AUA American University of Armenia
AVEDISIAN CENTER for BUSINESS RESEARCH & DEVELOPMENT

AUA American University of Armenia
MANOOGIAN COLLEGE of BUSINESS & ECONOMICS

CORONAVIRUS STATISTICS: LEARNING FROM THE SUCCESS STORIES

CENTER FOR BUSINESS RESEARCH AND DEVELOPMENT RESEARCH NOTE

JULY 2020
YEREVAN, ARMENIA

TITLE

VAHRAM GHUSHCHYAN, PhD
Manoogian Simone College of Business and Economics
American University of Armenia

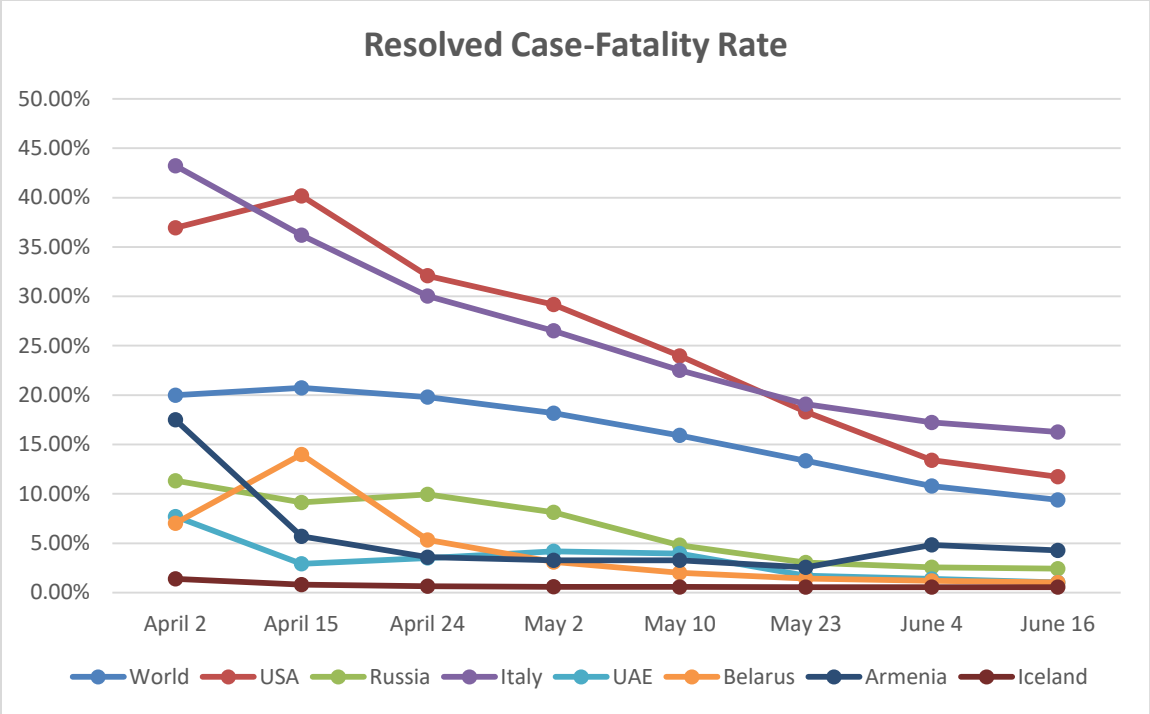
Key Findings

- Countries that achieved and kept the resolved case-fatality rate at 0.5%-1% by performing extensive testing were able to stop or significantly slow down the spread of the virus.
- Average mortality rate of COVID-19 is <550 per 100,000 infected, and overall population mortality risk is around 0.2-0.45%.
- Asymptomatic cases are common in coronavirus infected patients, every second patient is likely to show no symptoms when infected.
- If Armenia wants to stop the spread of the virus it has to significantly increase the level of testing

For the last 2.5 months, I have periodically collected data on coronavirus statistics from the worldometers.info website. The data from this site is often used in the media and government reports. Thus, as of June 16, 2020, there are more than 8 million confirmed cases of COVID-19 and 440K reported deaths globally. There is a huge asymmetry of information flows and contradicting reports on the mortality of the virus. The aim of this research is to analyze reported numbers, identify the sources of bias in the data, find success stories that are informative for decision making and health policy crafting, and develop a lasso machine learning model to estimate the incremental effect of testing on the resolved case-fatality rate. In this analysis, I use reported global statistics on the COVID-19 confirmed cases and clinical research recently published as a letter to the editor in the *New England Journal of Medicine*, one of the most respected scientific journals in the field. The results suggest that 1) reported resolved case-fatality rates of COVID-19 in many countries are biased upwards due to insufficient and/or delayed testing, 2) about half of the patients infected with the virus remain asymptomatic, 3) older patients with comorbidities who get infected by coronavirus can recover and even be asymptomatic, 4) countries that were able to perform more than 100,000 tests per 1 million population in a short period of time succeeded in containing the spread of the coronavirus, and 5) if Armenia wants to stop the spread of the virus it has to significantly increase the level of testing.

Bias in the RCFR. Resolved case fatality rate (RCFR): is calculated as confirmed deaths/resolved cases (deaths + recovered). Given that the COVID-19 outbreak is still active and different countries are in different stages of the outbreak, I find the RCFR to be a more dynamic measure of disease severity. The table and figure below present the dynamics of the RCFR for the last few months for selected countries.

	April 2	April 15	April 24	May 2	May 10	May 23	June 4	June 16
World	19.99%	20.73%	19.80%	18.17%	15.92%	13.34%	10.79%	9.38%
USA	36.94%	40.17%	32.09%	29.17%	23.96%	18.30%	13.40%	11.73%
Russia	11.32%	9.12%	9.95%	8.12%	4.81%	3.04%	2.56%	2.42%
Italy	43.22%	36.20%	30.03%	26.52%	22.51%	19.08%	17.22%	16.26%
UAE	7.69%	2.91%	3.51%	4.18%	3.96%	1.72%	1.38%	1.02%
Belarus	7.02%	13.98%	5.33%	3.09%	2.00%	1.41%	1.18%	1.02%
Armenia	17.50%	5.69%	3.58%	3.27%	3.27%	2.56%	4.83%	4.27%
Iceland	1.39%	0.80%	0.64%	0.59%	0.56%	0.56%	0.55%	0.55%



The downwards trend in the RCFR is noticeable. For many countries, it is also highly correlated with the number of tests performed per million population. For example, the correlation coefficient between the resolved case-fatality rate and tests per 1 million population is -0.69 for Belarus, -0.90 for Iceland, and -0.04 for Armenia. The RCFR is based on the number of infected. Since as many as 40-60% of cases may be asymptomatic, the RCFR decreases significantly with the increase in the number of performed tests. Thus, the "true" average mortality rate of COVID-19 is <550 per 100,000 infected, and overall population mortality risk is around 0.25-0.45%, given that kids and adolescents in most cases are not infected and often not even tested.

Asymptomatic cases are common in coronavirus infected patients. For example, 44% of infected patients on the Diamond Princess cruise ship were asymptomatic. Diamond Princess cruise ship started the well-known cruise on January 20, 2020. The first positive cases of COVID-19 were confirmed on the ship in February 2020. 712 people, 567 out of 2666 passengers and 145 out of 1045 crew, tested positive for the virus, and 14 died. All deaths were reported among the passengers. The median age of passengers was 69. The cases on the ship can be viewed as a natural experiment and are well studied and

reported[1]. This study found that even older patients with comorbid conditions can be asymptomatic. Out of 712 patients that tested positive for COVID-19, 311 remained asymptomatic, which accounts for 44% of infected. Thus, given the average age of people on the cruise, we can conclude that in the general population, the asymptomatic cases are going to be well above 44%.

Among the reported statistics, the Faeroe Islands and Iceland's cases stand out the most. The Faeroe Islands has a small population of about 60 thousand people. As of June 16, they have performed 225,192 tests per 1 million population, have identified 187 infected patients, and were able to treat all of them successfully. The country has developed and performed all the tests locally, and the combination of high-volume testing and quarantining gave the desired outcome: they have had zero fatalities, and the spread of the virus has stopped completely in the country.

As of April 15, 1 month after the lockdown, Iceland had 1,720 and Armenia had 1,067 confirmed cases of COVID-19; however, Iceland had already performed 106,490 tests per 1 million population while Armenia has performed 2,806 tests per 1 million population. About two weeks later, on April 24, the number of completed tests per 1 million was 143,988 in Iceland and 7,484 in Armenia. The corresponding confirmed cases of COVID-19 were 2,148 for Armenia and 1,798 for Iceland. Thus, massive, high-volume testing is needed to curb the virus outbreak. Selective and targeted testing did not prevent the spread of the virus, as is evidenced by the Armenian case.

Several factors influence the treatment outcome of COVID-19 patients. Age, gender, and existing chronic conditions are important predictors of the treatment outcome. On the country level, testing volumes, age distribution, life expectancy, health care management, social behavior, lockdown enforcement methods, and several other factors influence the RCFR. In this research, I want to estimate the relationship between RCFR and testing volumes; in particular, I am interested in the incremental effect of testing on the case-fatality rate. Given the numerous factors that influence the case-fatality and limited data available, I have used the least absolute shrinkage and selection operator (lasso) machine

[1] Sakurai et al. June 12, 2020. https://www.nejm.org/doi/full/10.1056/NEJMc2013020

learning method for inferential statistics[2] to accurately identify the incremental effect of testing on the RCFR. In this analysis, resolved case-fatality is the outcome variable, the number of tests per 1 million population is the key variable of interest, median age of population, life expectancy, health care expenditures as a percent of GDP, and the interactions of these variables are used as controls.

The analytic sample included more than 200 countries for which the coronavirus related statistics were collected seven times during 2.5 months. The final sample used for the analysis had 922 observations, and lasso selected 31 control variables. In the linear-log model specification, lasso estimated the marginal effect of the number of tests per 1 million population to be -4.87 (95% c.i. -5.66, -4.07). Thus, for each 1% increase in the number of tests per million population, the RCFR decreases by approximately 0.05 percentage points. This information is essential for crafting an anti-coronavirus testing policy and defining the optimal number of tests needed to slow down and stop the spread of the virus. The testing for coronavirus is an ongoing process, and our estimates are valid for a 2-week time interval. Countries that achieved and kept the resolved case-fatality rate at 0.5%-1% by performing extensive testing were able to stop or significantly slow down the spread of the virus.

[2] Lasso Reference Manual. https://www.stata-press.com/manuals/lasso-reference-manual/