"The Implementation of Non Pharmaceutical Interventions (NPIs) in smaller to large communities and its relation to R0 and R(t) during H1N1 pandemic 2009"

By

Sahar Hashmi

M.D. Doctor of Medicine

Submitted to the System Design and Management Program in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Engineering and Management

at the

Massachusetts Institute of Technology

November 2010

© 2010 Sahar Hashmi
All rights reserved

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author

Sahar Hashmi
System Design and Management Program
November 2010

Certified by

Dr. Stan Finkelstein
Thesis Supervisor
Engineering Systems Division

Certified by

Professor Richard Larson
Thesis Supervisor
Engineering Systems Division

Accepted by

Patrick Hale
Director
System Design & Management Program
# Table of Contents

Abstract ........................................................................................................................................

Acknowledgements .....................................................................................................................

Chapter 1: Introduction ..............................................................................................................

Chapter 2: The overview of previous NPI Studies: .................................................................

  Computer Model .....................................................................................................................

  Laboratory Studies ................................................................................................................

  Observational Studies ..........................................................................................................

  Historical Studies ................................................................................................................

  Miscellaneous Studies ...........................................................................................................

Chapter 3: NPIs and R0 and R(t) in La Gloria and Mexico: .....................................................

Chapter 4: Google Insights and R0 .........................................................................................

Chapter 6: Conclusion ..............................................................................................................

References ................................................................................................................................
“The Implementation of NPIs in smaller to large communities and its relation to R0 and R(t) during H1N1 pandemic 2009”

by

Sahar Hashmi

M.D. Doctor of Medicine

Submitted to the System Design and Management Program in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Engineering and Management

at the

Massachusetts Institute of Technology
Abstract

This thesis focuses on the use of non-pharmaceutical interventions (NPIs) during the time of the 2009 H1N1 pandemic and its possible relation to R0 and R(t). R0 is defined as the mean number of people that a newly infected person will subsequently infect in a completely susceptible population whereas R(t) is the average number of new infections by an infectious individual at time t. R0 is important for understanding the severity of an influenza outbreak while R(t) is a necessary tool to measure the progression of infection rate over time. A high R0 value (more than 2) generally corresponds to a more serious outbreak. This thesis discusses a town in Mexico named La Gloria, which is thought to be the place where the H1N1 pandemic started, and the subsequent implementation of NPIs in Mexico City as the virus spread and people became aware of its novelty. An evaluation of Mexico’s response to H1N1 suggests that the emphasis on the use of NPIs may have related to a decreasing R0 value. Further investigation of this relationship using news articles and Google Insights also shows interesting potential correlations. In short my thesis focuses on the possible relationship between R0s and NPIs in a pandemic setting.

Thesis Supervisor: Dr.Stan Finkelsetein and Prof.Richard Larson

Title: Professors at the Harvard Medical School/Engineering Systems Division of MIT
Acknowledgements

It is a pleasure to thank the many people who made this thesis possible.

It is difficult to overstate my gratitude to my thesis supervisors, Dr. Stan Finkelstein and Dr. Richard Larson. Throughout my research time, they provided encouragement, sound advice and guidance, good teaching, good company, lots of good ideas and support of my education. I consider myself lucky to have found this magnificent intellectual team.

I would like to especially thank Pat Hale for his support, advice and mentorship for not only my work but throughout the SDM program and his help to make my experience at SDM a very pleasant and enjoyable one. In particular, I would like to thank Bill Foley and the whole SDM staff for their endless support during my stay at the SDM program. I would like to thank Anna Teytelman enormously for her help in calculating the R0 for La Gloria using Markov models. I would like to also thank Professor Tivnan for helping me understand the exploratory statistics of my data. I want to thank Julia Hopkins for her help in gathering data from Google insight and to plot the graphs. In addition, I want to especially thank Julia and Kallie for their feedback on my thesis. I, of course, owe gratitude to everyone on the flu-team whose endless insight, feedback and valuable work has helped support my thesis.

I would like to express my deep appreciation to all of my sisters for their support, encouragement and love. Lastly, and most importantly, I wish to thank my parents, Dr. Nasim Hashmi and Aisha Hashmi. It is their upbringing and continuous support that has helped me to reach this far.
Chapter 1: Introduction:

Non-pharmaceutical interventions (NPIs) play an important role in mitigating the severity of a pandemic. While vaccines and the administration of antiviral are also considered effective actions against a pandemic, they take time and money to produce. NPIs, on the other hand, can be more easily executed. NPIs include practices which can either be implemented by individuals, such as hand washing, cough etiquette, wearing face masks, using hand sanitizers, and social distancing, or practices which can be enforced by governments, such as quarantine and school closure during an outbreak.

Based on their accessibility alone, NPIs have the potential to be very effective tools in mitigating and impeding the outbreak of pandemics. This thesis begins by presenting numerous papers which have attempted to quantify the effectiveness of NPIs at mitigating the effects of a pandemic through several study methodologies. These methodologies vary in the ways in which they model a pandemic and quantify the impact of an NPI. Currently, there is no single metric for the effectiveness of an NPI. Recently, however, the basic reproductive number, R0, has been suggested by Dr. Richard Larson of the Massachusetts Institute of Technology to serve as a quantitative metric to determine NPI impact.
R0 represents the average number of individuals infected by one single individual during a pandemic when the population is entirely susceptible. While it is generally treated as constant in a given epidemic outbreak, recent research by Dr. Larson has shown that R0 can change over the course of a pandemic. His research deviates from the majority of the published literature by introducing R0 as a product of two important factors: the frequency of contact with an infected person (denoted as $\lambda$) and the probability of contracting the disease from any given contact (denoted as $p$).\footnote{1} With this definition, R0 can be seen to be dependent on the specific population being modeled. It is no longer just a constant parameter that must be arbitrarily assigned for different runs of the same model. Either of these variables could change in a pandemic situation, such as when susceptible members of a population begin implementing NPIs.

This definition of R0 has yet to be widely used as the basis of a method for determining the effect of non-pharmaceutical interventions on a pandemic. Thus this thesis seeks to inform and support establish such a basis by examining the relationship between NPI use and R0 based on $R0 = \lambda p$. Such a relationship could prove to be very informative in formulating policies and practices.
Chapter 2: The overview of previous NPI Studies

In order to determine the appropriateness of using R0 as a means to measure NPI effectiveness, a review of past metrics used for measuring NPIs was conducted.

Laboratory Studies

*In vitro* studies encompass all studies of NPIs which can be performed in a laboratory under controlled, carefully monitored conditions. These studies are the most straightforward method of testing an NPI on the small scale. As a result of their scale, *in vitro* studies are limited to testing the NPIs of hygiene. For example, one study demonstrated the effectiveness of face masks and respirators on reducing viral transmissions in a lab. This study instructed influenza patients to cough onto petri-dishes wearing either a face mask, a respirator, or coughing without any physical barrier as a negative control. The amount of infectious particles collected was used to quantify the effectiveness of prevention. It was found that face masks and respirators reduced the amount of virus particles that a cough released into the air, since the petri-dishes which were inoculated while the patient was wearing either face covering did not grow any virus.

Similar in scale, another experiment tested the effectiveness of hand hygiene on reducing the spread of viruses. Participants in this particular study were given Norwalk viruses on their hands and then treated with various hand sanitation agents (e.g. hand soap, sanitizer, etc). The amount of virus on the hands both before and after the treatments was measured with RT-PCR (real-time polymerase chain reactions) in all cases. It was found that hand soap was most effective in reducing virus particles, while hand sanitizer did not significantly reduce the number of virus particles on a hand.
One of the greatest limitations with *in vitro* studies is that the environment of the experiment is controlled, and thus non-endogenous to what is noted in a pandemic setting. The small scale of these experiments allows them to be controlled, which results in reasonable quantitative measurements, but it also makes them difficult to scale up to a pandemic scenario. Furthermore, it is difficult to incorporate R0 values, which are essential to NPI studies, into the experimental results. Regardless, these studies offer concrete evidence that hygienic NPIs have the potential to reduce the presence of bacteria on human hands and in the air; a focused view of the effectiveness of certain NPIs on a reducing the impact of a potential pandemic scenario.

**Observational Studies**

Observational studies expand the scale of laboratory studies to experiments which monitor NPI use in daily life. Unfortunately, the environment and scale of these studies makes it difficult to clearly define, control, or quantify experimental variables. However, these studies are the closest a researcher can get to monitoring pandemics in a natural setting. As such, they are often subject to criticism due to their non-controlled environments. In a case study researching the effect of NPIs on the infection rate among students during the H1N1 pandemic in 2009, researchers assigned different student dormitories either an NPI of face masks and hand washing or each NPI individually. It was found that a combination of the NPIs significantly reduced infection rates while the individual practice of NPIs did not.
In a similar study, a researcher attempted to quantify the effect of NPIs on reducing infection rates among college students. Students were similarly subdivided with one group implementing a combination of the NPIs of hand hygiene and coughing etiquette while other groups implemented only one of the NPIs. Students were asked to complete a weekly online survey to monitor their use of the NPIs. This study, in contrast to those mentioned previously, concluded that NPIs had no effect on reducing the infection rate of influenza. However, it is important to consider that the results of this study were based solely on subjects filling out online surveys, a method prone to bias or misinformation.

In an effort to qualify the accuracy of self-reporting in a pandemic setting, one study experimented with surveys intended to monitor the habits of health care workers and another study tested self-reporting surveys on NPI practices utilized in Hong Kong during the influenza season of 2008. The results of these studies were inconclusive in terms of NPI effectiveness on pandemic scenarios, but they further emphasize the need for an unbiased and quantifiable means of measuring NPI effectiveness.
Computer Models

Computer models can achieve the largest scales for testing a pandemic scenario. As a result, they are one of the most prevalent techniques used to study NPIs. They can be used for modeling numerous pandemic situations with simulations capable of yielding entire epidemic curves and a complete record of infection rates for a given, specific scenario based on a set of pre-defined parameters. Unlike observational studies, computer models are programmed with specific variables in mind, and as a result are not subject to outside interventions which could skew the results. However, since each model deliberately explores a finite number of parameters in order to be simple enough to construct, the limitation of these simulations is that they are often unrealistic. Recent findings have attempted to rectify this issue.

One study utilized census data of the United States and Great Britain to simulate an influenza pandemic. The infection sites from which the researchers selected and collected census data from for this study were schools, social events and work places. The model used the distance between susceptible and infected individuals as a means of determining the likelihood of infection. International travel was also considered in calculating the outbreak of influenza. Using an individual-based simulation, the researchers were able to set pandemic conditions, such as an R0 value, and implement a combination of NPIs. They examined the resulting epidemic curve (total number infected, number infected at peak) to determine how effective these NPIs were in mitigating the impact of a pandemic. This study found that NPIs such as school closure yielded lower epidemic peaks, and that the early implementation of quarantine was effective in preventing the spread of infection as it resulted in a reduction of both cumulative attack rate (usually around 4%) and in at least a week delay of the epidemic peak. This study also concluded that border control did not impede cross-country outbreaks unless implemented with near-perfect
efficiency. Although this study indicated how certain NPIs could mitigate pandemic outbreaks, the model utilized pre-determined R0 values that did not vary as each scenario progressed, despite the fact that this variance occurs in real pandemics. This study treated R0 as a constant parameter of a pandemic. Several other computer models presented below have done the same.

When researchers simulated the effect of closing schools during an influenza outbreak in France, they also considered R0 as a constant. Their model was constructed using surveillance data of influenza rates in France and the timing of school holidays. The results of this model indicated that holiday school closures reduced the infection rate of influenza by 13-17%; the data were then extrapolated to simulate prolonged school closure in the event of a pandemic, an act which appeared to inversely correlate to lower infection rates. One limitation of this model was that it considered only a simplified school system to determine NPI effectiveness and therefore could not fully simulate a pandemic scenario, which impacts more than just schools. Despite the large scale of a computer simulation, the scope of the model must often be reduced to make the model programmable and its results comprehensible.

In a more complex approach, the effect of school closures on the H1N1 epidemic of Allegheny County in America was evaluated using a compartmental SEIR (susceptible, exposed, infected, recovered) model. The results of this study were not as straightforward as the results of the French study described previously. It was found that school closures that exceeded 8 weeks in length only slightly affected the epidemic curve; they pushed the peak back a week at most. Furthermore, there was no indication that school closures decreased the total number of individuals infected. The disparity between the French and Allegheny County studies is evidence of further limitations of computer models. While both studies found the NPI of school closure to be important, they differed in the quantifiable details of how this NPI should be implemented.
The different construction and quantification of the models (individual-based vs. compartmental, infection rate vs. epidemic curve) resulted in conclusions about the impact of school closure that do not match. It should be noted that each study also had to use constant R0 values for their models. If they had instead assumed that R0 is context specific, their results would likely have resembled a real pandemic scenario more and, as a result, the validity of each model may have been easier to determine.

In a different approach, an SEIR model was utilized to predict the spread of the H1N1 pandemic in Germany.\textsuperscript{11} This study attempted to determine the effects of combining the NPI of isolation and the administration of antiviral drugs on the influenza epidemic. For their model, the researchers utilized R0 values from the H1N1 pandemic in the United States and took into consideration varying infection rates among different age groups. The model implied that intervention techniques helped alleviate pandemics, even though it was never tested using the NPI of isolation alone.

In a similar study, the combined effects of NPIs and antiviral treatments on influenza outbreaks were researched using an agent-based model.\textsuperscript{12} The important difference between this study and the above German study was that this study examined the individual effects of NPIs and antiviral treatments as well as their effects in concert. The results of this study are slightly different from those of previous models. It indicated that the quarantine of infected individuals was not effective in impeding a pandemic outbreak; rather a combination of NPIs such as social distancing and school closures with antiviral treatments was more effective based on the percent of the population infected after the model was run. Interestingly, two other studies, one which utilized InfluenzaSim\textsuperscript{13}, and the other which created FluTE\textsuperscript{14}, two open source internet flu simulators, published similar conclusions.
The effects of social distancing have also been studied alone with computer models. In one such study, researchers attempted to quantify these effects with constant R0 values using a model that simulated a less-than-realistic approach by constructing a population in Australia from state census data and assigning the population a set number of individuals, then using a finite number of factors (age, R0 value) to determine how many people were infected per one cycle of the model. The researchers measured the impact of NPIs using the change in the daily and final attack rates in a pandemic scenario. This model indicated that beyond a certain R0 threshold value of 3.5, social distancing became ineffective.

It is evident from these studies that a standard for measuring NPI effectiveness is needed, as the lack of one makes it difficult to gauge the reliability of the results of computer simulations. Some models used attack rates to measure the severity of a pandemic while others studied the peak of the epidemic curve and more looked at the total number infected by the end of the pandemic. None of the models presented thus far considered the relationship between NPI use and the traditional measure of the severity of a pandemic, R0; each makes the same basic assumption that R0 is constant throughout a pandemic. Despite this, recent research has shown that R0 serves as an ideal metric for the success of an NPI.

Dr. Larson began working on a novel method of modeling a pandemic and testing the effectiveness of the NPI of social distancing as early as 2007. He used \( R0 = \lambda p \) to construct several populations within his initial model; each population had a different R0 based on the frequency of social contact and infectivity. With this distinction, the model revealed that social distancing was very effective at reducing the number of infected individuals in a pandemic when implemented by the group with the highest contact frequency. It also was seen to impact the infection rates of the groups with lower contact frequency, if somewhat less dramatically.
In addition, the model revealed that people with high contact frequency often dominate the results of a model that treats all of its subjects the same regardless of their activity levels in a community.

The usefulness of such results lies in policy implications; once the determinants of R0 are made clear, NPIs such as social distancing can be tailored to reduce these determinants and thus the value of R0 for a given population. This model was expanded to study several communities at once, accounting for the implementation of NPIs such as travel restrictions and hygiene in addition to social distancing. The multi-community approach allowed for a more thorough exploration of the impact of NPIs on a more realistic pandemic scenario, in which infections can be transmitted between communities through as little as one infected individual. The results from this more complex model revealed a clear distinction in the effectiveness of individual NPIs; social distancing and hygienic interventions were found to be much better at decreasing the impact of a pandemic than travel restrictions.

The usefulness of the concept of $R_0 = \lambda p$ has been further investigated by Karima Nigmatulina, a former PhD student at the Massachusetts Institute of Technology. In her thesis, she explored the role of R0 in a pandemic in a model similar to those presented above: a compartmental model of several communities which account for varied activity levels of individual members. She expanded on the concept of R0 as it relates to the time-dependent reproductive rate $R(t)$, noting how the latter is less than R0 and changes as more people become infected and subsequently immune to the pandemic virus.
This distinction is important in considerations of R0 calculations. R0 is a contextual value, while R(t) is entirely dependent on R0 and the number of people infected over time. Correctly calculating R0, therefore, is necessary for the determination of R(t), the progression of the reproductive rate as less of the population is susceptible to a given pandemic virus.

The model used by Nigmatulina deviates from previous compartmental models in that it accounts for the inconsistent behavior of individuals. In addition, the model was shown to be applicable to scenarios in which vaccines are available at a given time in a pandemic. The conclusions were akin to that of the previous papers which used similar models; reducing the rate of contact reduces R0 and therefore reduces the impact of a pandemic on a given community. We note that the versatility of this model was due in large part to the fact that it discounted assumptions that several previously formulated models made, such as that of constant R0 values. As a result, these three papers which used various forms of this compartmental model were able to quantitatively measure the impact of NPIs on pandemic scenarios, measurements which suggest important health policy implications.
Historical Studies

The 1918 influenza pandemic was the most deadly pandemic to affect mankind in recent history. Much of the data collected during this time period are currently being used in studies to determine what NPIs were implemented and the impact they may have had on the spread of the pandemic. In one such study, data from the various statistical and newspaper reports of the outbreak of the 1918 virus in Sydney, Australia were used to make a mathematical model that attempted to study and explain the variability noted between the city’s three waves of influenza outbreaks. The standard assumption about the waves of influenza had been that they were caused by a change in the infectivity of the virus. This model found that the variability in the outbreaks did not result from changes in the infectivity of the virus, and thus the researchers hypothesized that it was the inconsistent practices of social distancing that resulted in the variability between the three outbreaks. Their model, when compared with historical infection rates and epidemic peaks, was found to resemble the 1918 outbreaks closely by keeping the infectivity of the virus constant and varying the practice of social distancing.

In a study conducted in the United States of the 1918 pandemic’s effect on America, researchers utilized mortality data and public health data to determine the impact of NPIs on the infection and mortality rates of various cities. The study surveyed 43 cities for the NPIs of school closure, isolation and quarantine, and cancellations of public gatherings. They were interested in the quantitative measure of NPI effectiveness, which was determined based on the correlation between death rates and NPI use in each of the cities. It was found that fewer deaths were reported in cities where more NPIs were implemented.
Previous findings on NPI use in historical, deadly pandemics such as the 1918 Influenza Pandemic point to the possible benefits from NPI implementation in a pandemic scenario. However, it is worth quantifying such data with a set metric and standard, such as R0, to more concretely calculate effectiveness.

**Miscellaneous Studies**

In an independent study that set out to balance the economical costs of implementing NPIs with pandemic mortality rates, researchers utilized a mathematical model that was similar to the SEIR model, but also accounted for mortality. This study found that it is necessary to immediately implement NPIs during a pandemic for maximum effectiveness, measured by the number of work days per person lost vs. the death rate. Another study combined a literature review of NPIs with a questionnaire given to experts in the field to determine the effectiveness of specific NPIs. Due to the paucity of data in the literature review, this study focused on the survey and found that experts agreed on the effectiveness of self-isolation and hand hygiene, while they opposed the use of face masks and forced social distancing in decreasing the infection rate of a pandemic.

Another study monitored the impact of the 2009 H1N1 Pandemic in Hong Kong. It attempted to measure the effectiveness of NPIs in controlling the outbreak of the pandemic using data collected from Hong Kong school records. The data was used to create a model that compared school closures to the number of reported H1N1 cases. These researchers found that school closures resulted in a 25% decrease in flu transmission. The transmission factor is similar to the basic reproductive number, R0, though it is calculated differently.
The final study considered in this thesis did not relate the practice of NPIs to R0. However, it did explore the calculation of R0 for the entirety of the United States during the first wave of the outbreak. The study was working with imprecise data, for it was conducted as the pandemic was progressing, and as a result it did several sensitivity analyses to its data to determine what assumptions would alter the calculation of R0 the most. Use of the method adopted by this study to calculate R0, a likelihood-based method, resulted in several different R0 values which highly depended on assumptions about the serial interval and underreporting of number of cases. This study serves as a reminder that good data is essential for accurate measurements of R0. And as we have seen, good data is especially difficult to find or quantify during a pandemic scenario.

In this review of previous research on the effectiveness of NPIs on a pandemic scenario, several patterns can be readily identified. Most of the studies support the notion that NPIs can reduce infection rates or otherwise combat a pandemic. They differ on the specifics of which NPIs are the most effective and the effectiveness of a single NPI, but the general trend is clear. In addition, most of the studies consider R0 to be a constant factor in a pandemic, an assumption that Dr. Larson has shown to be inaccurate. By treating R0 as a variable in a pandemic scenario, more specific information can be gathered from computer and other mathematical models. It is clear that NPIs are likely favorable in a pandemic scenario, but more research must be conducted in order to make this conclusion applicable to current pandemic health policies.
Chapter 3: NPIs and R0 in La Gloria and Mexico

In the spring of 2009, an influenza outbreak caused by a novel Type A(H1N1) virus called the “swine flu” attracted considerable news media attention. Much of this attention was likely derived from the world health community’s earlier preparation for a global pandemic of H5N1, the “bird flu” from Southeast Asia with a reported death rate of 60%. The novel swine flu virus was first identified in La Gloria, Mexico on April 23, 2009 and shortly thereafter it was also reported in the United States, Canada, and the rest of the world. However, there had been speculation that the first cases actually appeared in California in late March.1,2

The incidence of this infectious disease soon caused the World Health Organization (WHO) to declare the outbreak to be a worldwide pandemic.

In Mexico, initial estimates of about 2,400 cases of H1N1 and 150 related deaths caused Mexican officials to report a flu death rate of around 6% (150/2400 = .0625), a rate close to the 8-10% death rate of SARS and about three times higher than the worst-known pandemic influenza, the “Spanish Flu” of 1918-1919.3,4 This estimation of 6% was later found to be much higher than the actual rate.3 The initial estimate did not account for all of the mild cases of H1N1 in Mexico at the time, a number which no one reliably knew and which caused the denominator in the death rate calculation to be too small. In addition, both the numerator and the denominator accounted for not only confirmed but also ‘probable’ or ‘suspected’ cases of H1N1.3 However, once news of the misleadingly high death rate spread, a world poised for a deadly pandemic flu was spurred to action, driven by memories of the horrors of the Spanish Flu and SARS.
Now with the world watching, by late April testing became more rigorous and more reliable data were gathered. Confirmed cases were reported initially to the WHO and the Centers for Disease Control and Prevention (CDC). However, after the initial scare, testing for the virus became sporadic and the accuracy of the number of reported cases questionable. The widespread lack of materials and expertise to correctly diagnose H1N1 influenza caused many cases of the flu to go undetected. By July, doctors were more inclined to send sick individuals with influenza-like illness (ILI) home to rest without confirming the presence of H1N1 unless the persons were considered "high risk" due to age, pregnancy or pre-existing medical conditions.

As a result, the most accurate reports for confirmed H1N1 came between April and mid-July. These reports have proven to be the best evidence for understanding the epidemic curve of H1N1.

The following reviews a set of events that took place in La Gloria and, more broadly, Mexico during the early days of the H1N1 outbreak. We are particularly interested in understanding if and when non-pharmaceutical interventions (NPIs) played a role in mitigating the spread of the infectious disease. We rely on what we believe to be the accurate reports of H1N1 cases from April to mid-July to track the spread of the virus. In order to track the progression of cases of swine flu during the months of the first outbreak of H1N1, we used the basic reproductive number R0.
As discussed earlier in this thesis, the measure of R0 is contextual, and therefore highly dependent on population density, hygienic behaviors, and other factors which impact the number of infected people in a given area. Consequently, we can use R0 to compare two different regions and can gain insight into what factors in the region influence its value. In addition, if an infection curve changes over the course of a pandemic, R(t) is a tool to measure that change. By definition, R(t) is the average number of new infections by an infectious individual at time t.

With this in mind, R0 was analyzed in conjunction with news articles from the same time period in order to understand its response to the emphasis on non-pharmaceutical interventions during the early days of the pandemic.

We begin our analysis by examining a village named La Gloria, located in the state of Veracruz, in the Perote valley of Mexico. This village reported the first confirmed case of H1N1, and supporting evidence from news articles and interviews with members of the community suggest that the H1N1 outbreak actually originated there.
La Gloria, Mexico.


La Gloria:

A qualitative study of H1N1 in La Gloria affords a description of how the influenza virus spread through its population and then went on to infect residents of Mexico City. La Gloria has a population of 2,243 individuals and is flanked by both mountains and a collection of pig farms belonging to the United States pork company, Smithfield Farms. These farms raised about a million pigs each as of 2008. The community of La Gloria believes that the unhygienic conditions of the closest farm five miles away, Granjas Carroll de México, led to the start of the H1N1 pandemic. This claim has yet to be substantiated by rigorous testing of the farm’s animals and workers.
From February to early April of 2009, many members of the community of La Gloria fell ill with respiratory illnesses, exhibiting flu-like symptoms such as high fevers and chills, body aches and sore throats.\textsuperscript{7,8,9} As of the end of April, the infection had spread to about 60% of the La Gloria community (about 1,350 individuals).\textsuperscript{7,10,11,12} By this time, 218 individuals in Mexico City were also reported to be showing symptoms of respiratory illness. In addition, H1N1 was suspected as the cause of the deaths of two people in La Gloria and five in Mexico City.\textsuperscript{13}

The spread of the virus from La Gloria to Mexico City is readily explainable. Reports of infections date back to February 18, 2009 in La Gloria.\textsuperscript{14,15} Approximately half of the residents of La Gloria work in Mexico City; their commute via bus is an ideal vehicle for the spread of the flu.\textsuperscript{11} From February to the beginning of April, those with work in Mexico City continued to commute from La Gloria despite the outbreak of illness in their city. Some left at the beginning of the week, others left later in the week. Residents typically remained several days in the city before returning home.\textsuperscript{11} However, this trend altered during the week before Easter, which fell on April 12, 2009. Most La Gloria residents and relatives working in Mexico City came back to La Gloria to celebrate the holiday.\textsuperscript{15} When Easter was over, many of them returned to work, creating a concentration of commuters who had been in contact with H1N1, heading towards Mexico City (see Figure 2). Not surprisingly, the peak of swine flu in Mexico occurred during the final days of April.\textsuperscript{16}
The first death confirmed to have been caused by H1N1 could be attributed to this commuting network, as well. Maria Adela Gutierrez from Oaxaca, Mexico was partnered with a coworker in Mexico City who reportedly had a persistent cough and who was from the Perote region where the swine flu was first discovered. Gutierrez fell severely ill with pneumonia-like symptoms shortly after working with this person and died on April 13, 2009.17

Methodology:

In villages like La Gloria with poor documentation of the early H1N1 cases, it is difficult to know the progression of an epidemic curve in order to calculate R0. However, we can provide a reasonable estimation of R0 by using backwards induction on the total number of suspected H1N1 cases incurred by La Gloria residents during the spring of 2009. Since H1N1 was not diagnosed until well into the outbreak in La Gloria, we assume that behavioral patterns, in terms of human to human contacts and illness prevention practices (which both give rise to the value for R0) remained unchanged in La Gloria before herd immunity was reached in the population. Under this assumption of behavioral constancy, we calculated a baseline estimate of R0, with the help of Anna et al, a simple probabilistic Markov Chain model.

The model considers La Gloria as an isolated community consisting of identical, independent individuals who spread the flu within discrete generation time periods. It designates one of the members of the community as a “patient zero” with the first case of the infection. Patient zero has numerous infectious contacts who in turn contribute to the spread of the flu.
In any given generation, a person in the community is designated as susceptible (no prior immunity or infection), infectious, or immune. Individuals are infectious for one generation. During that time, they have $R_0$ numbers of contacts with random, potentially susceptible people in the community. If a contact is already infected or immune, there is no newly infected person, but if the contact is susceptible, he or she becomes infected. Once the $R_0$ contacts are picked and delineated as either newly infected or not, one discrete time period ends and all infected persons become immune. The newly infectious people individually repeat the same process with the same average $R_0$. This continues until no new infections arise, either because everyone in the community is already immune or because the infected for a generation did not come in contact with any susceptible members of the community.

We found the distribution of the total number of infectious people in the community by implementing the Markov Chain model, in which the state space describes the number of infectious and immune individuals. We estimated the value of $R_0$ in La Gloria by applying the model to a community with the same characteristics as La Gloria; we ran different possible values of $R_0$ until we saw a distribution of 60% infected at the end of the run. Anna et al, a doctoral student at ESD, were able to calculate $R_0$ for La Gloria using the 60% estimate of infection rate at that time.
In order to gain insight into the use of NPIs in La Gloria, we interviewed physicians who worked in La Gloria at the time of the outbreak to get firsthand impressions of the state of the village during the pandemic. We also verified information found in several Spanish language news articles dealing with bus routes and commuter densities from La Gloria to Mexico City through interviews with the Autobuses de Oriente (ADO) bus company and ten Mexicans who lived in La Glory at the time (see Figure 2).

Results:

Data collected from Mexican newspapers suggests that illnesses due to what was later attributed to H1N1 may have started as early as February 2009 in Veracruz with the death of two babies. Other media accounts suggest that the illness infected 60% of the population of La Gloria by the end of March.\(^{7,10,11,12}\)

The pig farm closest to La Gloria, Granjas Carroll, is five miles north and its huge pits of pig waste and dead pigs cause massive fly infestations in La Gloria.\(^{7,8,9,10}\) Mexican officials orchestrated a spraying operation to kill the flies on April 6, 2009 as the initial infection control method.\(^{18}\) The government also placed a “health fence” around La Gloria on April 28th to monitor people leaving and entering the village for signs of the flu, hoping to catch those who exhibited symptoms before they could infect others in the village.\(^{12}\) These two NPIs were the only notable infection control strategies in La Gloria and were implemented after the worst of the outbreak had hit the village.
Our calculated R0 value, 1.64, is corroborated by other studies on the H1N1 pandemic in La Gloria. Fraser et al. estimated the value to be between 1.4-1.6 in the early month of April.¹⁹ Cruz Pacheco et al., a research group from Mexico, found that the R0 value for La Gloria was 1.716 during the H1N1 pandemic using the Kermack and McKendrick SIR model. Their approach took data from the Secretaría de Salud de México for April 10 to April 20, 2009, a time period which overlaps the weeks we chose to analyze.²⁰

Discussion:

Though early incidence of the flu suggests its origin in La Gloria, the widespread use of NPIs did not occur in this village until most of the flu had run its course. Dr. Armando Romeo Aguilar Cano, director of the Calidad de la Atención Clínica (the Quality Medical Care Clinic), M.D., worked in La Gloria during the outbreak and in his interview on April 12, 2010, he validated that almost no precautions in the form of NPIs were taken in La Gloria between February and April 2009 to stop the spread of the flu. The lack of NPIs during the outbreak likely contributed to the 60% of La Gloria that became infected. Once the first case of swine flu was confirmed, however, Dr. Cano corroborates that measures of isolation, quarantine and segregation, such as the “health fence,” were imposed on the population of La Gloria late in April 2009 (after 60% had become infected) and other individual measures were likely implemented by many in the population. These measures, unfortunately, came too late to significantly impact the epidemic curve of La Gloria.
Mexico

In contrast, the R0 values for the entirety of Mexico as a country were calculated using the exponential growth method, since the documentation of the daily number of cases was more consistent as the larger pandemic occurred after late April, when people were aware of the novelty of H1N1 and could test for it.

Data on the day-to-day progression of the disease came from the databases of the CDC, the WHO, and the European Center for Disease Prevention and Control (ECDC).\textsuperscript{21} The ECDC website provided daily situation reports on the H1N1 virus, whereas the WHO and CDC websites only gave updates every few days, so we worked primarily with data from the ECDC. Its reports provided an overview of the pandemic through epidemiological updates, recent publications, newly confirmed cases, and a regularly updated cumulative number of cases for selected countries, including Mexico.\textsuperscript{21}

To model the growth of the H1N1 pandemic over time, we used an exponential function $y = ae^{bx}$, where $x$ represents the number of days since the start of the pandemic and $y$ denotes cumulative number of cases. We imported discrete data points of the number of cases per day to Excel to calculate the best fit values of $a$ and $b$. From this, we computed the predicted number of cases taking into account the reported 2.5-day incubation period.\textsuperscript{22}
To find the growth rate, we computed the ratio of each day’s estimated number of cases to the previous day’s estimate. The uncertainty in this kind of analysis is that we assumed exponential growth for all data points, when in reality it cannot continue forever. Thus we also assumed that our R0 calculations were based on data occurring before the peak of the epidemic, making exponential growth a reasonable approximation.

To relate these calculated R0 values to the implementation of NPIs during a specific time period, we looked for indicators of increasing interest in NPIs. We used the HealthMap database and Google to find news articles from Mexico related to NPI recommendations and practices in Mexico and the United States. HealthMap is a website that collects daily reports on diseases from all over the internet. It marks the articles by country of origin and date; articles from Mexico about NPIs were easily tracked down on this website. However, HealthMap does not archive its articles. In order to retrieve more Mexican newspaper articles from Spring 2009, we used Google with search terms like “La Gloria H1N1” and “Granjas Carroll La Gloria.” Google allowed us to search within specific time periods for articles published during the pandemic outbreak in Mexico. Using both of these methods, we located a significant sample of the articles and news alerts about NPI use in Mexico published between the months of April and July 2009. The number of articles was counted for each time period for which we were able to estimate an R0 value and the specific NPIs of these articles were noted. This analysis allowed us to know which NPIs were written about and how many different articles were written about NPI usage during the H1N1 pandemic.
Results

R0 was estimated to be 2.168 for early April 2009 in the entirety of Mexico. R(t) later decreased to 1.7 by mid-July (see Table 1). The steady decrease in the value of R(t) in Mexico appears to be consistent with the government initiative to test and diagnose those exhibiting influenza-like symptoms. It is also consistent with the implementation of several NPIs in Mexico, notably the shutdown of Mexico City (see Table 2). Though it is not possible to know if NPIs such as hand hygiene and cough etiquette were implemented by the citizens of Mexico, the emphasis on their implementation, as evident in news articles within months of the initial outbreak, is assumed to have encouraged the general trend of NPI use among the population.

Discussion:

A glance at the more widespread emphasis on NPIs in Mexico after the experience in La Gloria reveals that self-isolation, social distancing and school closures were the only methods emphasized during the beginning of the pandemic in April. Mexico City was then almost shut down by the beginning of May. From May onwards, Mexican newspapers and government reports emphasized hand washing, coughing etiquette, and social distancing (see Table 2). In conjunction with the focus on NPIs and in the absence of any vaccines or other medical interventions, R(t) decreased very quickly and steadily during this time period (see Figure 1). The negative trend or relationship between NPIs and R(t) in Mexico is consistent with the hypothesis that NPIs decreased disease-generating contacts and thus both R0 and R(t) for the population of Mexico during the time period studied.
Our calculated R0 value for Mexico (see Table 1) are supported by other studies such as that of Boelle et al., who suggest the R0 for Mexico ranged from 2.2 to 3.1.\textsuperscript{23} Cruz Pacheco et al. also studied the spread of the pandemic in Mexico City with and without various sanitary control measures, using the Kermack and McKendrick mathematical model as they did with La Gloria. From this, they estimate an R0 value of 1.72 for Mexico City during the early weeks of the pandemic.\textsuperscript{20}

**Conclusion:**

The R0 value we calculated for La Gloria and Mexico is consistent with what has been reported in other research studies of the 2009 H1N1 Pandemic. Hence we believe our calculation method produced a reliable trend in R(t) values over time for Mexico.

The value calculated for La Gloria provides an excellent baseline case for the spread of H1N1. By the time the world had focused on La Gloria, herd immunity had already been achieved in the town. The residents themselves had no idea that 60% of their population had fallen ill with H1N1 until after the fact. As such, there was no attempt to implement non-pharmaceutical interventions or acquire antiviral medication for those who were sick during the outbreak. The only NPI implementation, the health fence surrounding the town and the spraying of flies, most likely occurred after herd immunity had been reached and the worst of the outbreak had passed. By that time, however, the virus had migrated to the rest of Mexico and was well on its way to becoming a global pandemic.
The spread of H1N1 cases from La Gloria to Mexico City and beyond appears to be readily explainable by travel patterns between the two. Many residents of La Gloria commute to work in Mexico City regularly throughout a given week. When Easter arrived in the middle of April 2009, the majority of the commuters headed back home to spend time with their families. They then departed *en masse* to Mexico City after the holidays. H1N1 likely followed these commuters back to their workplaces in Mexico City, where the outbreak of swine flu was garnering much more media attention and, as a result, a greater impetus to use NPIs. Three weeks after Easter, Mexico City effectively shut down for five days to prevent further spread of the H1N1 virus.

There appears to be a trend between the observed speed of $R(t)$ decline for Mexico and increasing NPI use. $R(t)$ decreased significantly faster over time then it would have in a natural progression of the disease. The government and individual citizens of Mexico focused more on NPIs such as hand hygiene and social distancing. The evidence for this comes from several newspapers published at the time, detailing both government actions and popular opinion about the use of NPIs during the pandemic.

With this in mind, it appears that the implementation of NPIs early on in a pandemic can best help decrease its infection rate and its potential to become a serious health threat. As a result, recognition of a pandemic and timely implementation of NPIs could save thousands of lives.
Table 1: Calculated R0 and R(t) values for Mexico in 2009 with exponential growth considerations

<table>
<thead>
<tr>
<th>Time Period (days)</th>
<th>April 1 to April 20</th>
<th>April 20 to May 1</th>
<th>May 15 to June 1</th>
<th>June 1 to June 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>R0 &amp; R(t) Values</td>
<td>2.060174</td>
<td>1.784268</td>
<td>1.108769</td>
<td>0.937434</td>
</tr>
</tbody>
</table>

R0 and R(t) values for Mexico

![Graph showing R0 and R(t) values for Mexico]
Table 2: Emphasis of non-pharmaceutical interventions over time

<table>
<thead>
<tr>
<th>Type of NPI</th>
<th>April 1 to April 20</th>
<th>April 23 to May 15</th>
<th>May 15 to June 1</th>
<th>June 1 to June 15</th>
<th>June 15 to July 1</th>
<th>July 1 to July 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Distancing</td>
<td>Enforced (scattered)&lt;sup&gt;49&lt;/sup&gt;</td>
<td>Advised&lt;sup&gt;19,40&lt;/sup&gt;</td>
<td>None</td>
<td>None</td>
<td>Advised&lt;sup&gt;54&lt;/sup&gt;</td>
<td>None</td>
</tr>
<tr>
<td>School Closing</td>
<td>None</td>
<td>Enforced (Mexico City) &lt;sup&gt;31,33,34,35,37,40&lt;/sup&gt;</td>
<td>None</td>
<td>None</td>
<td>Enforced (scattered)&lt;sup&gt;47&lt;/sup&gt;</td>
<td>None</td>
</tr>
<tr>
<td>Public Space Closing (i.e. museums)</td>
<td>None</td>
<td>Enforced (Mexico City) &lt;sup&gt;31,33,34,35,37,40&lt;/sup&gt;</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Hand Hygiene</td>
<td>None</td>
<td>Assisted&lt;sup&gt;32,37,40,42&lt;/sup&gt;</td>
<td>Assisted&lt;sup&gt;52&lt;/sup&gt;</td>
<td>Advised&lt;sup&gt;49,53&lt;/sup&gt;</td>
<td>Advised</td>
<td>Advised&lt;sup&gt;51&lt;/sup&gt;</td>
</tr>
<tr>
<td>Isolation</td>
<td>Enforced (hospitals)&lt;sup&gt;48&lt;/sup&gt;</td>
<td>Enforced (hospitals)&lt;sup&gt;31,33,34,35,37,40&lt;/sup&gt;</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Face Masks</td>
<td>None</td>
<td>Assisted&lt;sup&gt;31,33,40&lt;/sup&gt;</td>
<td>Advised&lt;sup&gt;52&lt;/sup&gt;</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Coughing Etiquette</td>
<td>None</td>
<td>None</td>
<td>Advised&lt;sup&gt;52&lt;/sup&gt;</td>
<td>Advised&lt;sup&gt;50,53&lt;/sup&gt;</td>
<td>Advised&lt;sup&gt;54&lt;/sup&gt;</td>
<td>Advised&lt;sup&gt;51&lt;/sup&gt;</td>
</tr>
<tr>
<td>City shutdown</td>
<td>None</td>
<td>Enforced (Mexico City)&lt;sup&gt;43,44,45&lt;/sup&gt;</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

The metric used for the table follows: Enforced (laws mandating the NPI, likely occurred), Assisted (NPIs were recommended and items such as hand sanitizers were distributed to help the public implement them), Advised (NPI implementation was recommended by officials), None (no advisories advocating the implementation of NPIs)
Figure 1: Relationship between R0 and R(t) values and NPI emphasis in Mexico

**Comparison of R0 and R(t) to NPI**

![Graph showing the comparison of R0 and R(t) to NPI emphasis in Mexico. The graph displays the number of articles and R0 values for different periods, starting from April 1 to April 20 and ending with July 1 to July 15.](image-url)
There are two possible routes from La Gloria to Mexico City by bus. An interview with the bus company Autobuses de Oriente (ADO) on June 29, 2010 revealed that buses run regularly from Mexico City to Ciudad Cardel and Zempoala, both of which are geographically close to La Gloria. An interview with Donny Holaschutz on July 13, 2010 revealed that there are buses running from each of these cities to La Gloria. The interview also relayed conversations with ten people who lived in La Gloria, each of whom stated that about 50-70% of the residents of La Gloria commuted to Mexico for work. Given the capacity of ADO buses (about 50 people) and their frequency, we calculated that about 29% of these commuters travel per day from La Gloria to Mexico City. With the exception of Easter Week, these commuters stay in the Mexico City for several days to work before heading back to La Gloria.  

Figure 2: Travel between La Gloria and Mexico City for February-April 2009.
Chapter 4: Google Insights and R0

Introduction:

Measuring the impact of NPIs on a pandemic relies heavily on being able to measure the implementation of NPIs by a population. Current studies tend to use extensive surveys, interviews and observational field studies for this purpose, methods which can be unreliably implemented or generally difficult to use given a large enough sample population. To tackle large populations such as those present in a pandemic scenario, mathematical models have been used to plot the epidemic curve for a specific set of pandemic parameters. These studies generally reveal that NPIs impact an epidemic curve in a positive manner. However, very few can give concrete measurements of NPI use and even fewer have attempted to relate NPIs to the value of R0. Given the uncertainty in any measurement of NPIs, new methods of analysis must be implemented to bolster current findings. The use of Google Insights is one such method that has the potential to offer statistically significant insights when applied to a pandemic.

The effect of NPIs on R0 has become the object of heightened interest ever since the 2009 H1N1 pandemic. However, this relationship is difficult to assess as it is almost impossible to reliably measure whether people in a population are actually implementing NPIs.

The intuitive methods of directly surveying the implementation of NPIs become impractical with large populations, and computer simulations are often simplifications of pandemic scenarios that could discount several important factors.
Therefore, this thesis proposes an alternative method to measure NPI use, a method stemming from the prevalence of internet-based information, most of which is found via the popular search engine Google. The premise of the search engine is simple: a user types in keywords into a search bar, and the search engine runs through billions of websites looking for matches to the keywords or variations thereof. Searches on a particular keyword can be tracked using the website “Google Insights,” a subset of Google which allows internet users to see roughly how many searches were done on a particular term in a given time period, as well as compare searches on terms. Google Insights has proven to be a valuable tool by Journals like Nature and Munich Personal RePec Archive. Authors have used this valuable tool for tracking the interest in NPIs of a given population; the theory is that those who search for certain NPIs are more aware of them and therefore more likely to use them. As a result, we used Google Insights to track trends in NPI interest (and, by extension, NPI use), comparing this to the R0 value for the spring wave of the 2009 H1N1 Pandemic to evaluate the effect of NPIs on the value of R0 in a pandemic.

Methodology:

We compared R0 values and NPI searches for each state in the United States of America. R0 and R(t) values were calculated using data on the number of infections between May 1st and June 8th. This time period was the most accurate in terms of recording actual H1N1 cases, and therefore is the focus of our analysis. We assumed an exponential growth model for the segments of epidemic curves generated for each state with their infection data, during this time interval.

We used data from the World Health Organization (WHO) and the CDC to track the number of cases occurring in the spring outbreak of the H1N1 pandemic. We tracked cases daily for each individual state, primarily with the WHO data.
The number of infections per day was assumed to ideally follow an exponential growth curve as long as the infections we considered were assumed to be restricted to the first half of the epidemic curve. Therefore, we took the number of infections for discrete segments of time (two weeks) and found the best-fit exponential curve for each segment. Assuming that this curve represented the entirety of the initial growth of the epidemic if it had been unhindered by outside influences such as NPIs, we extrapolated the initial number of cases from the curve’s equation, taking into account the reported 2.5 incubation period of the H1N1 virus. The actual value of R0 and R(t) was then approximated by finding the ratio of a single day’s number of cases to the previous day’s. This conceptual frame of R0 is integral to our method of calculating the basic reproductive number. It also is easily confused with the progression of the reproductive number over time, R(t). There are some key differences between the R0 we calculate and R(t), however. The latter is defined as a reproductive rate which varies over time because it is entirely dependent on the number of people infected over time in a pandemic. It relies on a particular R0 value and does not account for the changing behavior of a given population.

Simply put, the method of calculating R0 is to take a daily distribution of infected cases for a given pandemic and break this distribution into discrete time periods. For both of my chapters, I have used a time period of roughly two weeks. Since these numbers are taken directly from observed values during the H1N1 pandemic, it is more than reasonable to assume that the populations observed were altering their behaviors as the pandemic progressed. For Mexico, in fact, as mentioned earlier, there is evidence of government enforcement of NPIs such as social distancing, and encouragement of NPIs like hand washing and wearing face masks. Therefore, the value of R0 is changing as the pandemic progresses, for at the very least the frequency of contact is decreasing.
In light of this, the distribution of infected individuals was plotted for every two weeks and estimated a best-fit exponential curve to the data. All else equal (i.e. no behavioral changes), an exponential approximation would be valid for the initial segment of an epidemic curve, which is where this data is assumed to be located in the H1N1 pandemic curve.

This best-fit curve was used to extrapolate backwards, accounting for an incubation period of 2.5 days, in order to determine what the R0 value was at the beginning of each of these curves; in other words, what the R0 value was given the set of behaviors attributed to the population at a given point in time.

This approach resulted in several R0 values which decrease in value as an epidemic curve progresses because they were taken from different States at the beginning of the infection in those States. Note that these values are not exclusively dependent on the number of people infected, like R(t). Instead, they are calculated with the fundamental notion of \( R_0 = \lambda \rho \); as the behavior of a population changes, so does at least one of these two variables, and thus so does R0.

In order to track NPI interest for the same discrete time intervals, we used Google Insights with specifically chosen search terms. We focused on the NPIs of hand washing, the use of face masks, and the use of hand sanitizers because these three were the most likely NPIs to be searched for the purpose of understanding how to perform them correctly. In addition, they are simple tasks that can be done individually, unlike NPIs such as school closings.
The above mentioned terms were likely to have the best correlation between Google Insights’ interest rating and actual implementation. The specific search terms used were “hand washing”, “face masks”, and “hand sanitizer.” We searched within a particular state for the entire year of 2009 in order to obtain values indicating the interest in an NPI during the H1N1 Pandemic.

It should be noted that Google Insights reports a normalized value of searches on a specific search term in order to protect the privacy of Google users. Therefore, the data we received from Google Insights was not in the form of specific numbers of searches. Rather, it was normalized across an entire year and therefore cannot be further predicted mathematically. Also, Google Insights has a threshold under which it does not report the interest in a search term. Some states such as West Virginia were consistently under this threshold, likely as a result of low population density coupled with low internet use/interest in NPIs.

In order to account for some of these issues, we looked at the availability of internet access in each state from the most recent survey by the United States Census Bureau, a survey conducted in 2007. We compared R0 to NPI use graphically by normalizing R0 as well—we divided by the highest value and multiplied the result by 100. This gave us R0 values which were not correct numerically, but the trends of which could be compared visually. We also looked for correlations between R0s and NPIs using SPSS software, to analyze the significance of R0 to the implementation of NPIs.
Results:

Our results revealed that a higher peak of Google hits occurred right before the peak of the value of R0 corresponding to a higher interest in NPI use in each state. We believe that, as the government through the media tried to enforce NPIs over the population, people looked up these NPIs online using Google out of either fear or curiosity. We assume that a high incidence of Google hits for a particular NPI means that people are implementing a particular NPI more. This implies that people who educate themselves about NPIs actually change their behavior as a result. We have taken into account that internet access was available to 68-78% of the populations in 2007 in all the seven states we considered. The Google Insights searches were for 2009, and recent publications reveal that internet use has only increased in the past years, largely due to the prevalence of smart phones. Therefore, it is likely that a greater portion of the population had the ability to educate themselves about NPIs through the internet.

We chose to include seven states out of the forty-nine (due to the availability of data from these 7 States) that returned Google Insight results to best illustrate the general trends of these results. These states showed an interesting relationship between R0 and Google hits. New York State had an R0 value of 1.6 in the first two weeks of May 2009, a value which dropped to 1.1 in weeks following. Right before the decline in R0 there was a peak in NPI interest (Figure 1).
This trend can be seen in several different ways, but one plausible explanation is that there is a delay between implementation and the change in R0; as people implement NPIs, R0 decreases over time. Granted, the R0 values we calculated were for every two weeks, which ignores fluctuations in those discrete time segments, but overall the trend of NPIs decreasing as R0 decreases bears out our explanation.

People begin implementing NPIs to decrease R0, and once R0 begins to decline so does their use of NPIs. To finish illustrating this trend, we see that Massachusetts had an initial R0 of 1.7 in the first two weeks, which dropped to 1.2 in the later two weeks of May 2009 (Figure 2). Colorado, a state with a high percentage of the population with internet access, had an initial R0 of 1.8. This value became 1.2, also having a peak before R0 peaked (Figure 3). These similar possible relationships between R0s and Google hits were also seen in states like California, Washington DC, Arizona and Texas (Figures, 4, 5, 6, and 7).

Discussion:

Our data shows that R0 may relate to the interest in the NPIs of hand washing, face masks, and hand sanitizers as tracked by Google Insights. With the assumption that increased internet searching corresponds to increased use of NPIs, our findings are consistent with that possibility. In short, the use of NPIs potentially decreases the severity of a pandemic.

This conclusion is far from concrete. There are likely other reasons for typing the phrase “hand washing” into Google, reasons that have little to do with implementing NPIs during a pandemic.
There is a possibility that because of initially reported high death rates, people were initially frightened and as it became clear that the pandemic was relatively mild, concern wained. Our finding is counter-intuitive of what one might have expected that NPI interest would grow as R0 falls. With more study, Google insight may very well prove itself to be a reliable tool for measuring the activities of a population as well as their interests in a particular search term.

**Conclusion:**

We believe that because technology can play a large part in transferring information from one person to another via search engines like Google and social networking sites like Twitter, we can use this technology to track a population’s interest in a particular NPI. More research on the relationship between Google Insights and the activities of individuals will likely shed light on the reliability of our assumption that NPI interest corresponds to NPI practice.

As it stands, our data relates R0 to the possible implementation of NPIs in a population and our interpretation is suggestive of a positive significant causation. In light of this and previous studies, I suggest that the emphasis on the implementation of NPIs should be made readily apparent in news sources and government recommendations as soon as a pandemic is recognized in a country. Media technology is becoming more and more prevalent. Getting information about the proper implementation of NPIs to a population should not require extensive Google searches on the part of an individual. Government agencies can harness the power of the internet to effectively inform citizens of the importance of NPI implementation. As a result, the severity of a pandemic is likely to decrease, sparing many the risk of being infected by a novel virus.
Figure 1

New York RO vs. NPI Hits

* = significant at 0.01 level

68% Internet Access in 2007

Figure 2

Massachusetts RO vs. NPI Hits

72.9% Internet Access in 2007
**Figure 3**

*Colorado R0 vs. NPI Hits*

* = significant at 0.01 level

78.9% Internet Access in 2007

<table>
<thead>
<tr>
<th>Weeks in May 2009</th>
<th>R0 Value*</th>
<th>Hand Washing</th>
<th>Face Mask</th>
<th>Hand Sanitizer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>

**Figure 4**

*California R0 vs. NPI Hits*

* = significant at 0.01 level

73.6% Internet Access in 2007

<table>
<thead>
<tr>
<th>Weeks in May 2009</th>
<th>R0 Value*</th>
<th>Hand Washing</th>
<th>Face Mask</th>
<th>Hand Sanitizer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>
Figure 5

Washington D.C. R0 vs. NPI Hits

* = significant at 0.01 level

74.7% Internet Access in 2007

- Hand Washing
- Face Mask
- Hand Sanitizer

Figure 6

Arizona R0 vs. NPI Hits

* = significant at 0.01 level

71.7% Internet Access in 2007

- Hand Washing
- Face Mask
- Hand Sanitizer
Figure 7

Texas R0 vs. NPI Hits

* = significant at 0.01 level

68.1% Internet Access in 2007

- Hand Washing
- Face Mask
- Hand Sanitizer

Normalized Values

weeks in May 2009
Chapter 5: Thesis Conclusion:

We conclude that all of our data relates R0 to the possible implementation of NPIs in a population and is suggestive of a positive significant causation. We think, NPIs are important to implement at the early stages of a pandemic, as they could potentially decrease the infection rate. NPIs have been shown to have an effect on R0 in several case studies mentioned in this thesis. La Gloria, which is thought to be the first place where the H1N1 pandemic started, had a relatively high R0 since people did not know that they were getting sick from a novel influenza virus, and thus no NPIs were implemented. Later, in Mexico City people became aware of the novel virus and there was more emphasis on the use of NPIs in the media. This emphasis may possibly link to the speedy decrease in R(t) over time. News articles and Google Insights also have shown interesting relationship between the two.
References for Chapters 1 and 2:


References for Chapter 3:


30. Sypsa V, & Hatzakis A. School closure is currently the main strategy to mitigate influenza A(H1N1)V: A modeling study. *Eurosurveillance.* June 18, 2009; 14(24).


**References for Chapters 4:**

7. “Google Searches as a Means of Improving the Nowcast of Key Macroeconomic Variables.” Authors: Konstantin A. Kholodilin, Maximilian Podstawski, Boriss Siliverstovs, Constantin Bürgi. Published German paper
8. “A Poisson Regression Examination of the Relationship Between Website Traffic and Search Engine Queries.” Authors: Heather Tierney, Bing Pan. Published article in *Munich Personal RePec Archive*
