Journal of Business and Social Science Review
Issue: Vol. 1; No.10; October 2020 pp.35-43
ISSN 2690-0866 (Print) 2690-0874 (Online)
Website: www.jbssrnet.com
E-mail: editor@jbssrnet.com

COVID-19 Positivity Rate as a measure of the potency of the virus in Humans' War against the virus

Daniel Bocanegra Diaz, Bisrat Kinfemichael & Frank T. Lorne
New York Institute of Technology
School of Management

The spread of an infectious virus is not only a natural phenomenon, but an outcome depending on a large number of socio-economic and political factors. There is a need to educate the public of the progression of an infectious virus in a simple, easy-to-understand metric that minimizes misleading guidance for lifestyle and mobility. The objective of this paper is to emphasize the importance of the COVID-19 positivity rate, or percentage of confirmed cases among the tests performed in a population. This measure can be considered as potency rate in the context of the Coronavirus pandemic, because the war with virus is not only an internal body biological battle of how the virus invades the cell of a human body, but also the results of decisions made by individuals living in a society. In other words, it is also intrinsically a socio-economic-political problem.

In related papers, we described a five-pronged traveling map of virus as a reading the enemy combat strategy for humans to consider. To facilitate a discussion of testing and its implications on the potency rate, we modify that traveling map depending on the status of the individual: asymptomatic, positive, and with non-clinical symptoms. The model shows the journey on the virus, in humans, surfaces and geographies. With reliable data, the model can be included in tracking dashboards used in policy making decisions, and general information. The rate is used as a criterion for the readiness of opening an economy. We calculate the cumulative positivity rate for the United States, Canada, Chile, and Colombia as an illustration to show significant difference in regional trend relative to daily reported cases. Regardless of which region is being tracked, we argue that it is the trend of the positivity rate that is a more reliable statistics to track than the daily positive cases, a metric more commonly constantly used in the media.

Tracking vs. Epidemiology Studies of Virus

The spread and the severity of the COVID-19 virus is often narrated by media's daily new confirmed cases and deaths. A statistical measure of a virus spread is traditionally measured as Rnaught, which measures how contagious is a virus from a health reporting system based on new cases spreading from of existing cases. This paper argues that confirmed cases can increase partly because of an increase in the number of tests, and thus highly misleading. We examine a more informative simpler measure, the positivity rate, defined as confirmed cases divided by the total number of tests in a defined geographical area as reported by the correspondent health authority. It is assumed that most of asymptomatic cases, while not being able to be tracked by health systems uniformly, and given a virus incubation period of 14 days on average, are not being reported to the health authority statistics. The cumulative positivity potency rate (not the daily potency rate) could be a better measure for monitoring the spread of a virus over just confirmed new cases.

We wish to pursue the inquiry of the virus spread as different snapshots of the virus war across different states in United States. The first date of reference is April 23, 2020, when the total confirmed cases in the U.S. approached the 1 million mark (2020, Johns Hopkins University). A similar ratio, the true prevalence ratio, ideally measured, would require testing the entire population. That measure being impossible to implement, the cumulative positivity ratio serves as an estimate of the potency power of virus in an area in a given point of time. The positivity rate can also be understood as a summary statistic on the war fought between humans and the virus in a geographic area. For example, if the rate is one, it could mean that the virus has completely overwhelmed the population, and that its community has missed milder of asymptomatic cases (2020, Johns Hopkins University, Testing Hub). If the ratio is zero, it means the humans in a geographical area has completely defended themselves against the invasion. Virus, like humans, has productivity. Their productivity could be measured as how potent they can invade humans. Humans, like virus, also have their ways to defend the invasion.

The effectiveness of human's productivity defense method against the virus is 1 minus the potency of a virus. The potency of a virus, the test positivity rate, is a like other economic productivity measures, which usually refers to output per person. The relative productivity measure here is a zero-sum game.

A potency rate of virus can be estimated as broad for the world, or a country, or as narrow as a local community. The convergence of different geographical potency rates towards the same number is only relevant if travels between different geographical areas are 100% free with zero transportation costs, i.e. virus travels across regions without delays. The current reality is that there are large regional differences in methods of detecting the virus and the culture of fighting of it. Even within a smaller regional district in a country, the method of testing and the laboratories used may be different. According to Carter et. al., the RT-PCR is currently the gold standard for identification of SARS-CoV-2 / COVID-19 virus (2020, p.592). However, some studies had advised about the considerable proportion of false negative cases using this technique. Kucirka, et. al. (2020) states that the most used coronavirus tests, produces false negatives at least 20% of the time. This is where the power of statistics cannot tell us if something is necessarily "the truth and the whole truth."

As a matter of understanding aggregate data based on heterogenous settings, let the number of positive cases be equal to $f(X, Y)$, where X and Y are the total number of tests conducted in two different laboratories in a given time. Our proposed indicator for a potency of a virus, measured by the potency rate, suggests that the spread of the virus in a community is $f(X,Y)/(X+Y)$. The specification of $f(X,Y)$is highly determined by the methods used by particular lab, and it is a relevant question to ask if the medical researchers are "after the truth"

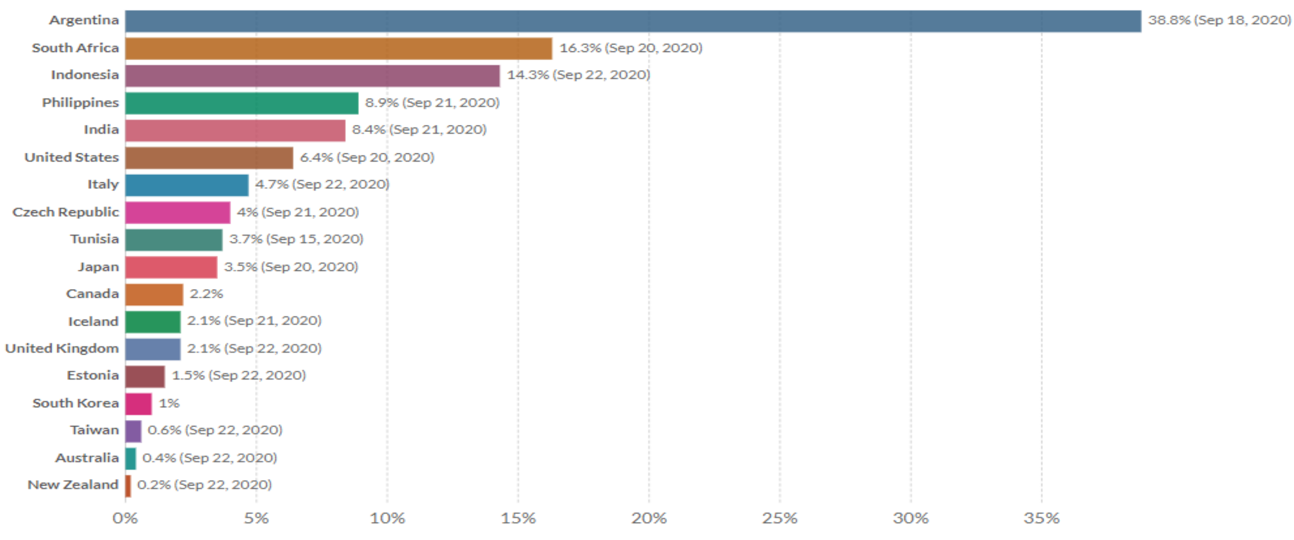
Catching every single infected individual in the population. The researcher interested in tracking the virus progression would want to track $f(X,Y)$ very cautious, making sure that every test is robust and standardized. The potency rate is a function of the sum of laboratory tests. This measure is different from other epidemiological ratios that has been used in various studies.

For example, the OurWorldinData (2020) COVID-19 tracking website calculates the inverse of potency rate as another relevant statistics to be monitored. It asks the question "how many tests are needed to find a COVID-19-positive case?" The website explains,

"No country knows the true number of people infected with COVID-19. All we know is the infection status of those who have been tested. The total number of people that have tested positive – the number of confirmed cases – is not the total number of people who have been infected. The true number of people infected with COVID-19 is much higher."

The same site also compares the positivity rate among countries, defined as the share of total COVID-19 tests that were positive. Figure 1 shows the total confirmed cases as a share of the total number of people tested, or the number of tests performed. As of September 18, 2020, Argentina is the country with the highest accumulated positivity rate for Covid-19, followed by South Africa and Indonesia (OurWorldinData, 2020).

Figure 1: Share of total COVID-19 tests that were positive, September 18, 2020



Adopted from OurWorldInData: COVID-19 Positive Rate Bar, 2020

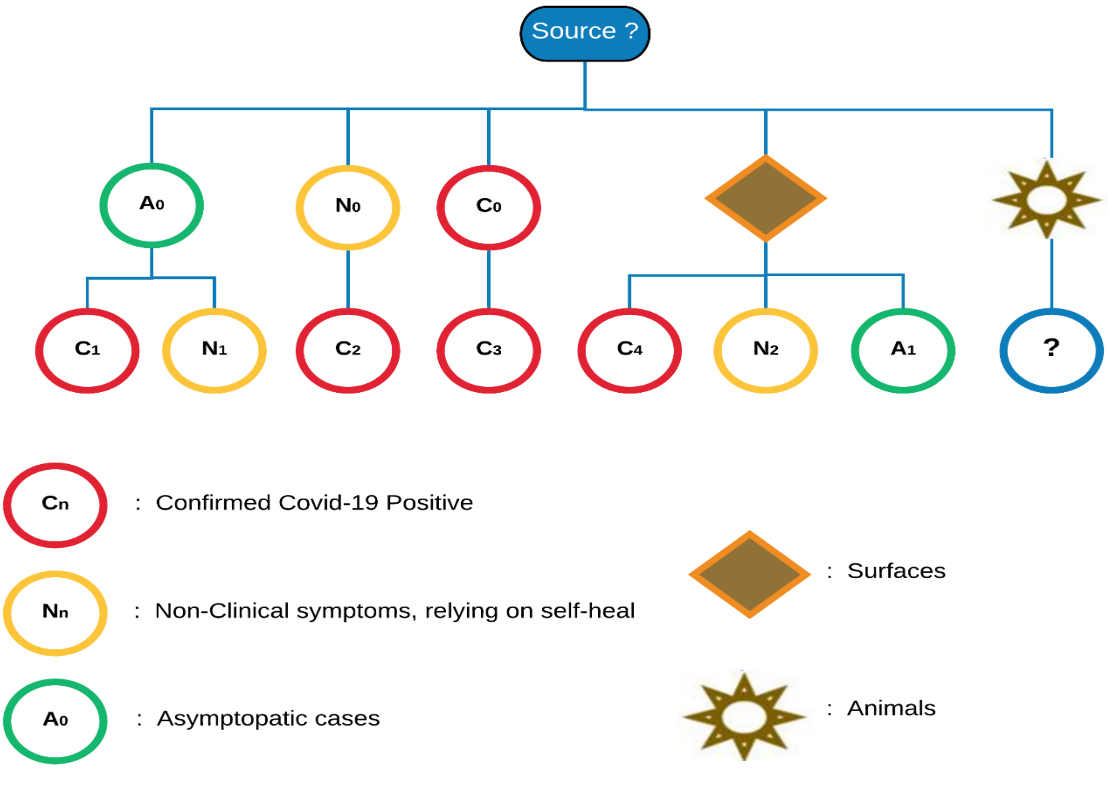
Independent of whether testing an entire population is doable or not, noticing that several economists have advocated that (with some arguing that it should be done repeatedly, at least twice a week), the potency rate method we propose in this paper accepts the fact that the war on virus can only be won on a certain extent, and not on the goal of complete elimination of it. The monitoring method of the potency rate can be useful for (1) tracking whether the virus effectiveness in penetration in a region is increasing (exponentially), or decreasing (with diminishing returns), and determining if the testing is missing mild or asymptomatic cases (2020, Johns Hopkins University Testing Hub), (2) tracking regional differences and examine whether underlying socio-economic factors and the warriors fighting the war have been effective.

The Traveling map of Virus outside of a human body

The reason why potency rate has so far not been used initially by most health jurisdictions, is probably because policy makers did not consider that the rate could capture missing cases, i.e. the unreported cases and the asymptomatic cases. For a better understanding, we propose a 5-prongs attack model of the virus, that can provide a justification and limitations of the measure approach. The traveling map of a virus is described in Figure 2, denoting the ways a virus can attack individuals in their different health status: confirmed COVID-19, non-clinical symptoms / relying on self-heal, and asymptomatic case patient; represented by a red, yellow, and green lights accordingly. Surfaces, and an animal transmission channel are also listed as possibilities in addition to the human lights. The 5-pronged attack will happen successively in different vertical levels representing a spread over time. A green light asymptomatic case can never be discovered unless tests are randomly conducted for the whole population. This seems an impossible task at first glance. However, if we understand the nature of the 5-pronged spread, we can say that at some stage, the green lights are bound to affect some, resulting in yellow and red lights.

Alternatively, if the green color-coded individuals are not able to infect others, by turning lights to yellow or green, either the virus is not that potent and likely to be diminished at one point, or the virus is hibernating. Basically, the potency rate measure assumes that the virus in the asymptomatic individuals, given that they are potentially contagious, will eventually be transmitted to someone else, turning to a yellow or a red light. When a case is red, and the individual checks into a health system clinic, he/she will be counted as a "confirmed case". In other words, the hidden virus in the body of asymptomatic will eventually appear and caught into the system, although some system can catch it faster than others.

Figure 2: ATraffic Lights model of the COVID-19 virus



The journey of a virus, turning individuals to red status, may take some twists and turns before getting caught. The infected yellow light cases may or may not report to the health system, as some patients might decide to rest at home or self-isolate. If they are reported, they are counted into the system, whether the test shows positive or negative. If they do not report and rest at home, the system will not know there is a virus out there *resting at someone's home*. The resting at home virus could have the same potency waiting period as the green lights. They may affect more individuals if the yellow lights patients do not self-isolate. The 5-prongs attack diagram shows that successive layer transmissions would imply that eventually it will be caught in terms of a red light. And for those viruses that never turn red, they are not sufficiently potent in those individuals.

Limitations on Socio-Economic Potency Rates for Tracking the COVID-19 Virus

There are many aspects of the traveling map of virus and the traffic color light approach that can be explored. One relevant aspect of this analysis is that irrespective of the origin of a virus, a medium of spread is the surface, denoted as diamond as one prong of the 5-prongs attack. Surfaces can be narrowly interpreted as landing sites of virus on a geographic area. They can be as narrow as a kitchen counter, or as broadly as the area covered by a particular health system. The point to note is that many health systems catches only the red lights individuals. For example, Dr. Bonnie Henry, lead health authority officer of British Columbia, Canada, in a June 17, 2020 letter to the province's business community explicitly said,

"At this time, it is recommendation that only people with symptoms or people otherwise identified by a heal professional should be tested for COVID-19" (Vancouver is Awesome, June 22, 2020). Widespread testing *"would create a significant burden for the public health system and would provide little value in protecting your business and could impede out ability to protect the health of all British Columbians."*

We do not know how many health systems around the world also shares the same sentiment. If that is the case, positivity rates as reported by health systems are likely to be a *lower bound* estimate of the true potency rates of virus for a population. Obviously, the assumption of 14 days isolation (worldwide practice now of checking travelers crossing borders) rests on a presumption that the COVID-19 virus in asymptomatic cases would lose its spreading potency if there are no symptoms during isolation. That makes the positivity rate calculation based on the cases reported to a health system a more reliable estimation of population potency rate as far as cumulative (not daily) positivity rate is concerned.

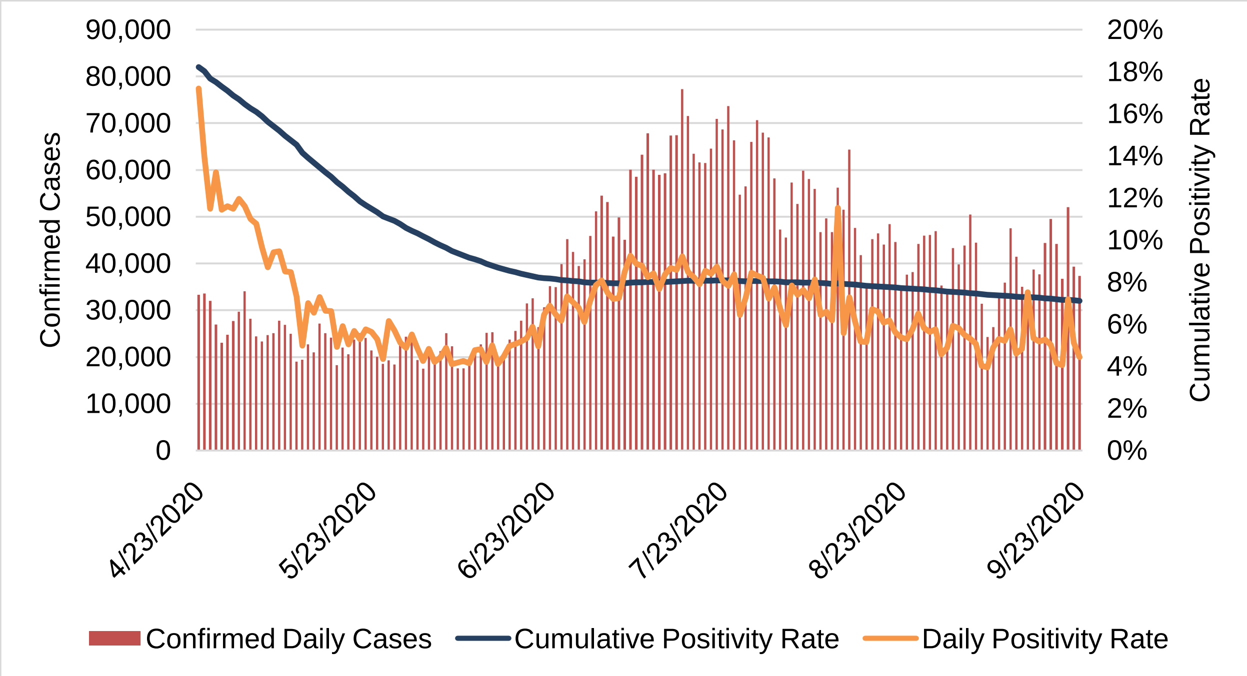
Another criticism for estimating potency rates on a given date, has to do with the turn-around time between the dates when the test is performed and its outcome. Hence, for any given date, the confirmed cases on that date are not the outcome of the tests conducted that same day. On the other hand, the number of days it takes for test results to be reported varies according the type of the test and the lab that is running it. The efficiency of the tests is likely to increase over time for a particular locality with the improvements of the method of analysis. Therefore, the error caused by the tests turn-around time lag is likely to decrease for given locality, with some testing kits now promising a 30-minute turnaround time for test results. Regardless whether fast turn-around time is achievable or not, the cumulative positivity rate is likely to be more accurate as the time history of tracking is longer for a locality. For all these reasons, even though the measure of the potency rate is not perfect, we expect the estimate error to be less consistently biased for a given locality over time, compared to others. To analyze the pattern of the positivity rate in a certain region, we calculated the rate in four specific countries.

Cumulative Positivity Rate versus daily Positivity Rate over time

Figures 3, and 4 compares the confirmed COVID-19 daily cases, against the cumulative, and daily positivity rate in the United States and Canada, from April 23, to September 23, 2020. According to Figure 3, the confirmed daily cases in the U.S. have progressively incremented from the middle of June, after a decreasing trend shown by the end of May of the same year. According to our calculation of the cumulative positivity rate, there was no second wave. The calculated for the U.S. is decreasing by more than double during the period, from 18.23% to 7.11% during this period. The seven-day moving average positivity rate calculated for the U.S. follows a similar trend, reducing from 15.60% to 4.94%. The decreasing cumulative and daily positivity rates shows that the potency of the virus is declining, even though there is a recent increase on the daily confirmed cases.

On May 12, 2020 the World Health Organization -WHO, issued a recommendation to governments and authorities, stating that the recommended positivity rate should be maintained lower than 5% for at least the last 14 days before reopening economies (2020, Johns Hopkins Testing Hub). The WHO considers that epidemiological criteria to indicate that the COVID-19 pandemicis being controlled (May 12, 2020, World Health Organization). After the recommendation from the organization, many cities and states in the U.S. are publishing the positivity rates in their health authority websites.

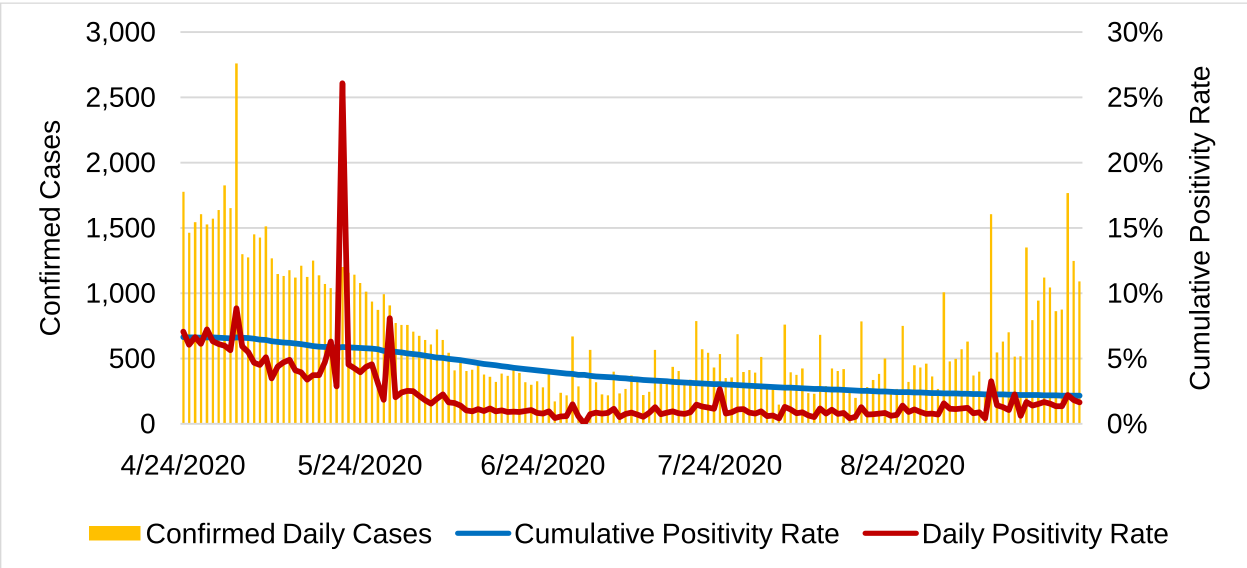
Figure 3: Daily Confirmed COVID-19 Cases, Cumulative Positivity Rate, and Daily Positivity Rate in the United States



Source: Johns Hopkins University CSSE (Confirmed cases), The Atlantic: The COVID Tracking Project (Tests performed).

In Canada, the confirmed daily cases, and both the cumulative and daily positivity rates, follows the same downtrend trajectory for the analyzed period according to Figure 4. Two peaks in the daily positive rate are registered on May 21 and 29, explained by a suddenly drop in the new tests reported. Considering the possibility of missing data, statistical adjustments, and other causes for daily fluctuations, some health authorities calculate the 7-day moving average of the positivity rate.

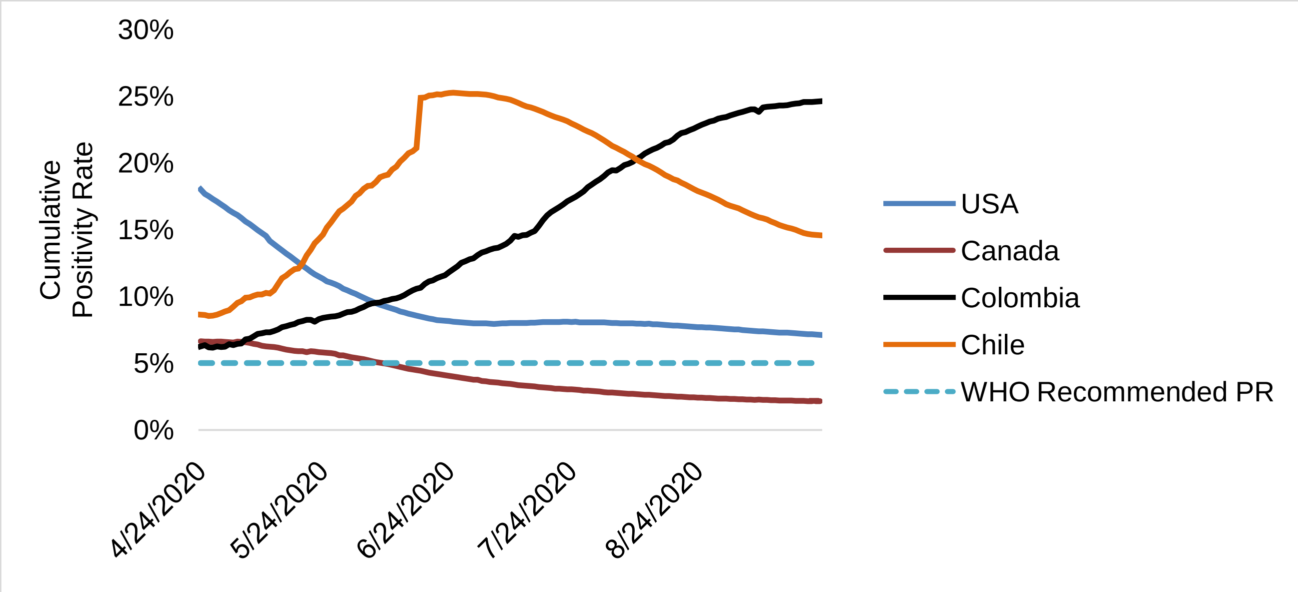
Figure 4: Daily Confirmed Covid-19 Cases, Cumulative Positivity Rate, and Daily Positivity Rate in Canada



Source: Public Health Agency of Canada: Public Health Infobase

Figure 5 compares the cumulative positivity rate between Canada, Chile, Colombia, and the United States in the same period. The data indicates that the cumulative positivity rate is higher in the two Latin American countries as compared to the U.S. and Canada. According to Our World in Data, Colombia has the lowest total tests per million people compared to the other three countries, as shown in Table 1. This could indicate that the South American country is not doing enough testing to find new positive cases. Considering that Chile had performed COVID-19 tests per capita in a number close to Canada, it is very likely that the COVID-19 virus has had a higher potency in Chile as compared to Canada.

Figure 5: Cumulated Positivity Rate: Canada, Chile, Colombia, and the United States



Source: Johns Hopkins University CSSE (U.S., and Colombian confirmed cases), The Atlantic: The COVID Tracking Project (U.S. tests performed), Public Health Agency of Canada: Public Health Infobase (Canadian cases and tests performed), Government of Colombia – National Health Institute (Colombian tests performed), Science, Technology, Knowledge and Innovation Ministry, Health Ministry of Chile (Chilean cases and tests performed).

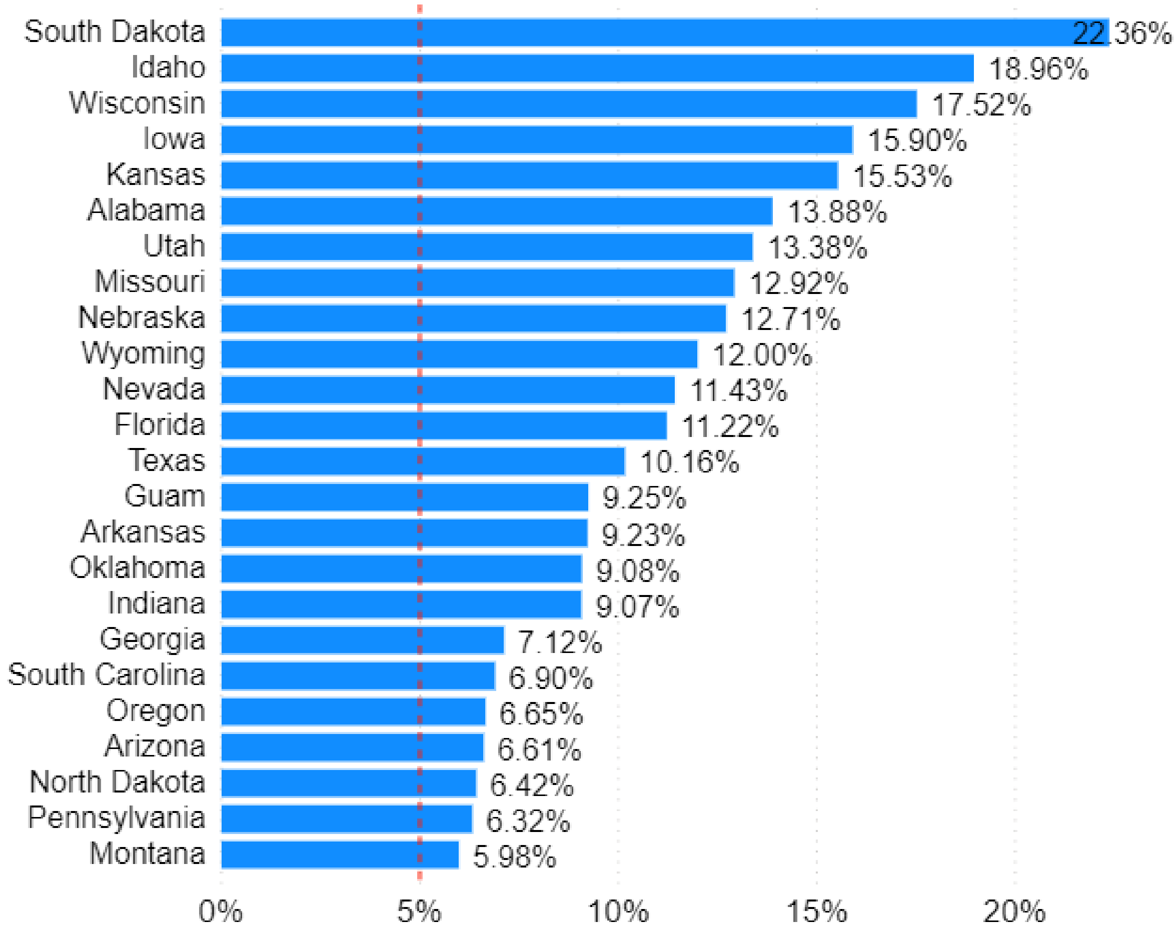
Table 1: COVID-19 tests and confirmed cases per million people.

Country	**Test per Million people**	**Total Confirmed cases per million people**	**Cumulative Positivity Rate**	**As of date**
United States	320,703	20,835	7.15%	September 20, 2020
Canada	179,295	3,886	2.16%	September 24, 2020
Chile	160,101	23,463	14.62%	September 22, 2020
Colombia	62,115	15,281	24.60%	September 22, 2020

Source: Our World in Data

In addition to the graphs presented in this paper, we created the site covidcharts.wixsite.com/powerbi which is a tracker for different COVID-19 statistics, including the cumulated positivity rate by available countries, the 7 and 14 day moving average positivity rate for U.S. states, and the trajectories of confirmed cases by number of days since the 100^{th} case per country. Figure 6 shows the U.S. states with highest positivity rate on a 7-day moving average, as of September 24, 2020. South Dakota is the state with the highest 7-day positivity rate (22.36%), followed by Idaho (18.96%), Wisconsin (17.52%), and Iowa (15.90%). The data indicates that the virus is gaining potency primarily in the Midwest. This regional differences are significantly different from that we tracked on April 23, 2020, which ranked New Jersey as the highest positivity rate state, with ratio of 0.498, New York at second with 0.363, and Connecticut third at 0.317, all adjacent and located next to each other in the North-east corner of U. S.

Figure 6: U.S. States by Positivity Rate: 7-day Moving Average (as of September 23, 2020)



Source: covidcharts.wixsite.com/powerbi using data from The COVID Tracking Project

COVID-19 Positivity Rate across states in United States

In this section, we present an empirical study of the cumulative positivity rate across U.S. states, using cross-sectional data reported on April 23, 2020. Using a regression model, we tested the dependency of the number of confirmed cases (C), the number of COVID-19 tests, and the cumulative positivity rate (C/T) on the following state's variables: Total area (Land-Km²), GDP per Capita (GDP/Capita), Population (POP), African American population of the State as a number (B-BOP), and as percentage of the population (B-POPP), prisoners population (P-POP), prisoners population as a percentage of the population (P-POPP), senior population (S-POP), senior population as a population (S-POPP), percentage of health care workers as a percentage of the population (H-Care%), state's total health care professionals (H-POP), health care professionals as a percentage of the population (H-POPP), health care location quotient compared with the national average (H-National), healthcare professionals per thousand habitants (H-Density), health care annual mean wage (H-Wage), and a dummy variable for the political party of the current State Governor (Governor) where 1 = Democrat, 0 = Republican.

Table 2 shows the regression results, using STATA. Columns 1, 2, and 3 shows the significance of the coefficients for the total number of confirmed cases (C), total number of tests performed, and the cumulative positivity rate (C/T), respectively. According to the results, there is a positive relationship between the African American population as a percentage of the total population (B-POPP), and the cumulative positivity rate (C/T). B-POPP is the only significative variable explaining the cumulative positivity rate, in this case at 10% (* p<0.1). On columns 1 and 2, the coefficients marked with * are statistically significant at 10%, those with ** are significant at 5%, and coefficients marked with ***at 1%. We reran the regression using data reported in July, 2020, 2 weeks after the long weekend, with no noticeable difference of results.

Table 2: Regression Analysis

	(1) **C**	(2) **T**	(3) **C/T**
H-Care%	1.82e+06**	9.97e+06	0.669
	(8.60e+05)	(6.31e+06)	(0.585)
Land-Km²	-0.053	-0.660	-0.000
	(0.115)	(0.841)	(0.000)
GDP/Capita	3.607**	20.408*	0.000
	(1.373)	(10.073)	(0.000)
B-POPP	1.61e+05	6.68e+05	0.186**
	(1.09e+05)	(7.97e+05)	(0.074)
P-POP	0.732*	13.258***	-0.000
	(0.376)	(2.757)	(0.000)
P-POPP	-1.26e+07*	-1.15e+08**	0.600
	(7.02e+06)	(5.15e+07)	(4.775)
S-POPP	1.43e+05	1.16e+06	-0.532
	(6.47e+05)	(4.75e+06)	(0.440)
H-POPP	-7.12e+06*	-4.83e+07*	0.923
	(3.59e+06)	(2.63e+07)	(2.441)
H-Wage	-1.050	-6.895	0.000
	(1.322)	(9.696)	(0.000)
Governor	12320.771	1.38e+05	-0.000
	(15948.227)	(1.17e+05)	(0.011)
Number of observations	47	47	47

* p<0.1, ** p<0.05, *** p<0.01

The purpose of the exercise in Table 2 is to show that potency rate of the virus as we define in this paper, measured by the positivity rate, can also be a socio-economic variable similar to the productivity of labor for an economy, except that in this case, the virus ability to infect, spread, and make people sick, should also be depended on many other variables. For example, had we have data on the percentage of people in a state wear masks in public space (which may be impossible to collect), the coefficient of that variable on positivity rate is likely be significantly negative.

It is unlikely that we can capture all the variables that could affect the outcome of the war against the COVID-19 virus, but that is not the purpose of this study. We do not intend to prove that socio-economic variables *can explain* the potency of a virus. It is our intention to show that a better proxy of the positivity rate based on cumulative confirmed case can more effectively capture asymptomatic cases and thus more useful. Projecting forward, we believe that by understanding more about positivity rate implications on the society, health authorities in the world may want to calculate the positivity rate by age, comorbidities, income, ethnicity, and other socio-economic factors. More effective and targeted public policy could be implemented by analyzing the potency rate trend in specific populations.

Conclusion

This paper argues that positivity rate is a better statistic measure than the daily confirmed case as a monitor for a virus spread, as asymptomatic cases will remain dormant for some period before being counted into the Health Authority statistics. The positivity rate of virus is like the productivity of a "soldier group", agents in a war against the humans. Because humans' measures against the enemy are different across different regions and different health authority policies, we see regional difference in positivity rates. We suggest how positivity rates can potentially be studied by various socio-economic variables.

On theoretical grounds, we can hypothesize that if some potential factors could have an impact, we should examine its effect on positivity rates. However, we could only gather proxies. The limitation to our study is that the variables used may or may not be accurately accounting for the potency rate. This is an inherit difficulty in social science research. We can only examine the data patterns and analyze the results. The positivity rate is a measure that tracks the potency of the spread of the virus in a population.

Policy makers, health authorities and media should increase their attention to this measure, making them more accurate, and examining its implication over human activities.

Reference:

Bocanegra, D. (2020). COVID Log Charts. Retrieved from https://covidcharts.wixsite.com/powerbi

Carter, L. J., Garner, L. V., Smoot, J. W., Li, Y., Zhou, Q., Saveson, C. J., Sasso, J. M., Gregg, A. C., Soares, D. J., Beskid, T. R., Jervey, S. R., & Liu, C. (2020). Assay Techniques and Test Development for COVID-19 Diagnosis. ACS central science, 6(5), 591–605. https://doi.org/10.1021/acscentsci.0c00501

Johns Hopkins University. (2020). COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE). Retrieved from: https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv

Johns Hopkins University. (2020). Testing Hub: Which U.S. States Meet WHO Recommended Testing Criteria? Retrieved from: https://coronavirus.jhu.edu/testing/testing-positivity

Lorne, Frank T. (2020) The 5-pronged Attack of the Coronavirus War, Mimeo

Lorne, Frank T., A.C. Lai, Sairam Katla, Krish Kale, D. Bocanegra, "Community-Based Solution for a Community Spread requires Incentive-Compatibility Considerations" *Con ScienS Conference Proceedings*, Sept 28-29.

Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D., &Lessler, J. (2020). Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure. Annals of Internal Medicinedoi: 10.7326/M20-1495

National Health Institute of Colombia and *DatosAbiertos* - Open Data. (2020). COVID-19 tests performed in Colombia (*Muestras procesadas de COVID-19 en Colombia*). Retrieved from: https://www.datos.gov.co/Salud-y-Protecci-n-Social/Muestras-procesadas-de-COVID-19-en-Colombia/8835-5baf

Our World in Data. (2020). Total COVID-19 tests conducted vs. confirmed cases per million. Retrieved from: https://ourworldindata.org/grapher/covid-19-tests-cases-scatter-with-comparisons?tab=table&yScale=linear

Our World in Data. (2020). Tests conducted per new confirmed case of COVID-19. Retrieved from: https://ourworldindata.org/grapher/tests-per-confirmed-case-daily-smoothed?tab=chart

Public Health Agency of Canada. (2020). Public Health Infobase, Interactive data visualizations of COVID-19. Retrieved from: https://health-infobase.canada.ca/covid-19/

Science, Technology, Knowledge and Innovation Ministry, Health Ministry of Chile. (2020). COVID-19 Database : DP5 – Daily National Totals *(Base de Datos COVID-19: DP5 - Totales Nacionales Diarios).* Retrieved from: https://github.com/MinCiencia/Datos-COVID19/tree/master/output/producto5

Science, Technology, Knowledge and Innovation Ministry, Health Ministry of Chile. (2020). COVID-19 Database : DP17 –Cumulative PCR tests reported in the last date by type of establishment. *(Base de Datos COVID-19: DP17 - PCR acumulado e informado en el último día por tipo de establecimientos).* Retrieved from: https://github.com/MinCiencia/Datos-COVID19/tree/master/output/producto5

The Atlantic. (2020). The COVID Tracking Project: US Current and Historical Data. Retrieved from: https://covidtracking.com/data/api

World Health Organization. (May 12, 2020). Public health criteria to adjust public health and social measures in the context of COVID-19: Annex to Considerations in adjusting public health and social measures in the context of COVID-19. Retrieved from: https://www.who.int/publications/i/item/public-health-criteria-to-adjust-public-health-and-social-measures-in-the-context-of-covid-19