection among the tested was only 3%, then the PPV of the test would be only 68%. This would make about one third of the reported cases false positives, increasing the estimated attack rate and reducing the apparent case fatality rate.
3. In estimating infection fatality rates, instead of knowing the true number of infected who have died (D), we generally have to make do with the number who have tested positive and have died (D*), which can lead to overestimation or underestimation. This depends on the testing strategy. Some may falsely test positive and die

and thus be included among D*. Others may die with infection but falsely test negative or have not been tested at all and thus not be included in D*. Moreover, deaths among some who recently tested positive may be delayed for weeks after the calculation of rates. Complicating interpretation further is that it is possible to die with, but not of, the infection and that there is no sharp boundary between the two (e.g., how can we know whether a death from coronary insufficiency during COVID-19 pneumonitis would have occurred without the infection?) Thus the total death count of infected persons (D), even if known, will overstate deaths from the infection.

THE CASE OF THE UNITED KINGDOM

The difficulties of drawing policy decisions from inadequate data are illustrated by the situation in the United Kingdom, which initially took a “wait for herd immunity” approach but is now taking measures similar to those of other large European countries.¹ The change came because a report in March from Imperial College² used new data from Italy showing that the proportion of hospitalized patients needing intensive care was similar to that reported from China in January. If these estimates applied to the United Kingdom, the National Health Service would be overwhelmed and there would be about 250 000 deaths—similar to the UK death toll of the 1918/19 flu. The Imperial College estimates were much lower when assuming the United Kingdom followed the isolation approach of

other countries. However, a University of Oxford report explored other scenarios, one of which indicated there may already be substantial herd immunity and therefore the death toll was likely to be relatively low.³

These two sets of reports used essentially the same data but fed different assumptions into the models and thus came to starkly different conclusions. The key difference was in the assumptions about the proportion of infected individuals who have been undiagnosed, because they were either asymptomatic or otherwise untested. Most have estimated the infection fatality rate to be about 1% (as in the Imperial College analyses) based on deaths among test-positives,⁴ whereas if there is a large pool of undiagnosed nonsymptomatic infections, then the true rate (D/I) would be much lower.

SURVEILLANCE NEEDS

We need testing strategies to estimate population numbers and rates of infection and death—and not just in people with symptoms or people testing positive. We also need accurate immunologic tests to see who has been infected and may have developed immunity. Ideally, surveillance would involve the following:

1. Repeated representative sampling of diverse parts of the population is necessary. This can be approximated if countries follow the World Health Organization’s recommendations with the caveat that the implications of results from test surveys depends on the stage of the epidemic.⁵ Epidemic control requires detecting even minor symptoms and testing the immediate contacts of those who test

positive. Test-negatives also provide important information (e.g., about protective factors and false negatives). More representative testing would enable reasonable estimates of current reproduction numbers as well as attack rates; ideally this would be a continuous process throughout the epidemic.

2. Validation data for each test brand, laboratory, and country are needed because the tests cannot be identical in performance across sources and field administrators.
3. Estimation of infection fatality rates from those who test positively for infection in representative samples of the population followed for a sufficient length of time. These estimates need to account for test error rates and should ideally clarify whether and how deaths were classified as being from (as opposed to with) infection.

We will eventually get through this pandemic, but in the process the world will change. One positive change should be the recognition that we need good surveillance systems permanently in place for both infectious and chronic disease. Because emerging conditions and some established conditions cannot be identified from routine health system data, regular population health and medical surveys are vital. Thus, surveillance and descriptive epidemiology remain vital foundations for sound health science and policy. **AJPH**

Neil Pearce, PhD

Jan P. Vandenbroucke, MD, PhD

Tyler J. VanderWeele, PhD

Sander Greenland, DrPH, MS

CONTRIBUTORS

The authors contributed equally to this editorial.

ACKNOWLEDGMENTS

We thank Elizabeth Brickley, Alfredo Morabia, and Christina Vandembroucke-Grauls for their comments on the draft of the editorial.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Hunter DJ. COVID-19 and the stiff upper lip—the pandemic response in the United Kingdom. *N Engl J Med*. 2020; Epub ahead of print.
2. Ferguson NM; Imperial College of Science, Technology and Medicine. *Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID19 Mortality and Healthcare Demand*. London: Imperial College COVID-19 Response Team; 2020.
3. Lourenco J, Paton R, Ghafari M, et al. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. 2020. Available at: <https://www.medrxiv.org/content/10.1101/2020.03.24.20042291v1.article-info>. Accessed April 8, 2020.
4. Kucharski A. Can we trust the Oxford study on COVID-19 infections? *The Guardian*. March 26, 2020. Available at: <https://www.theguardian.com/commentisfree/2020/mar/26/virus-infection-data-coronavirus-modelling>. Accessed April 8, 2020.
5. Lipsitch M. Far more people in the US have the coronavirus than you think. *Washington Post*. March 23, 2020. Available at: https://www.theeagle.com/opinion/columnists/far-more-people-in-the-u-s-have-the-coronavirus/article_767acb63-0299-50ec-a142-06d9ba32042a.html. Accessed April 8, 2020.
6. Caplin B, Jakobsson K, Glaser J, et al., International collaboration for the epidemiology of eGFR in low and middle income populations—rationale and core protocol for the Disadvantaged Populations eGFR Epidemiology Study (DEGREE). *BMC Nephrol*. 2017;18(1):1.
7. Ellwood P, Asher MI, Billo NE, et al. The global asthma network rationale and methods for phase I global surveillance: prevalence, severity, management and risk factors. *Eur Respir J*. 2017;49(1):1601605.