Optimization of Public Health Policies at Institutes of Higher Education Using Statistical Analyses of COVID-19 Data

A Capstone Proposal Submitted to the

Department of Biomedical Engineering
School of Engineering and Applied Science
University of Virginia

Catherine M. Pommersheim
Fall 2020

Signed _____ Date: 11.17.2020
Catherine Pommersheim

Signed _____ Date: 11/17/20
Faculty Advisor, Brent French

I. SPECIFIC AIMS

The COVID-19 pandemic spread rapidly throughout the entire world resulting in 34.1 million cases and over one million deaths worldwide in less than a year while completely disrupting global socioeconomic systems. The emergence of this disease, caused by the SARS-CoV-2 virus, challenged public health experts to protect the population and develop strategies to prevent further hospitalizations and casualties resulting from infection. One necessary piece of information that became urgent to identify early in the pandemic is the method through which COVID-19 is transmitted from person to person. Originally, many believed that SARS-CoV-2 could remain on surfaces and infect people for long periods of time which led to panic and a shortage of household cleaning supplies across the United States. However, it is now known that SARS-CoV-2, like other coronavi surfaces for more than a few hours or days and is very unlikely to be the sour e most common way that infections occur is through respiratory droplets passed from person-to-person. This information helped shape public policy to encourage 6-foot separation and mandatory face mask policies within the larger strategy of social distancing to limit the spread of COVID-19.

major mode of transmission?

The novel coronavirus is a difficult disease to control, as evidenced by the ongoing pandemic, particularly because the primary mode of transmission is person-to-person contact even when infected individuals may be asymptomatic. Asymptomatic transmission is thought to be the leading way that COVID-19 is spread amongst the population. As a result, public health efforts to control the virus have little effect if people are actively contagious and infecting others without their knowledge. This project aims to examine the spread of COVID-19 through the population and hopefully gain insight into the best ways to limit its transmission through statistical analyses on positive case data (**Aim 1**). The data of interest has been provided to the public by the New York Times and is sampled from colleges and universities across the United States. Higher education can be slowed? unique window into disease transmission because of the number of individuals within a controlled geographic region. To further understand how the spread of COVID-19 is slowed, a thorough review of the preventative measures implemented at the schools in this study will be conducted. Information needed from this exploration includes the details of any mitigation efforts as well as the time at which they were instated. Furthermore, regional mask-wearing tendencies must be considered when accounting for effectiveness of COVID-19 preventative measures which will be extracted from data assembled by the New York Times. (**Aim 2**). The team will then compare the results of our analyses in **Aim 1** with the qualitative information from **Aim 2** to form actionable recommendations on the efficacy of public health policies in mitigating COVID-19 on higher education campuses. (**Aim 3**).

Aim 1: Perform statistical analyses on COVID-19 data from higher education institutions
- A. Analyze the positive case data from schools provided by the New York Times
  - i. Test for regional variations and type of institution

Aim 2: Identify and assemble qualitative data on public health policies and mask-wearing compliance
- A. Perform a thorough review of the preventative measures enacted at institutions of interest
- B. Extract and organize regional mask compliance information from the New York Times

Aim 3: Compare the results of the analyses in Aim 1 with that of Aim 2 to form actionable recommendations on public health at higher education institutions
- A. Contrast the results of **Aim 1** with health policies and mask compliance both separately and together
- B. Use observable trends from **Aim 3A** to serve as the basis for developing recommendations on the efficacy of public health policies at higher education institutions

The recommendations from this study will be directly applicable to COVID-19 mitigation efforts across campuses in the United States. Additionally, this information will help inform future public health decisions in the event of a future disease outbreak on higher education campuses to more rapidly implement preventative measures that are proven to be effective.

II. RESEARCH STRATEGY

(A) Significance

The emergence of the novel coronavirus, COVID-19, in late-2019 shocked and confounded pubic health experts as it rapidly spread throughout the world. By October 2020, over 41.3 million cases of COVID-19 had been reported globally as well as 7.8 million cases nationally (Times, 2020). This disease has been particularly difficult to contain due to its relatively high rate of asymptomatic cases as well as the wide array of potential symptoms. With symptoms ranging from a cough and headache to gastrointestinal issues, COVID-19 may go unnoticed in individuals who are not attentive to minute changes in their health. In March 2020, the pandemic first reached the United States and created widespread panic throughout the public over a novel virus and its potential ramifications. At this point, those studying the virus itself, SARS-CoV-2, were unsure of the mechanism by which it spread from person-to-person. This uncertainty resulted in public panic and, consequently, nationwide shortages of basic cleaning products including disinfecting wipes and hand sanitizer over the fear of COVID-19 spreading on surfaces. After more extensive research, it is now known that the primary mechanism for the spread of SARS-CoV-2 is through respiratory droplets spread in close contact between two people (CDC, 2020). While SARS-CoV-2 may survive on surfaces and remain infectious for up to a few hours or days, this is not the primary mechanism of infection. This new information on the spread of COVID-19 helped influence public health policy including the implementation of the face covering and mask requirements which have been proven effective against this disease.

The importance of accurate information in the fight against COVID-19 is exemplified in the scenario presented above of identifying the transmission mechanism of this disease. While major public health measures have been taken to slow and ultimately end the spread of COVID-19, the number of positive cases is currently increasing again (Times, 2020). There are currently multiple promising vaccine candid[illegible]V-2 in development and clinical trials; however, their distribution is n[illegible]e immediate future. As a result, the key to containing this disease remains [illegible]ead. The returned increase in positive case numbers indicates that alternate m[illegible]er consideration in the fight against this disease. Some such methods have been developed and implemented already including asymptomatic testing and wastewater-based epidemiology. In addition to increased testing procedures, it is important to understand the specific dynamics of a community in which COVID-19 is prevalent. For example, a rural environment, an urban environment, and a college campus may all experience varying severity of the spread of this disease by how quickly it is able to move through a community. Factors that may influence the transmission of this disease include population density, current case numbers, and other information such as public perception of the pandemic as well as personal views on face coverings.

third surge in positive case numbers?

additional?

The proposed project will investigate how COVID-19 has spread throughout different communities and consider confounding factors that may have affected its rate of transmission. In particular, this project will focus on higher education institutions in the United States, with a focus on the University of Virginia (UVA). Additionally, the effect that preventative measures and regional mask compliance have on COVID-19 transmission will be considered. This project aims to analyze existing COVID-19 case data using a statistical methodology adapted from a previous study performed by R. Gonzalez et al. on COVID-19 genomic data from wastewater surveillance in Southeastern Virginia. The information resulting from this analysis will then be used to compare public health policies and mask compliance in order to provide actionable intelligence on successful methods to combat COVID-19 at higher education institutions.

Are we sure this will be the right methodology?

(B) Innovation

The analysis of viral data, even for COVID-19, is not innovative in itself as similar studies have been published. For example, the study performed by R. Gonzalez et al. implemented a statistical analysis on SARS-CoV-2 concentrations in wastewater to identify the spread of COVID-19 through

daily case count data?

Southeastern Virginia (Gonzalez, 2020). While this project intends to work with preexisting COVID-19 case data instead, we plan to adopt a similar methodology to that used in this study to extract information on viral transmission patterns. The limitation of current work is that there has not yet been a correlative study done with COVID-19 data between health policies and virus transmission in the context of university campuses. This gap of knowledge limits the strength of policy recommendations as there is little basis for their efficacy among higher education institutions, let alone any association with regional mask compliance. The novelty of this project comes from the information we hope to uncover and its potential applications to public health. The knowledge of how college campuses react to the pandemic in conjunction with information on their COVID-19 policies will provide public health experts with evidence for the effectiveness of preventative methods. Additionally, comparing the results of our analyses with mask compliance can reveal the effect that regional differences may have on COVID-19 transmission of at colleges. **The proposed research will provide novel insights on the spread of COVID-19 across higher education campuses and the success of preventative measures in order to inform public health policies to better protect the public from this ongoing pandemic**.

(C) Approach

The general strategy to achieve the aims of this project involves performing a variety of statistical tests on COVID-19 case data from universities followed by an organized qualitative analysis of related health policies and regional mask compliance. Finally, the results of these analyses will be compiled to identify trends in COVID-19 transmission in response to both public health policies and regional mask compliance. The results of this study will be used to form recommendations for future public health measures based on the efficacy of current preventative measures at universities.

Aim 1: Perform statistical analyses on COVID-19 data from higher education institutions

Rationale. The importance of analyzing existing COVID-19 transmission data is to not only track the number of positive cases, but to compare this statistic between different sources. In this study, the sources of interest are colleges and universities in the contiguous United States. The unique population dynamics on a college campus provide an ideal lens into COVID-19 transmission patterns. Additionally, the frequent monitoring and testing implemented at most colleges provides fairly representative data on the true number of positive cases compared with more sporadic testing amongst the general population. Combining the information above with the preventative measures implemented at each school creates a clear picture of the effect that COVID-19 is having on college campuses.

percentage?

Methods. The tests that will be performed on the COVID-19 case data from universities will compare the mean number of positive cases between schools. To test for regional differences in schools, 20 schools will be selected from each geographical region in the United States. For the purposes of this study, geographical regions of interest include: Northeastern, Mid-Atlantic, Southeastern, Southwestern, Pacific Northwestern, Rocky Mountain Western, and Midwestern. To more accurately represent the schools within a given region, 10 will be randomly selected from private institutions while the other 10 will be randomly selected from public institutions. For the first part of this study, the public or private status of a school is not important and the mean number of positive cases from all 20 schools will be aggregated to create an average from each region. These averages will then be compared to test for significant differences in positive case numbers between geographic region. Additionally, the mean number of positive cases from UVA will be included in this analysis as its own variable to be compared with the regional average whether or not UVA is randomly selected from the Mid-Atlantic region of schools. The next part of this analysis will compare average case numbers between public and private institutions within each region and as well as nationally. To control for the usually significant size difference between public and private institutions, the proportion of positive cases within a university's student body and staff will be compared instead of simply the average number of mean cases. This requires the extra step of

identifying the number of people enrolled or employed at each school in the study. Once the proportions of positive cases are assembled, the average proportion from each school type within each geographic region will become its own variable. First, tests to compare the mean proportion within each of the seven geographic regions will be performed. Next, the national average proportion of positive cases from each type of school will be calculated and then compared to test for a significant difference. The proportion of positive cases within the UVA community will also be included in each of these analyses to compare the transmission of COVID-19 on Grounds to regional and national averages from different types of schools. To create the most representative picture of viral transmission, data will be collected from the New York Times' publicly available database at the end of the Fall 2020 semester. As a result, analysis will not begin on this portion of the study until mid-to-late December 2020.

Expected Results. A preliminary investigation into COVID-19 cases among U.S. colleges and universities revealed that schools in the Southeastern and Midwestern regions have reported the highest number of positive cases as of November 2020 (Times, 2020). Based on this information, one expects to observe high average positive case numbers from these two regions at the end of the Fall semester as well. Based on prior knowledge of private universities' typically small number of students, one can expect that these schools will produce lower average positive case numbers than public institutions which are usually characterized by having a much larger student body.

Limitations. The main limitation in our methods for **Aim 1** is that the positive COVID-19 data we will be examining is restricted to only the positive tests that were reported to each school. For example, if students at a university tested positive for COVID-19 but did not self-report to the school, that information will not be accounted for in this analysis. Similarly, if positive cases were reported more than once, our team will not be aware of duplicate entries in the case of multiple positive tests from one individual. Finally, the New York Times data only contains information on schools with at least one known positive case so if a school is unaware of a positive test or has no cases, it will not be included in the population of schools to be selected from for this study. This has the potential to skew our analysis to report a disproportionally high number of positive cases as zero-case schools will not be included in our population. As a result, a few assumptions must be made to simplify our analysis: 1) Every positive COVID-19 case from a school is reported, 2) Every positive COVID-19 case that is reported belongs to distinct individuals, and 3) Every school in the contiguous U.S. has had one or more positive COVID-19 case and it has been reported to the New York Times.

Aim 2: Identify and assemble qualitative data on public health policies and mask-wearing compliance

Rationale. The knowledge of average positive COVID-19 case data provides little information without the context of public health policies and mask wearing compliance at the schools involved in this exploration. Regional views on face coverings policies could be attributed to cultural views or simply a lack of COVID-19 cases which renders face coverings unnecessary. These preferences are necessary to understanding a regional group's predisposition to following face covering policies and more general COVID-19 guidelines as a whole. Additional context needed to understand COVID-19 transmission is the type of public health policies implemented at the schools in this study. This information will give a backdrop to how COVID-19 transmission has been slowed or how it has been allowed to spread. Once these policies have been identified for the schools in this study, their specific restrictions can be compared with average positive case data from **Aim 1** to determine which policies were the most effective in limiting disease transmission.

Methods. To acquire the information needed to provide context to our results from **Aim 1**, mask wearing compliance data will be qualitatively extracted from the New York Times' collection of regional mask wearing tendencies. Likelihood of wearing a face covering in public is provided at the county-level for the United States based on a survey of 250,000 respondents conducted in July (Times, 2020). These tendencies will be aggregated to represent the average likelihood to wear a face

covering within a certain state and then will be combined again to represent preferences within an entire geographic region. This information will provide context to how COVID-19 transmission differs based on regional views on masks when combined with our results during **Aim 3**. Assembling public health policies for the schools in this study involves a thorough review of school websites and announcements on their COVID-19 policies. Once these policies are assembled for schools in this study, they will be grouped into similar tactics to simplify our analysis in **Aim 3**. The COVID-19 policies for UVA will also be assembled to judge their effectiveness on mitigating transmission. The search for preventative measures at schools in this study can only begin after the random selection of schools occurs at the end of the Fall 2020 semester.

Expected Results. This aspect of the study focuses more on acquiring qualitative data so there are no quantitative results to anticipate. However, when aggregating the mask compliance data, one can expect to see trends of more masks being worn in the Northeast compared to the Rocky Mountain Western or Southeastern regions. Additionally, one may expect to find stricter COVID-19 policies in the Northeast and Midwest due to the number of cases and high mask compliance.

Limitations. The major limitation for **Aim 2** is that our data on mask compliance is restricted to only qualitative likelihoods of wearing a face covering from a survey. Responses from a survey have an inherent bias based on the types of people who choose to and have the time to fill out a survey. Additionally, survey responses can be affected by social desirability bias in which respondents will adjust their answer to be more socially acceptable. This survey was conducted in July 2020 so it is possible that mask compliance tendencies have changed dramatically since then which will not be reflected in this study. The accumulation of public health policies from schools in this study will be limited to only what information is available to the public.

Aim 3: Compare the results of the analyses in Aim 1 with that of Aim 2 to form actionable recommendations on public health at higher education institutions

Rationale. The combination of average positive case data at U.S. colleges and universities with their respective health policies and likelihood of mask compliance will provide the most comprehensive picture of COVID-19 transmission at higher education institutions. This correlative study will reveal which preventative measures implemented by colleges have been the most effective in limiting disease transmission throughout the Fall 2020 semester. By assembling these insights into a series of actionable recommendations on the effectiveness of COVID-19 policies, **Aim 3** will produce a thorough review of how to best limit disease transmission on college campuses. This information is undoubtedly relevant to monitoring the current pandemic, but can also be applied to future disease outbreaks. Knowledge of the most effective preventative measures against a transmissible disease will inform policymakers in the future to more rapidly identify a successful strategy to protect a community from a disease.

Methods. This final portion of the study aims to connect the results of **Aim 1** and **Aim 2** to examine the trends in COVID-19 transmission in response to both regional mask compliance and individual health policies at the schools in this study. To perform this final analysis, the regional averages of COVID-19 positive case data will be compared with regional mask compliance tendencies to determine how they overlap. Specifically, we will look for any unexpected combinations of lower average positive cases in regions with low mask compliance tendencies or vice versa to identify any significant differences. The purpose for identifying these differences is that a significant gap in positive cases on college campuses versus regional mask compliance tendencies might indicate that the difference is attributable to health policies at these schools. The next part of this study involves a correlative analysis between the results of **Aim 1** and the types of health policies at each school from **Aim 2**. To tackle this analysis, the first step will be to identify the top policies from groups of schools comprised of same type from the same geographic region. These groups will be the same as those used in **Aim 1** to test for regional differences between public and private institutions. The top policies

from these groups will be compared to identify significant differences in positive case averages that can be attributed to different COVID-19 mitigation strategies. Additionally, the policies implemented at UVA will be compared to the positive case data in a more in depth analysis to determine if the implementation of new policies had an effect on positive case numbers. The results from this analysis on UVA's policies will be compared to the results of the same analysis between small groups schools and their top policies to identify the success or failure of measures used at UVA versus the larger network of higher education institutes. Once these comparisons have all been performed, actionable recommendations will be formed on the efficacy of COVID-19 policies at college campuses. The final product of this study will consist of the results of our analyses on positive case data from **Aim 1** combined with insights and recommendations from **Aim 3** to produce a comprehensive report on the optimal way to mitigate disease transmission at higher education institutions.

Expected Results. One can expect to see an association between low mask compliance and higher positive case numbers based on the knowledge that masks reduce person-to-person transmission of COVID-19. Additionally, schools with stricter health policies are expected to have lower positive case numbers if their prevention methods are effective.

Limitations. The main limitation of **Aim 3** is that the comparison of health policies with positive case data is restricted only to the health policies that are in effect and published at the end of the Fall 2020 semester. This will reduce the team's ability to make conclusions on the efficacy of policies in place at the beginning of the semester.

III. REFERENCES

Times, T. N. Y. (2020, July 20). Covid in the U.S.: Latest Map and Case Count. *The New York Times*. https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html

CDC. (2020, February 11). *Coronavirus Disease 2019 (COVID-19)*. Centers for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/faq.html

Gonzalez, R., Curtis, K., Bivins, A., Bibby, K., Weir, M. H., Yetka, K., Thompson, H., Keeling, D., Mitchell, J., & Gonzalez, D. (2020). COVID-19 surveillance in Southeastern Virginia using wastewater-based epidemiology. *Water research*, *186*, 116296. Advance online publication. https://doi.org/10.1016/j.watres.2020.116296