SARS-CoV-2 genetic surveillance recommendations for state, tribal, local, or territorial public health agencies

I. Purpose of this document:

The purpose of this document is to **provide recommendations for the minimum number of SARS-CoV-2 positive samples that should be sequenced to enable jurisdictions to detect SARS-CoV-2 variant lineages for situational awareness and to support national virologic surveillance**. The number of SARS-CoV-2 positive samples that should be sequenced depends on the purpose. This document lays out recommendations for jurisdictional situational awareness as to SARS-CoV-2 variant detection, proportional changes, the susceptibility of circulating viruses to therapeutics and inform potential need for vaccine antigen updates.

II. Background:

National SARS-CoV-2 genetic surveillance is used to identify and track variant lineages, identify mutations that may affect detection by diagnostic tests, determine if changes to the vaccine antigen are warranted, and provide situational awareness as to the susceptibility of the viruses to therapeutics (e.g., monoclonal antibody therapies). This large surveillance effort hinges on genomic sequencing approaches for SARS-CoV-2, but genetic changes in various variant lineages present many challenges to obtaining high quality data capable of producing a coding complete genome challenges (e.g., mismatches in primers resulting in overlapping amplicon failure(s)) or coding complete spike (1). Others are targeting the spike, which plays a major role in entry, pathogenesis, species specificity, immune evasion, sensitivity to some therapeutics, and is primary antigen in most vaccines. The approach(es) used depend upon the goals of the jurisdiction and support national SARS-CoV-2 surveillance efforts. By conducting local SARS-CoV-2 genetic surveillance, your jurisdiction will have timely local data on identification of emerging variants and changes in variant lineage proportions to guide public health action. Additionally, this will contribute to national surveillance efforts to identify and characterize emerging Variants of Interest(VOI), Variants of Concern (VOC), and Variants of High Consequence (VOHC) (https://www.cdc.gov/coronavirus/2019-ncov/variants/vriant-info.html).

III. Recommended criteria for selection of samples to be sequenced.

Selection of samples for sequencing can be challenging and we recommend guidelines similar to those used for the National SARS-CoV-2 Strain Surveillance (NS3) program (https://www.aphl.org/programs/preparedness/Crisis-Management/Documents/NS3-Submission-Guidance.pdf).

a. The quality of the specimen and/or RNA directly affects the likelihood of sequencing success. Ideally, specimens should have an RT-PCR Ct value of 28 or lower (this may vary depending on specific assay in use). If Ct values are not available, specimens that are positive for SARS-CoV-2 may be sequenced (avoid samples that are weak positives).

b. The time from specimen collection to sequence characterization has a large effect on your ability to quickly detect and track emerging variants. Prioritize specimens that have been collected within the last 7 days whenever possible. If the number of specimens collected within the last 7 days is insufficient to meet your jurisdiction's goals, sequence the most recent specimens possible (i.e., collected within the last 14 days).

c. Ideally, specimens should represent geographic, demographic (e.g., age), and clinical (e.g., disease severity or outcome) diversity from across the jurisdiction. The specimens should

be de-identified. We understand jurisdictions are using an array of approaches for surveillance, and many approached may be used to reduce bias. Some alternative approaches are listed below.

- Random selection of specimens collected within the last 7 days. For example, the desired number of positive samples may be selected by using random number tables or using random numbers generated by a computer program (e.g. Excel, =RAND()).

- Systematic/consecutive sampling. Select samples at regular intervals, for example, choose one of every third, fifth, or tenth sample to be sequenced, until the sampling target is achieved. This could be done each day to avoid groups of samples from the same location/submitter.

- Select samples in bulk to achieve your weekly goals. For example, select the first 60 positive samples arriving each day of the work week. Thought should be given as to how samples arrive. For example, if shipments of large numbers of specimens come from the same locality or facility each morning this method should be avoided.

d. Specimen metadata should be reviewed periodically to ensure that sampled specimens are generally representative of infections in the population. Deviations from randomness will impact the confidence of the observed proportions.

IV.   How many samples should be sequenced locally?

While the level of SARS-CoV-2 genetic surveillance conducted is dependent on the specific goals for your location, the strategy we suggest for detection of emerging variants, to understand the proportion of resistance/sensitivity of the variants to therapeutics and to optimally contribute to national surveillance is to target detection of variants circulating at 1% (1 variant/100 positive samples) or lower proportion (with 95% confidence). To achieve this, each jurisdiction should conduct SARS-CoV-2 sequencing of **300 SARS-CoV-2 positive original clinical specimens** each week (collected in the previous 7-day interval) and to maintain this level of sequencing consistently (i.e., every week). Recently identified specimens meeting surveillance criteria (defined in "III. Recommendation Criteria" above) enter the sequencing pipeline in a timely fashion so that the results are relevant and current for immediate public health decisions and action. For some jurisdictions the level of genetic

Figure 1. Detection threshold chart from https://covid-19.tacc.utexas.edu/dashboards/variants/.

surveillance may be too low for your desired detection limits or may be impractical. A variant detection calculator can be found at https://covid-19.tacc.utexas.edu/dashboards/variants/ that can be used to determine the number of specimens that should be sequenced in each interval to achieve the goals for

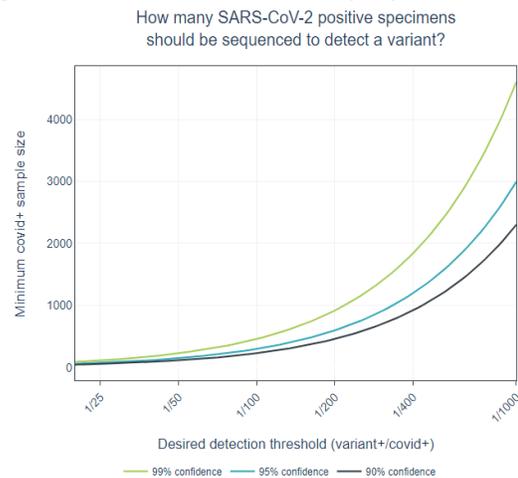## SARS-CoV-2 genetic surveillance recommendations for state, tribal, local, or territorial public health agencies

your jurisdiction per period assuming random sampling of positive specimens. Consistency and timeliness are very important for genetic surveillance, and there is a tradeoff between the total number sequenced and the period from swab collection to sequence publication. Rapid turnaround time between collection and sequence publication is very important for public health action.

Finally, jurisdictions with large populations (e.g., California, Texas, Florida, New York) should strongly consider detection of variants circulating at lower proportions such as 0.5% to 0.1% and should therefore target 598 to 2,995 positive specimens/week or use the calculator to determine the optimal number of specimens for jurisdictional detection and proportion tracking goals.

To meet the surveillance target numbers proposed in these recommendations, many jurisdictions may want to partner with groups (e.g., clinical diagnostic laboratories, pharmacies, or sentinel physician networks) that are performing diagnostic testing and reporting for SARS-CoV-2 so that positives are regularly submitted.

V.   What to do when there are not enough positive specimens available?

Continue to conduct SARS-CoV-2 surveillance during times of low prevalence, similar to influenza virologic surveillance. If there are not enough specimens available to meet the minimum recommendations, sequence as many as possible and do this regularly (e.g., preferably every week). In times of low prevalence, some flexibility on the criteria for sample collection (described in section III) is acceptable, but multiple samples from the same highly connected outbreak (e.g., congregate settings) should be avoided.

VI.   Other potential local public health goals that can be supported by genetic sequencing

SARS-CoV-2 genetic sequencing can be applied to many other factors related to the COVID-19 pandemic, including specific jurisdiction goals. Some of these are similar to the approach described herein (2, 3) and others are beyond the scope of this document but examples include: distinguishing ongoing transmission from introduction in a closed setting, identifying specific transmission routes, effect on diagnostics, outbreak investigations, understanding community transmission, or targeted epidemiologic studies to better understand variant effect on disease severity or vaccine breakthrough. For these examples, increased quantity, quality (coverage, depth), and/or appropriate representativeness may be needed. There are many resources and publications available to help with a variety of studies or improve upon these simplified guidelines(2-4), COVID-19 Genomic Epidemiology Toolkit | Advanced Molecular Detection (AMD) | CDC, Sequencing of SARS-CoV-2: first update (europa.eu), and https://apps.who.int/gb/COVID-19/pdf_files/2021/25_11/Item3.pdf.

VII.   How jurisdictional SARS-CoV-2 genetic sequencing supports our national surveillance system?

The CDC National SARS-CoV-2 Surveillance System (NS3) minimally aims to identify lineages/variants circulating at or below 0.1% (1 variant in 1000 samples) with 99% confidence. This information is used to assess the risk that variants pose and to develop or refine national and/or regional mitigation strategies and policies. CDC currently exceeds these detection thresholds using the NS3, and sequencing by commercial diagnostic labs across the United States that are contracted by CDC (i.e., https://www.cdc.gov/coronavirus/2019-ncov/variants/cdc-role-surveillance.html). We are also aided by

SARS-CoV-2 genetic surveillance recommendations for state, tribal, local, or territorial public health agencies

partners tagging their data as baseline (https://www.aphl.org/programs/preparedness/Crisis-Management/Documents/Technical-Assistance-for-Categorizing-Baseline-Surveillance-Update-Oct2021.pdf), particularly state and local public health labs, academic partners, and the SARS-CoV-2 Sequencing for Public Health Emergency Response, Epidemiology and Surveillance (SPHERES). Sequencing at the jurisdictional level will significantly improve regional, state, and local recommendations for use of specific therapeutics. Additionally, we anticipate that sequencing at the jurisdictional level will reduce the time frame from swab collection to sequence availability, which is very important to ensure that results are relevant for public health decisions and action.

VIII.   References

1.   Ladner JT, Beitzel B, Chain PS, Davenport MG, Donaldson EF, Frieman M, Kugelman JR, Kuhn JH, O'Rear J, Sabeti PC, Wentworth DE, Wiley MR, Yu GY, Threat Characterization C, Sozhamannan S, Bradburne C, Palacios G. 2014. Standards for sequencing viral genomes in the era of high-throughput sequencing. mBio 5:e01360-14.
2.   Wohl S, Lee EC, DiPrete BL, Lessler J. 2022. Sample Size Calculations for Variant Surveillance in the Presence of Biological and Systematic Biases. medRxiv:2021.12. 30.21268453.
3.   Anonymous. 2014. Influenza Virologic Surveillance Right Size Roadmap. 1-118.
4.   Hicks AL, Kissler SM, Lipsitch M, Grad YH. 2019. Surveillance to maintain the sensitivity of genotype-based antibiotic resistance diagnostics. PLoS Biol 17:e3000547.