Linking Statistics With Testing Policy to Manage COVID-19 in the Community

Lee H. Hilborne, MD, MPH,[1,2,3] *Zachary Wagner, PhD,*[1,4] *Irineo Cabreros, PhD,*[1] *and Robert H. Brook, MD, ScD*[1,5]

From the [1]RAND Corporation, Santa Monica, CA; Departments of [2]Pathology and Laboratory Medicine and [5]Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA; [3]Quest Diagnostics, Secaucus, NJ; and [4]Pardee RAND Graduate School, Santa Monica, CA.

Key Words: COVID-19; SARS-CoV-2; Laboratory; Surveillance; Public health; Coronavirus

Am J Clin Pathol 2020;XX:0–0

DOI: 10.1093/AJCP/AQAA099

ABSTRACT

Objectives: *To determine the public health surveillance severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) testing volume needed, both for acute infection and seroprevalence.*

Methods: *Required testing volumes were developed using standard statistical methods based on test analytical performance, disease prevalence, desired precision, and population size.*

Results: *Widespread testing for individual health management cannot address surveillance needs. The number of people who must be sampled for public health surveillance and decision making, although not trivial, is potentially in the thousands for any given population or subpopulation, not millions.*

Conclusions: *While the contributions of diagnostic testing for SARS-CoV-2 have received considerable attention, concerns abound regarding the availability of sufficient testing capacity to meet demand. Different testing goals require different numbers of tests and different testing strategies; testing strategies for national or local disease surveillance, including monitoring of prevalence, receive less attention. Our clinical laboratory and diagnostic infrastructure are capable of incorporating required volumes for many local, regional, and national public health surveillance studies into their current and projected testing capacity. However, testing for surveillance requires careful design and randomization to provide meaningful insights.*

Key Points

- Community testing for coronavirus disease 2019 (COVID-19) disease surveillance requires a different testing strategy from that used for individual health care purposes.
- The number of people who must be sampled for public health surveillance and decision making, although not trivial, is potentially in the thousands for any given population or subpopulation, not millions.
- Current and projected laboratory capacity is sufficient for surveillance testing but requires careful design and randomization to provide meaningful insights.

As the coronavirus disease 2019 (COVID-19) pandemic enters its sixth month in the United States, much of the public health and prevention discourse focuses on the need for increased diagnostic testing. The purposes of testing, however, receive less attention, leading to confusion about the testing capacity required. Different testing goals require different numbers of tests and different testing strategies.

There are many distinct roles that testing aims to address, including: (1) health care management for individual patients, (2) identifying exposed individuals through contact tracing to inform quarantine, and (3) disease surveillance.

Sufficient tests are generally available to meet the personal health needs of individual patients, and most of the nearly 16 million (as of May 28, 2020) molecular tests administered to this point have been for this purpose. Health policy experts argue that effective contact tracing requires drastic increases in testing capacity.[1] For instance, one highly publicized plan developed by a consortium of experts recommends 20 million tests per day.[2]

© American Society for Clinical Pathology, 2020. All rights reserved.
For permissions, please e-mail: journals.permissions@oup.com

In contrast, testing strategies for national or local disease surveillance, including monitoring of prevalence, receive less attention. As a result, there is uncertainty surrounding basic questions: What proportion of the population in a given area is currently infected? What proportion of the population already has been infected? Is acute infection prevalence increasing as shelter at home guidance relaxes? Despite the widespread testing for individual patient care, testing to date cannot answer these questions. Without evidence of the infection dynamics in the population, policy makers are in the difficult position of making decisions without a clear picture of the true prevalence and mortality rate of the virus.

The number of tests required for disease surveillance is manageable but requires carefully designed random testing. This report seeks to provide guidance for public health officials, local governments, and large employers developing testing strategies to track disease prevalence in their respective communities. The central message is that accurate monitoring of disease prevalence can be achieved by testing a relatively small number (typically, thousands) of randomly sampled individuals.

The Current Testing Landscape

Currently there are 3 types of diagnostic tests available: molecular assays (polymerase chain reaction [PCR]), serology assays, and antigen tests. Each test has specific applications for personal health management and public health surveillance. This discussion focuses on the first 2, because the antigen test has only recently received emergency use authorization from the Food and Drug Administration and is not yet widely available.[3] Molecular and serology testing complement each other, and both are necessary to paint a complete picture of the current state of the pandemic.

The molecular test identifies people currently infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus responsible for COVID-19. Those who test positive, whether symptomatic or not, are presumed to be contagious and risk transmitting the virus; however, some patients who test positive later in the course of infection may have noninfectious viral remnants detected.[4] These tests can be used to monitor active infection in the population and assess how infection dynamics change in response to policy modification (eg, opening schools) **Table 1**. However, testing for public health surveillance must be conducted in a careful and targeted way. Testing conducted for personal health or prevention cannot be used for public health surveillance. For example, in the initial phases of the pandemic, only the sickest patients were tested with the molecular test (due to limited availability). Using these test results to measure the number of positive cases drastically underestimates the extent of infection because many patients are asymptomatic. Furthermore, morbidity and mortality rates among tested individuals overestimate the true morbidity and mortality rates of COVID-19 because the denominator, rather than reflecting the total number of individuals infected, only incorporates those who were tested.

Table 1
Molecular (Polymerase Chain Reaction) Testing Applications

Personal health management and prevention
Evaluating/triaging patients with acute flu-like illnesses
Informing positive patients of necessary measures to reduce transmission during the infective period
Screening asymptomatic patients before medical procedures
Contact tracing: ensuring individuals with known exposure do not harbor the virus either asymptomatically or presymptomatically
Public health surveillance
Measuring acute infection rates (the number of people infected at one point in time)
Evaluating asymptomatic patients to understand the prevalence of acute, undetected infection
Monitoring the rise and fall of acute infections in response to easing or strengthening public health measures (eg, physical distancing)

Serum-based serology tests identify people who were previously infected or who are recovering. These tests demonstrate an immune response by most patients as they recover from infection. While it is unknown how long antibodies to SARS-CoV-2 will be detectable, based on data from SARS-CoV and MERS-CoV it is likely to be at least a few years.[5,6] Serology testing can identify how many people have been infected by SARS-CoV-2 and provide greater insight on exactly how harmful and deadly the virus actually is **Table 2**. This will help refine estimates of the true morbidity and mortality rates for different populations, evaluate the prevalence of an immune response following infection, and identify populations or locations with greatest spread. It will also help identify particularly resilient populations. If, for example, a large enough portion of people under 40 years of age are

Table 2
Serology Testing Applications

Personal health management
Determining resistance to reinfection due to immunity (not yet proven)
Identifying protected individuals who can venture back to work (not yet proven)
Public health surveillance
Identifying how many have been infected with the virus
Estimating true mortality and morbidity rates
Assessing how a community is progressing toward "herd immunity"

DOI: 10.1093/ajcp/aqaa099
© American Society for Clinical Pathology

asymptomatic and the mortality rate is sufficiently low, this population could play a larger role in the initial stages of reopening the economy, provided they are not personally or living with individuals at increased risk.

Although not fully proven, experts expect that those with antibodies have immunity, at least in the near term.[6,7] Thus, serology testing can also be used to monitor immunity in the population. At present, the primary application for serology testing is to monitor prevalence in communities, not to inform personal care or behavior of individual patients.[5]

To date, several studies have been conducted for the purpose of disease surveillance in communities. Studies conducted in large metropolitan areas found seroprevalence levels of 4.6% in Los Angeles County,[8] 2.8% in Santa Clara County,[9] and 21% in New York City.[10] If accurate, these studies have large public policy implications. First, they suggest that the total number of infected individuals could be far higher and the death rate far lower than official records. Second, despite the devastating toll the virus already has had on society, these findings suggest that most people remain at risk. Experts expect that herd immunity, the level of immunity required in the community for disease spread to be unlikely, is not achieved until at least 55% to 60% are immune.[11]

Moving From Statistics to Testing Policy

Policy makers and researchers need to understand the dynamics of both acute infection and seroprevalence in communities to inform testing policy decisions. But how many people must actually be tested in the US, in a state, in a city, in a large factory or plant, or in an at-risk group to estimate prevalence with sufficient accuracy to inform testing policy? It will never be feasible to test everybody, nor is it necessary. Careful random testing allows precise estimation of population values from relatively few total tests.

More testing alone is insufficient to estimate the true prevalence in the population; it is essential to draw random samples from the community to ensure the estimates represent the population. As discussed above, testing only the sick with molecular tests will not yield an accurate estimate of the rate of active infection in the population. If serology testing is not offered randomly, those who suspect they were previously infected will be more likely to volunteer for testing, leading to seropositivity overestimates. The serology testing study in Santa Clara used social media to recruit test volunteers; and the New York study recruited volunteers from grocery stores. Recognizing that not all individuals use social media or visit grocery stores, these samples are not representative of the entire community from which they were drawn.

Recruiting a random sample from a population and conducting a testing campaign is no small task. From a logistical perspective, this likely requires setting up many testing sites community-wide to ensure convenient testing access. Furthermore, all tests must be collected, analyzed and reported over a short time (eg, 1 or 2 days for molecular tests, a week for serology) to be actionable and because the prevalence can change quickly.

Even if the number of tests is manageable from the perspective of testing capacity, there are still significant recruitment and logistic challenges. For example, to perform random testing in a city, implementers would need a sampling frame of the entire city population to draw from (eg, census records). Sampled individuals may not agree to be tested or may not follow through even if they agree, which can lead to nonresponse bias. Thorough follow-up, sufficient incentives, and careful nonresponse modeling should be in place to reduce nonresponse bias and ensure accuracy. In practice, this also means that more people must be recruited for testing than are needed for the desired sample size.

These logistics are challenging. Public health officials should seek guidance from sampling experts either within their organizations or externally to ensure the process is conducted appropriately. Local leadership, working with community liaisons, should consider seeking endorsements from respected leaders and determine incentives to encourage public participation.

Recognizing the complexities of whole population based screening, a more focused approach that examines groups of interest would be more feasible and potentially more informative. Consider these examples:

1. Transportation. A potential strategy could randomly select transportation patrons for surveillance screening. A study design may examine random travelers passing through a Transportation Security Administration checkpoint for the presence of acute infection. Similarly, a collection could be established at several subway, railway, or other transportation hubs. While these sampling strategies might not be representative of the entire population, they are representative of a segment of the population that could be spreading the virus.
2. Large residential complexes. A sampling strategy may elect to focus on individuals living in high rises, assisted living facilities, or dormitories because these domiciles have a higher risk compared to those living in less crowded settings. A study from the Veterans Affairs Greater Los Angeles Healthcare System instituted

© American Society for Clinical Pathology

DOI: 10.1093/ajcp/aqaa099

serial skilled nursing resident screening, at approximately weekly intervals.[12]

3. Restaurant patrons. A sampling of different types of dine-in restaurants could be selected to understand the prevalence of infection among those occupying tables or providing service to patrons and whether the prevalence changes with increased restaurant patronage.
4. Education facilities. Examining students, faculty, and staff across public and private schools and universities could be informative as in-person classes resume. Sampling from younger populations (eg, elementary schools) could provide greater insight into those harboring asymptomatic infection yet risk bringing the infection home to adults and seniors located in the same household or with whom they have close contact.

The exact number of people to test depends on the degree of accuracy required, which is in turn determined by the nature of the questions posed. Developing a viable testing strategy for disease surveillance therefore requires both statistical input and testing policy expertise. Consider these illustrative examples that seek to determine the testing capacity needed to address key public health questions:

1. What is the current seroprevalence and acute infection prevalence?
2. How do infection rates vary by age, sex, ethnicity, population density, and comorbidities?
3. How is seroprevalence changing over time?
4. How do active infection rates change after a change in policy?

The testing considerations necessary to address each of these questions follow.

What Is the Current Seroprevalence and Acute Infection Prevalence?

The number of tests required to monitor prevalence depends on several inputs. The first thing to consider when designing a community testing program is the precision required for estimates to be meaningful. When prevalence is higher, estimates with large uncertainty may still be informative. When prevalence is lower, higher precision may be required. For instance, the difference between a prevalence of 0.5% and 2.0% may have huge implications (4-fold difference) while difference between 20.5% and 22% may be less important. Prevalence of acute infections is much lower than seroprevalence because acute infection exists for only 2 weeks or less, whereas antibodies remain much longer. Thus, larger samples are generally needed for molecular than for serology testing to produce informative estimates.

Another important consideration is test accuracy, typically described by measures of sensitivity and specificity. Low specificity or sensitivity can increase the number of tests required to attain adequate precision.

Perhaps counterintuitively, the number of tests needed to estimate prevalence does not depend sensitively on the total size of the population, when the population is large. For instance, whether estimating prevalence in the entire US, the state of California, or Los Angeles County, the required sample size is virtually identical. However, for smaller communities (eg, a small municipality or a large employer), population size does impact the sample size required. **Figure 1** demonstrates how the required sample size changes as a function of population size when other parameters are fixed. As population size approaches 1 million, rather than continually increasing, required sample sizes actually plateau (in this example, the maximum required sample size for any population is 2,377). This observation is a general result of statistical sampling, and applies equally to any random sample to determine prevalence, not specifically to COVID-19.[13]

Table 3 and **Table 4** show the number of tests required to obtain estimates of prevalence within specified degrees of precision for the molecular test and serology test, respectively. Sensitivity and specificity for each of the 2 tests have considerable variation reported in the literature. The statistical model accommodates these differences, incorporating sensitivity and specificity as explicit variables in the calculation. In this discussion, for the molecular test, a specificity of 99.8% and sensitivity of 80% is assumed, which is reasonable for currently available PCR tests (although the sensitivity of collection continues to improve from the 65% originally reported).[14] For the serology test, a specificity of 99% and sensitivity of 95% is used, which is also typical.[5,15] Both tables show required sample sizes for a range of values of common interest and assume that the studied population is large (eg, a large metropolitan region). We have provided an online tool, derived from Supplement 1 (all supplementary material can be accessed at *American Journal of Clinical Pathology* online), to compute required sample sizes for any prevalence, specificity, sensitivity, precision, and population size.[13,16]

An important takeaway from Tables 3 and 4 is that the number of tests required under random sampling for any given population is far less than what is required for effective contact tracing (thousands as opposed to millions). While testing capabilities must be enhanced substantially to support contact tracing, the current and projected testing infrastructure can handle the additional

DOI: 10.1093/ajcp/aqaa099
© American Society for Clinical Pathology

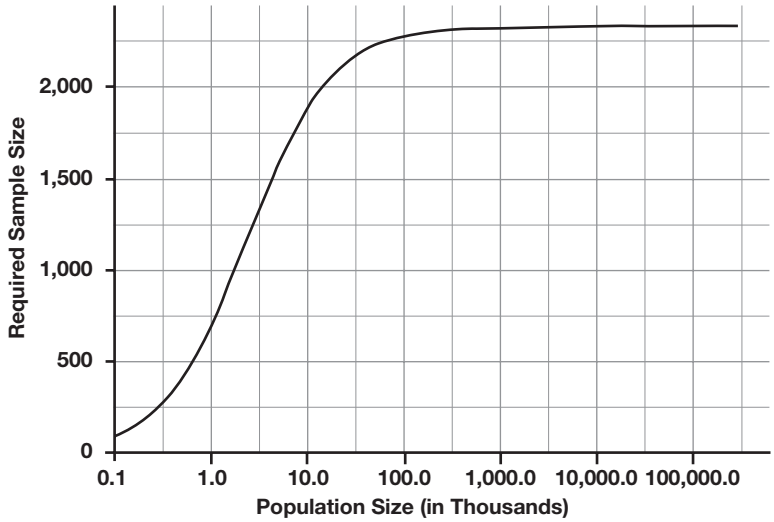

Figure 1 Required sample size as a function of population size (at sensitivity 95%, specificity 99%, true prevalence 5%, and precision 1% [ie, 4%-6%]).

Table 3
Sample Sizes Required for Molecular Testing

Prevalence	Precision (95% Confidence Interval)	No. of Random Sample Tests Required
0.2%	0.1%-0.3%	21,615
0.5%	0.3%-0.7%	8,980
1%	0.6%-1.4%	3,726
5%	4.0%-6.0%	2,422

Table 4
Sample Sizes Required for Serology Testing

Prevalence	Precision (95% Confidence Interval)	No. of Random Sample Tests Required
1%	0.6%-1.4%	5,170
5%	4.0%-6.0%	2,337
10%	8.0%-12.0%	1,013
20%	16%-24%	432
40%	35%-45%	413

testing required to estimate prevalence in multiple different settings simultaneously.

Since the required number of tests to determine prevalence depends on prevalence itself, an estimate of the expected prevalence in the population of interest can help guide the required level of precision.

Consider, for example, Los Angeles County, where 4.6% of the population tested positive for antibodies in a recent study based on a random sample.[8] Knowing that the acute phase of the illness is short, over the first 12 weeks of the pandemic approximately one-sixth of the total infected patients would have active infection at any one time (about 0.7%, or 70,000 people). This means that estimating active infection prevalence requires a very small margin of error for estimates to be informative (much less than 1 percent). For estimates to be within 0.2 percentage points of the true prevalence of acute infection, which would provide a prevalence range of 0.5% to 0.9%, would require testing approximately 8,728 people in the community.

How Do Infection Rates Vary by Age, Sex, Ethnicity, Population Density, and Comorbidities?

To understand disease prevalence within certain subgroups in addition to the population as a whole, the testing strategy will have to ensure that adequate precision is reached within each of these subpopulations. If the size of these subpopulations is large, then necessary sample sizes can be attained identically to the previous section (Tables 3 and 4). For instance, if the expected rate of seroprevalence is 5% and precision within 1% (ie, 4%-6%) is required for both men and women, then a sample of 2,337 men and 2,337 women is needed. If the size of the subpopulation is small relative to the size of the sample, the sample sizes reported in Tables 3 and 4 may be much larger than necessary. However, the online tool takes population size into account and can determine accurate sample size estimates for populations as small as 100 individuals.[13,16]

How Is Seroprevalence Changing Over Time?

As the COVID-19 pandemic progresses, community officials must monitor seroprevalence changes to better understand community experience with COVID-19, and in the near term reiterate that, assuming seroprevalence is protective, it is not yet sufficient to provide for herd immunity. However, as seroprevalence increases, viral transmissibility (R0) would

© American Society for Clinical Pathology

DOI: 10.1093/ajcp/aqaa099

decrease. This requires random sample serology testing at multiple points in time in the same communities. For example, departments of public health could conduct monthly or quarterly serology testing on a random sample to monitor seroprevalence locally. The Centers for Disease Control and Prevention has already proposed immunity surveillance nationally.[17] Using the online calculator,[16] for a national or large regional population, a sample size of 1,013 individuals would be needed to determine immunity assuming an expected 10% prevalence with a confidence interval between 8% and 12%. As seroprevalence increases, testing intervals can be extended. Continued tracking of immunity following the launching of a vaccination program will allow better assessment of community seroconversion and hopefully resultant herd immunity.

How Does Prevalence Change After a Change in Testing Policy?

As society begins to open, it is essential to understand the impact of easing different restrictions on prevalence. To accurately monitor fluctuations in active infections, random upper respiratory samples for molecular testing should be collected immediately before and within 3 to 5 days after the policy change. The population of interest will depend on the testing policy. To open restaurants, initial sampling should focus on restaurant workers. In anticipation of opening schools, students, teachers, and staff should be tested. While openings could lead to spread beyond restaurants and schools, identifying a spike in infections in these settings early on will allow policy makers to retreat from permissive policies before the spread becomes too severe (eg, increased hospital admissions). We have provided a second online tool to compute required sample sizes to detect any prevalence change given test specificity, sensitivity, and population size.[18]

For example, to assess changes in infection rates after public schools open in a large metropolitan area, the following steps should be taken:

1. Use student registers and staffing lists of all public schools to randomly select individuals from schools.
2. Test all selected individuals with molecular tests the day before schools open.
3. Repeat steps 1 and 2 a few days after schools open.
4. Compare active infection rates between the 2 testing waves.

Policy makers must carefully consider how much of an increase in infections among students, faculty, and staff would be an appropriate threshold to close schools down again; the required sample size is based on this value. For instance, if the active infection rate among 10,000 students, faculty, and staff was 0.5% before schools opened, and a 50% increase was the threshold policymakers established to close schools, then sampling approximately 2,200 students, faculty, and staff before and after opening is needed to have enough precision to detect this size change. Then, if the infection prevalence rose to 0.75%, the local school board would need to implement mitigation protocols.

Conclusion

Monitoring the evolution of the pandemic in our society is crucial. The above discussion highlights some very important principles of statistical sampling that are relevant to formulating policy. Perhaps most importantly, the number of individuals that must be tested is feasible and does not change when the population size is large. This observation highlights that the number of people who must be sampled for public health surveillance and decision making, although not trivial, is potentially in the thousands for any given population or subpopulation, not millions. Current clinical laboratory and diagnostic infrastructure is generally capable of incorporating required volumes for many local or regional studies into their current and projected testing capacity. However, testing for surveillance requires careful design and randomization to provide meaningful insights. Controlling the COVID-19 pandemic requires leveraging the tight cooperation between diagnostic laboratories and the public health infrastructure developed over the decades to assure that these insights effectively impact policy.[19] There are few investments that are as important as carefully designed testing strategies.

Corresponding author: Lee H. Hilborne, MD, MPH; lhilborne@mednet.ucla.edu.

References

1. Babiker A, Myers CW, Hill CE, et al. SARS-CoV-2 testing: trials and tribulations. *Am J Clin Pathol.* 2020;153:706-708.
2. Allen D, Block S, Cohen J, et al. Roadmap to pandemic resilience: massive scale testing, tracing, and supported isolation (TTSI) as the path to pandemic resilience for a free society. https://ethics.harvard.edu/files/center-for-ethics/files/roadmaptopandemicresilience_updated_4.20.20_1.pdf. Accessed May 19, 2020.
3. Venter M, Richter K. Towards effective diagnostic assays for COVID-19: a review. *J Clin Pathol.* 2020. doi: 10.1136/jclinpath-2020-206685.
4. Kim BG. Tests in recovered patients found false positives, not reinfections, experts say. *Korea Herald*, April 29, 2020. http://www.koreaherald.com/view.php?ud=20200429000724. Accessed May 29, 2020.

DOI: 10.1093/ajcp/aqaa099
© American Society for Clinical Pathology

5. Torres R, Rinder HM. Double-edged spike: are SARS-CoV-2 serologic tests safe right now? *Am J Clin Pathol.* 2020;153:709-711.
6. Payne DC, Iblan I, Rha B, et al. Persistence of antibodies against Middle East respiratory syndrome Coronavirus. *Emerg Infect Dis.* 2016;22:1824-1826.
7. Theel ES, Slev P, Wheeler S, et al. The role of antibody testing for SARS-CoV-2: is there one? *J Clin Microbiol.* 2020. doi: 10.1128/JCM.00797-20.
8. Sood N, Simon P, Ebner P, et al. Seroprevalence of SARS-CoV-2–specific antibodies among adults in Los Angeles County, California, on April 10–11, 2020. *JAMA.* 2020. doi: 10.1001/jama.2020.8279.
9. Bendavid E, Mulaney B, Sood N, et al. COVID-19 antibody seroprevalence in Santa Clara County, California. *medRxiv.* 2020. https://doi.org/10.1101/2020.04.14.20062463.
10. Goodman DJ, Rothfield M. 1 in 5 New Yorkers may have had Covid-19, antibody tests suggest. *New York Times.* April 23, 2020. https://www.nytimes.com/2020/04/23/nyregion/coronavirus-antibodies-test-ny.html. Accessed May 15, 2020.
11. Angulo FJ, Finelli L, Swerdlow DL. Reopening society and the need for real-time assessment of COVID-19 at the community level. *JAMA.* 2020. https://doi.org/10.1001/jama.2020.7872.
12. Dora AV, Winnett A, Kitt LP, et al. Universal and serial laboratory testing for SARS-CoV-2 at a long-term care skilled nursing facility for veterans — Los Angeles, California, 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69:651-655.
13. Cochran WG. *Sampling Techniques.* 3rd ed. New York, NY: Wiley; 1977.
14. Wikramaratna P, Paton RS, Ghafari M, et al. Estimating false-negative detection rate of SARS-CoV-2 by RT-PCR. *medRxiv.* 2020. https://doi.org/10.1101/2020.04.05.20053355.
15. United States Food and Drug Administration. EUA authorized serology test performance. https://www.fda.gov/medical-devices/emergency-situations-medical-devices/eua-authorized-serology-test-performance. Accessed May 19, 2020.
16. https://covid-testing-calculators.shinyapps.io/calculator/.
17. Centers for Disease Control and Prevention. COVID-19 serology surveillance strategy. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/serology-surveillance/index.html. Accessed May 20, 2020.
18. https://covid-testing-calculators.shinyapps.io/calculator2/.
19. Villanueva J, Schweitzer B, Odle M, et al. Detecting emerging infectious diseases: an overview of the laboratory response network for biological threats. *Public Health Rep.* 2019;134:16S-21S.

© American Society for Clinical Pathology

DOI: 10.1093/ajcp/aqaa099