

Modeling SARS-CoV-2 positivity using laboratory data: timing is everything

Running title: Laboratory data and pandemic modeling

Thor Larson¹, Karissa Culbreath², Dennis Chavez³, Richard Larson⁴, Michael Crossey², David G. Grenache²

¹Harvard College, Cambridge, MA

²TriCore Reference Laboratories, Albuquerque, NM

³Rhodes Group, Albuquerque, NM

⁴University of New Mexico Health Sciences Center, Albuquerque, NM

Corresponding author: David G. Grenache, TriCore Reference Laboratories, 1001 Woodward Place NE, Albuquerque, NM; 505-938-8647; david.grenache@tricore.org

As the pandemic caused by the novel coronavirus SARS-CoV-2 spreads across the globe, there is intense scrutiny placed on the clinical laboratory tests used to identify infected individuals. Apart from the analytical and clinical performance characteristics of these real-time reverse transcription polymerase chain reaction (rtRT-PCR) tests for the qualitative detection of the viral RNA, much attention is given to the number of positive cases reported by public health authorities each day. As it pertains to the current Lab 2.0 movement (1,2), there has never been a more clear example of the value of the laboratory in not only diagnostics and treatment of COVID-19, but in the real-time policy decisions that are made daily.

Across the United States, rtRT-PCR testing for SARS-CoV-2 was initially performed at public health laboratories. Such facilities have limited testing capacities which was quickly realized to be an impediment to meet the high demand for testing. Independent and hospital clinical laboratories quickly mobilized to alleviate the testing burdens that were accumulating.

TriCore Reference Laboratories is a regional clinical laboratory in the state of New Mexico (NM) that began rtRT-PCR for SARS-CoV-2 testing on 03/12/2020 using a platform that permitted testing approximately 300 samples per day. The number of samples submitted for analysis each day exceeded testing capacity, resulting in a backlog of specimens that was only cleared after high-throughput testing platforms were put into production.

All test results were reported to the NM Department of Health which updated the number of positive results on their website daily (3). However, the absolute number of positive cases reported by any institution on a given date does not represent the actual number of new cases on

that day. Due to backlogs, there was initial latency between the time a sample was collected from an individual through the time the result was reported to health authorities. The cumulative number of positive results from tests performed at TriCore as a function of the date of sample collection (Figure 1A) and the date of result reporting (Figure 1B) are shown. Performing a best fit analysis with either a quadratic (linear) or exponential model at day 17 yielded different results. Data based on date of sample collection best fit a quadratic ($y = 0.5435x^2 + 3.215x + 8$, $R^2 = 0.996$, Figure 1A), while data based on date of result best fit an exponential ($y = 3.0699e^{0.2355x}$, $R^2 = 0.99$, Figure 1B). On day 31 (04/12/2020), the actual case number was 696, which was in alignment with the quadratic model (630 cases, Figure 1) but not the exponential based on date of results (4,547 cases, Figure 1B). It is not unusual that data used to generate prediction models are based on date of results, as was the case in New Mexico and other states (3, 4). As a result, models based on date of result rather than date of collection may lead to overestimates of future case numbers. Our quadratic model (Figure 1A) has maintained $R^2 > 0.99$ on each day additional data are added, reinforcing the strength of the model and its projections. Using this model, we predict 1,470 positive cases on day 49 (04/30/2020). It is important to note that our model is only based on test results reported by TriCore and no other laboratories. TriCore performs ~65% of SARS-CoV-2 testing in New Mexico.

The reasons for non-exponential growth are multi-factorial and likely attributable to executive and public health orders issued by the state of New Mexico during the month of March, 2020. Orders such as the closing of educational facilities (March 13), prohibition of mass gatherings (March 19), and the closing of non-essential businesses and services (March 24) are probably key contributors.

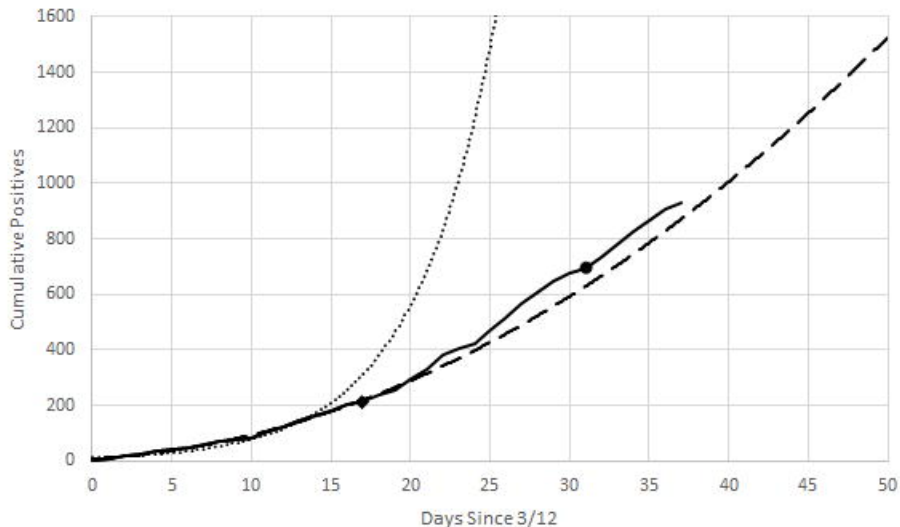
This COVID-19 pandemic has given a clear example of the critical role the laboratory plays in the healthcare system. As proposed in the Lab 2.0 movement, the value of the lab cannot be held at the cost per test level. Its value should be measured in the downstream impact of accurate, reliable and most importantly timely results that are used by not only by our clinical colleagues, but policy makers as they plan strategies to mitigate this pandemic.

References

1. Crawford JM, Shotorbani K, Sharma G, Crossey M, Kothari T, Lorey TS, et al. Improving American Healthcare Through “Clinical Lab 2.0”: A Project Santa Fe Report. *Acad Pathol.* 2017;4:1–8.
2. Swanson K, Dodd MR, VanNess R, Crossey M. Improving the Delivery of Healthcare through Clinical Diagnostic Insights: A Valuation of Laboratory Medicine through “Clinical Lab 2.0.” *J App Lab Med*; 2018;3:487–97.
3. New Mexico Department of Health. Coronavirus Disease 2019 in New Mexico. <https://cv.nmhealth.org> (Accessed April, 2020).
4. The COVID Tracking Project. <https://covidtracking.com> (Accessed April, 2020)

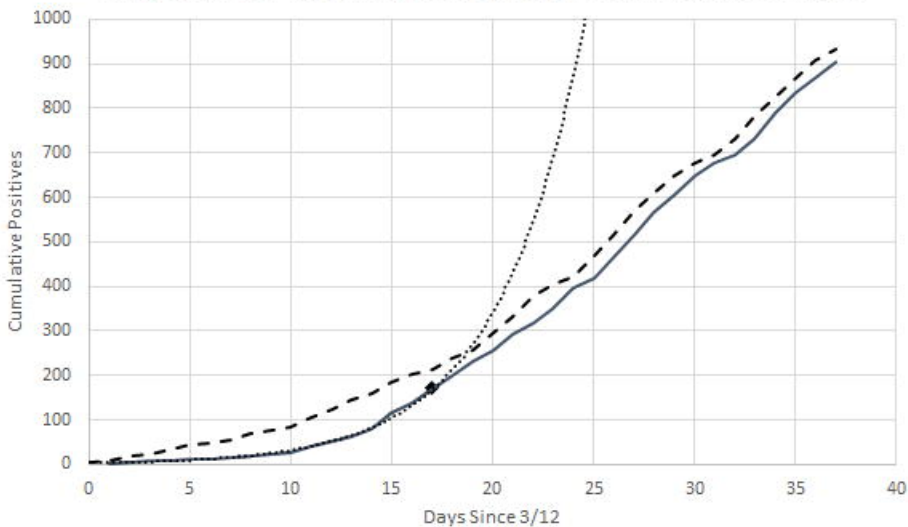
Fig 1. A) Actual cumulative positive cases, counted by date collected (black, solid). Using data collected through day 17 (◆), both quadratic (linear; dashed line) and exponential (dotted line) models were projected. The black dot (●) shows the actual number on day 31. B) Comparison of actual cumulative cases reported by date of collection (dashed line) and date of results (black, solid). The best-fit exponential projection based on the date of results on day 17 (◆) is shown. See text for details.

Models Based on Date of Collection



(A)

Comparison of Cases Reported by Date of Collection vs. Result



(B)