Accurate Statistics on COVID-19 are essential for Policy Guidance and Decisions

Neil Pearce[1], Jan Vandenbroucke[1,2,3],
Tyler J. VanderWeele[4], Sander Greenland[5]

(1) London School of Hygiene and Tropical Medicine, London, UK

(2) Leiden University Medical Center, The Netherlands

(3) Aarhus University Denmark

(4) Harvard T.H. Chan School of Public Health

(5) Department of Epidemiology and Department of Statistics, University of California, Los Angeles, CA

6 April 2020

N=1282

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. In this piece we review key epidemiologic 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 time period*: the number of identified Covid-19 cases (numerator) divided by the population size (denominator). See Table 1 for abbreviations and definitions. The size and source of the denominator is important; for example, a headline proclaiming "Italy surpasses China" which is based on total case counts is misleading because, compared to Italy, China has about 24 times the population; 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. As we write, the shape of the portrayed trends in case counts enables us to see how the UK, France, Italy, Spain are currently on similar trajectories, whereas Korea and other Asian countries are '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" of cases does not reflect that the U.S. attack rate was about a fifth that of Italy's and 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, a subgroup (I) will have been infected, some of them will have tested positive (T), some may have symptoms (S), and some of these will have died (D). Many non-infected may have similar symptoms to Covid-19 cases, and the

infected I will exceed the symptomatic S, possibly by a large amount. If testing for Covid-19 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 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

(i) Only testing people with symptoms: unfortunately the test-positive rate among the symptomatic (TSR=T/S) and the diagnosed-case fatality rate (CFR=D/T) are inflated estimates of the population attack rate (I/P) and infection fatality rate (D/I). Consider that the reported case fatality rate (CFR) in Italy has been over 10%, but this may largely reflect that only symptomatic people (who tend to be older) are being tested (D/T, not D/I). Germany employs more widespread testing and has had an apparent case fatality rate below 2%. While some of this difference is due to case management, the huge difference shows that death counts and fatality rates among symptomatic cases are grossly inadequate for determining infection fatality.

(ii) 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 near perfect specificity of a molecular test, due to technical errors there may be occasional false positives and considerable false negatives (due to 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 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 were 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 case-fatality rate.

The case of the United Kingdom

The difficulties of drawing policy decisions from inadequate data are illustrated by the United Kingdom, which initially took a 'wait for herd immunity' approach, but is now taking

measures similar to other large European countries[1]. The change came because a report in March from Imperial College [2] 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 UK, the NHS would be overwhelmed and there would be about 250,000 deaths - similar to the UK death toll of the 1918/9 Spanish Flu. The Imperial College estimates were much lower when assuming the UK 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 infections which have been undiagnosed, either because they were 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 (CFR=D/T) [4], whereas if there is a large pool of undiagnosed non-symptomatic infections, then the true rate (D/I) would be much lower.

The need for surveillance

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:

1. Repeated representative sampling of diverse parts of the population. This can be approximated if countries follow WHO'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 of the immediate contacts of positives. 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 population prevalence; ideally this would be a continuous process throughout the epidemic.
2. Validation data for each test brand, laboratory, and country since the tests cannot be identical in performance across sources and field administrators.

We will eventually get through this, but in the process the world will have changed. One positive change should be the recognition that we need good surveillance systems permanently in place for both infectious and chronic disease. Some conditions cannot be identified from routine health system data, and regular population surveys of both infectious and chronic diseases are needed. Surveillance and descriptive epidemiology remain essential foundations for sound health science and policy.

Neil Pearce, PhD, Jan Vandenbroucke, MD PhD, Tyler VanderWeele, PhD, Sander Greenland, MS DrPH

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

Correspondence should be sent to Professor Neil Pearce, Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London WC1E 7HT

Accepted on: April 7th, 2020

Contributors

All authors contributed equally to this manuscript.

Acknowledgements

We thank Elizabeth Brickley, Alfredo Morabia, and Christina Vandenbroucke-Grauls for their comments on the draft manuscript.

Conflicts of interest

None declared.

References

1. Hunter, D., *Covid-19 and the Stiff Upper Lip — The Pandemic Response in the United Kingdom.* New England Journal of Medicine, 2020: p. in press.
2. Ferguson, N.M., et al., *Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand*. 2020, Imperial College: London.
3. Lourenco, J., 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, University of Oxford: Oxford.
4. Kucharski, A., *Can we trust the Oxford study on Covid-19 infections?*, in *The Guardian*. 2020: London.
5. Lipsitch, M., *Far more people in the U.S. have the coronavirus than you think*. 2020, Waskington Post: Washington.
6. Caplin, B., 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 Nephrology, 2017. **18**: p. DOI 10.1186/s12882-016-0417-1.
7. Ellwood, P., et al., *The Global Asthma Network rationale and methods for Phase I global surveillance: prevalence, severity, management and risk factors.* European Respiratory Journal, 2017. **49**: p. 1601605.

Table 1: Epidemiological concepts and statistics used in this paper

Abbreviation	Term	Formula
P	*Population*	
I	*Infected*	
T	*Subgroup of testing positive*	
S	*Subgroup with symptoms*	
NS	*undiagnosed non-symptomatic infections*	
D	*Deaths*	
Se	*Sensitivity*	True Positive Tests/Diseased
Sp	*Specificity*	True Negative tests/Non-Diseased
AR	*Attack rate over a defined time period*	I/P
CFR	*Diagnosed-case fatality rate*	D/T
IFR	*Infection fatality rate*	D/I
PPV	*Positive predictive value*	I/Positive Tests