

# Use of EHR Data to Identify the Influence of Pre-existing Conditions on COVID-19 Outcomes

By: Anisha Tehim





# TABLE OF CONTENTS

## 01. Introduction & Literature Review

Current Research  
Available

## 02. Methods

Techniques & Code  
Utilized to Develop  
Analyses

## 03. Data Analysis

Results and Important  
Points of Notice

## 04. Conclusion

Results and their  
Implications

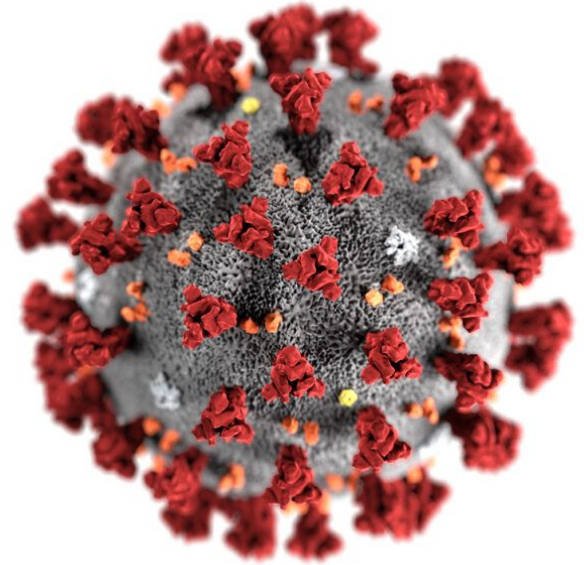
# 01.

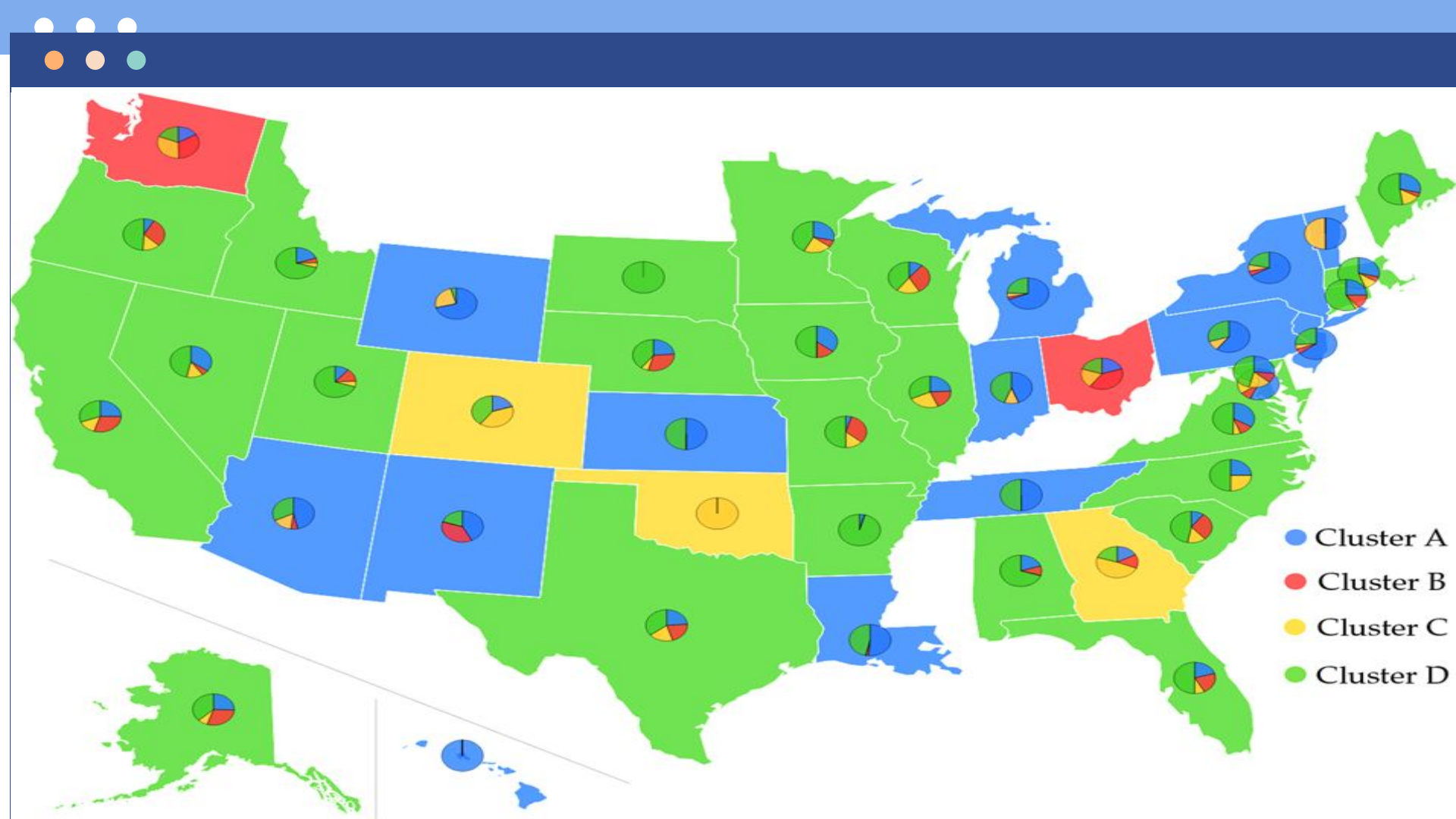
## Introduction & Literary Review



# Introduction to the SARS-CoV-2 Virus

- Transmits through virus particles and droplets when an infected person breathes, talks, coughs
- Currently thousands of Covid-19 variants that differ by at least one mutation
  - New variants have the ability to alter the virulence and transmissibility of the virus (Sauer, 2021).





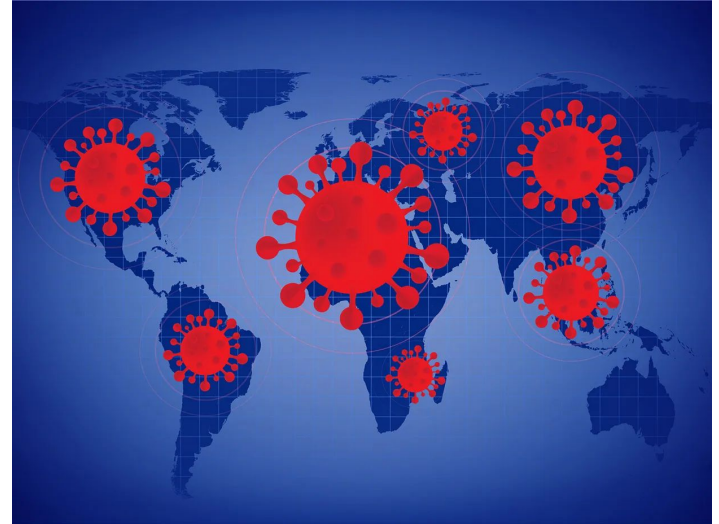
# Current Well Known Pre-Existing Conditions

- Heart Diseases:
  - Congenital heart disease, Coronary artery disease, and hypertension(Mayo Clinic, 2021)
- Other medical conditions:
  - Chronic lung disease, asthma, obesity, severe cough, fever, hypertension, and wheezing (CDC, 2021)



# The Focus of Current Research

- National or international perspective
- Majority of research also targeted towards older patients, usually 30+





# Addressing The Gap

- Filtered data → so that the sample population is specifically from New Jersey
- Organizes entirety of the data into several age groups (0-4; 5-17; 18-49)
- Allows for targeted analysis towards New Jersey populations, when understanding which diseases are correlated with increased outcomes of SARS-CoV-2
  - Categorizing by age also allows for the detection of discrepancies among different age groups



# 02.

## Methods



# Justification of Method

- Need for more research on how individual health characteristics make populations with certain diseases more susceptible to the virus than others
- Improving understanding of immune susceptibility will lead to more effective, targeted treatment



# Understanding how “R” can be Utilized to Analyze Patient Data in the Context of Covid-19

- R is a statistical programming language often used to mine and analyze large quantities of data.
- Several R Packages (such as those listed below) are used to analyze the data
  - dataframe
  - dplyr
  - ggplot

```
dens <- density(data, n = npts)
dx <- dens$x
dy <- dens$y
if(add == TRUE)
  plot(0., 0., main,
       ylab,
       if(orientati == yst)
         dx2 <- (dx - min(dx)) / (max(dx) - min(dx))
         x[1.]
         dy2 <- (dy - min(dy)) / (max(dy) - min(dy))
         y[1.]
         seqbelow <- rep(y[1.], length(dx))
         if(Fill == T)
           confshade(dx2, seqbelow, dy2
```



# Procedure

- Data Source: Electronic Health Record Data from Hospitals Across the United States
- Utilize data from only hospitals in NJ, filtered by several age groups (0-4;5-17;18-49)
  - Each age group has a sample size of 50
- For each of the following conditions: **Asthma**, COPD/emphysema, **Diabetes**, Coronary artery disease, Heart failure, Hypertension, **Obesity, Chronic Kidney Disease, Abdominal Pain**, Altered mental, Anosmia/decreased smell, Chest pain, , Congested/runny nose, **Cough, Diarrhea**, Dysgeusia/decreased taste, Fever/chills, Headache/bloody sputum, Muscle aches, Sore throat, Wheezing, Acute renal failure/acute kidney injury, Acute respiratory distress syndrome, Acute respiratory failure, Pneumonia, **Sepsis**

# Procedure Continued

- For each disease or medical condition, a correlational analysis was performed utilizing R, demonstrating which condition was tied to a positive SARS-CoV-2 outcome
- The data will be displayed utilizing mosaic bar graphs for each condition, organized by age group



# Mirrored Study

- Utilized data from “OpenSafely”
- Primary care records of 17,278,392 adults were pseudonymously linked to 10,926 COVID-19-related deaths
- Most relevant factors included being male, diabetes, severe asthma, as well as ethnicity

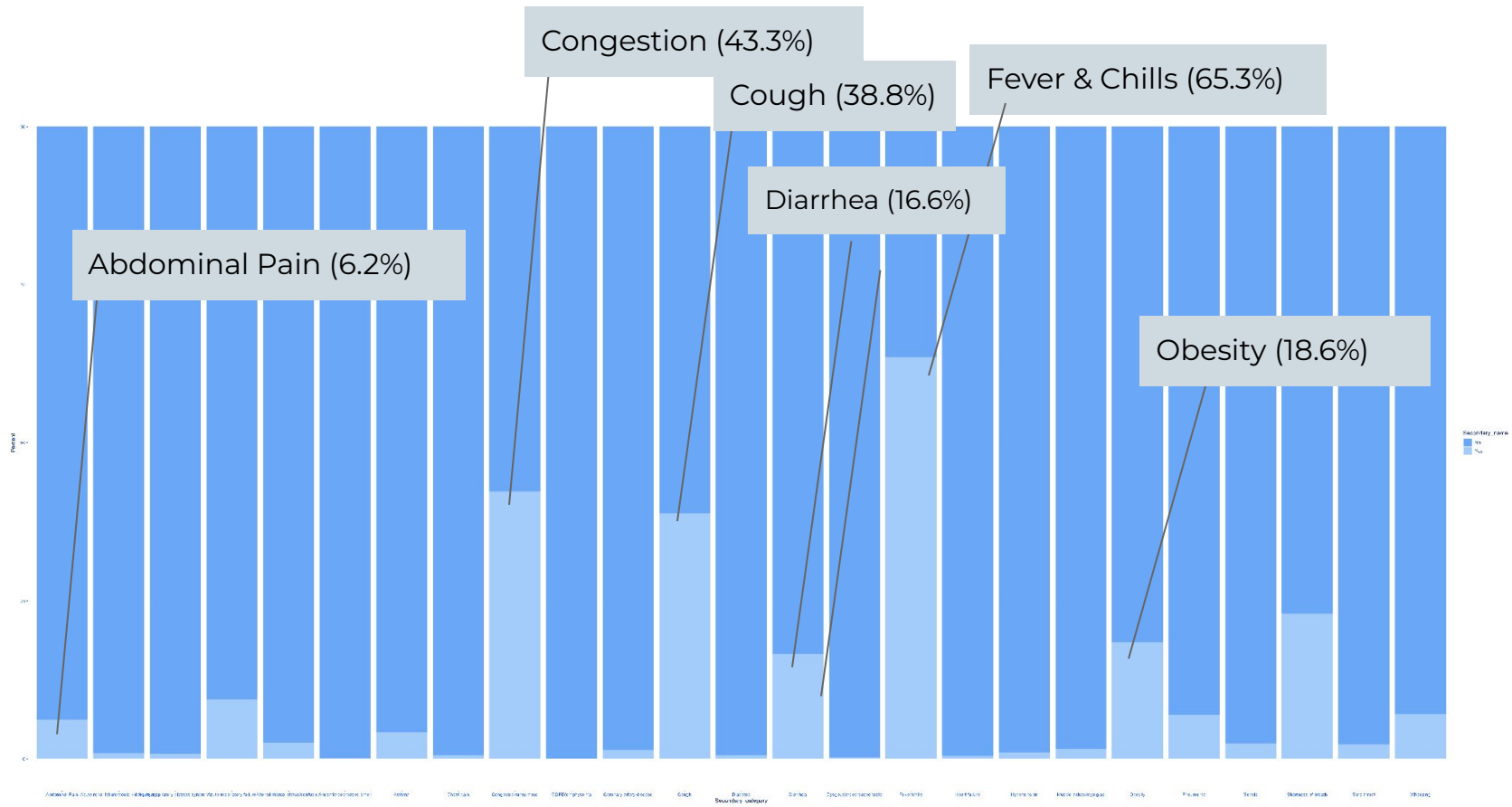


# 03.

## Data Analysis



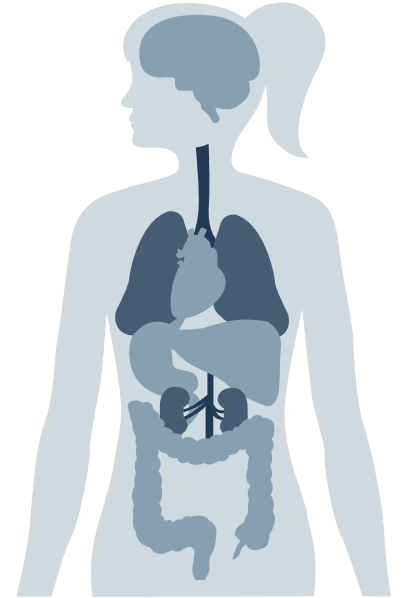
# Age Group: 0-4



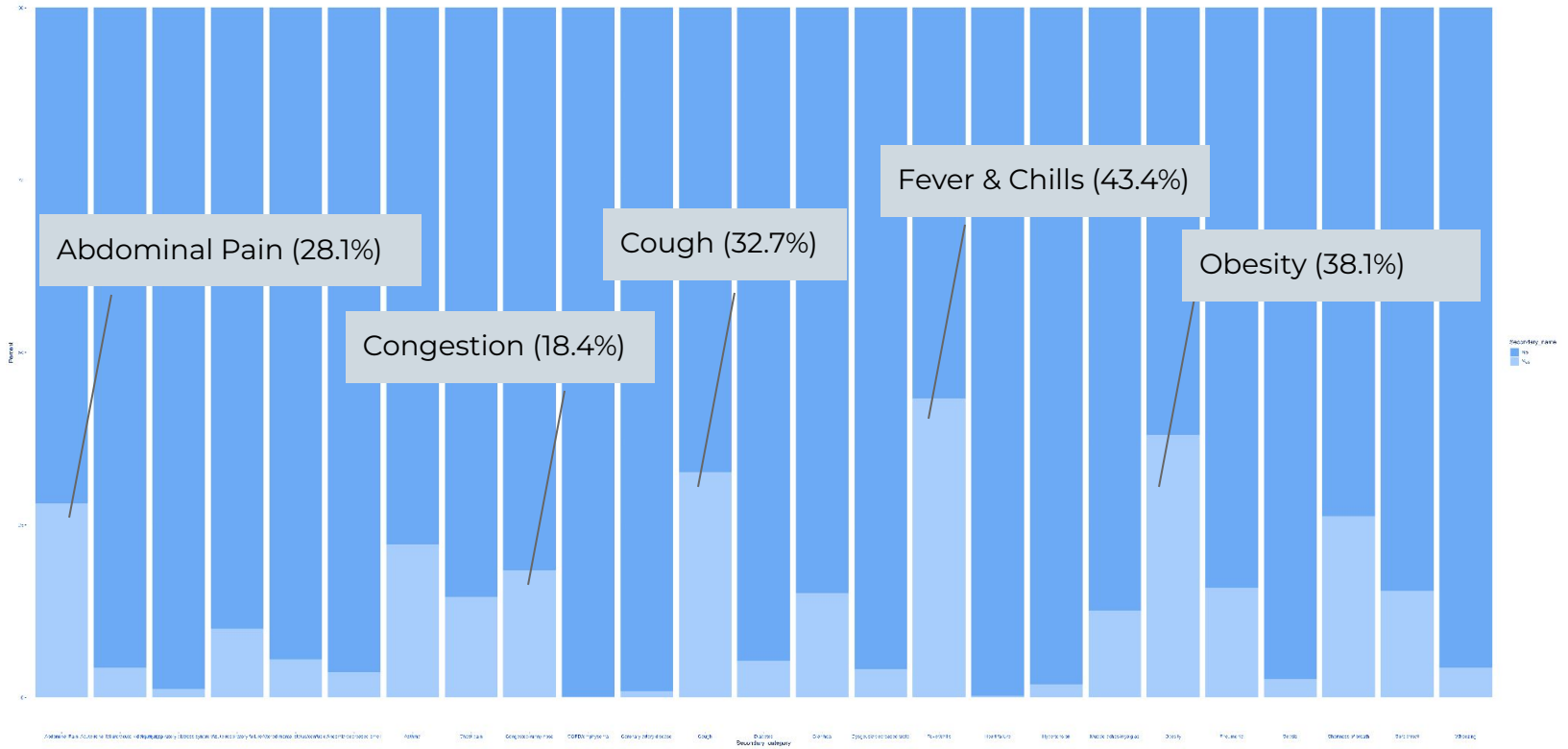


# Age Group: 0-4

- Most common symptoms: Fever and Chills (63.5%), Cough (38.8%), and Congestion (43.3%)
- Diarrhea (16.6%) was also relatively high
  - Only approximately 2% lower than Obesity and even 12.4% higher than Asthma
- Abdominal Pain at 6.2%.
- The prevalence of diarrhea and abdominal pain suggests the virus has an impact on other systems such as the digestive system



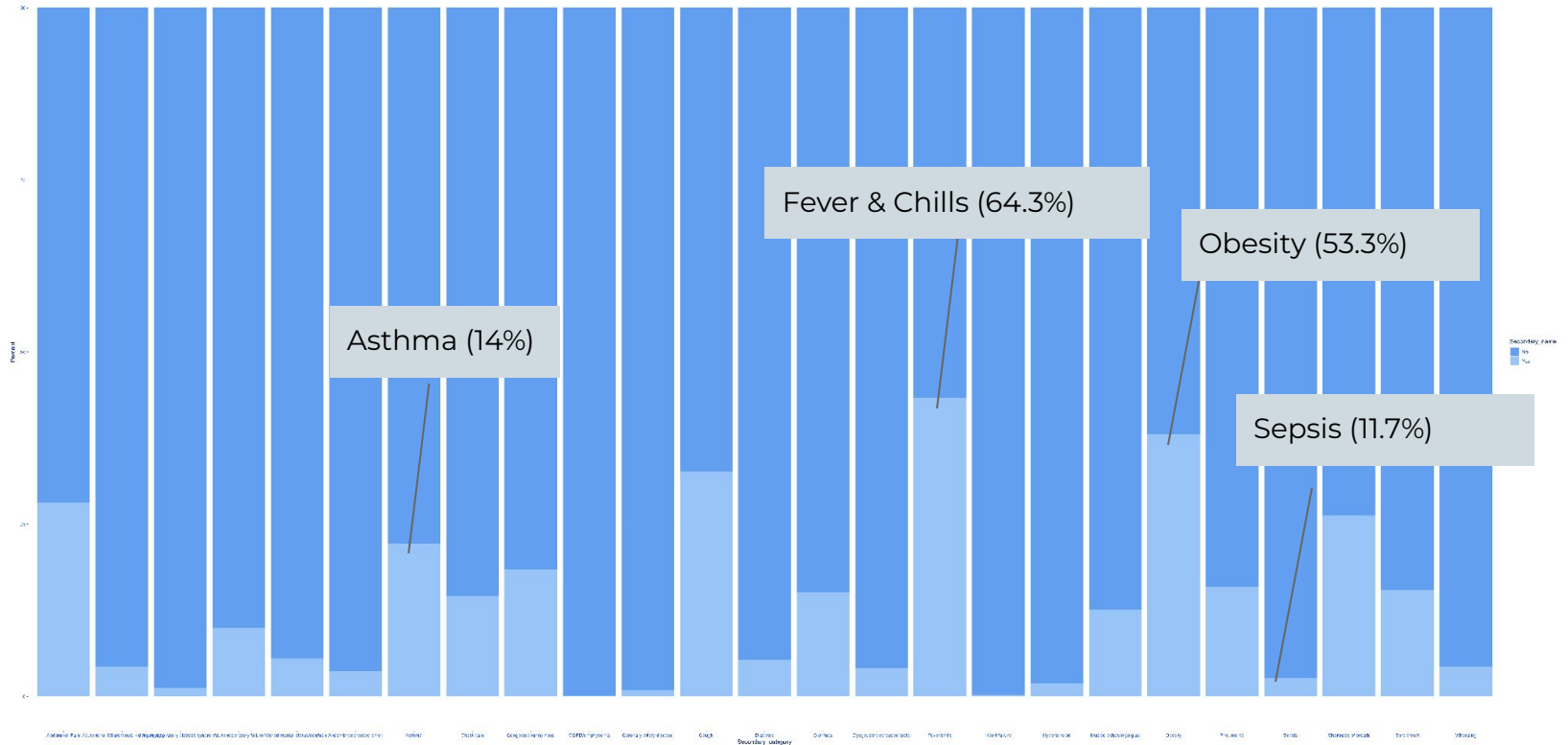
# Age Group: 5-17



# Age Group: 5-17

- Similarly common symptoms (0-4) included Fever and Chills (43.4%), Cough (32.7%), and Congestion (18.4%)
- Most common symptom after Fever and Chills was Obesity (38.1%)
  - Aligns with previous findings in the literature review
- Abdominal pain - (28.1%), approximately ten percent higher than the symptom of congestion.

# Age Group: 18-49



# Age Group: 18-49

- Obesity (53.3%) and Fever and Chills (64.3%) were among the most common within patients
- Sepsis (11.7%), was at 2.7% percent in the younger age group
- Most articles found in the literature review state Asthma as one of the most common pre-existing conditions to SARS-CoV-2
  - Only 14.% of subjects here had asthma.

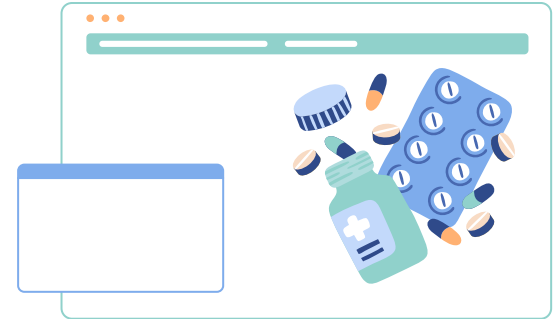
# 04.

## Conclusion



# Conclusion

- While predicted variables such as Fever & Chills, Asthma and Cough = most highly associated with increased Covid-19 outcomes, many unpredicted variables were also found
  - Abdominal Pain & Diarrhea had significantly higher percentages than expected
  - In addition, many results varied immensely among age groups, which affirms the unpredictable nature of the virus





# Future Study & Limitations

- International vs Regional Similarities
- Limitations: While sample size = 50 for all the age groups, ideally a large sample size could have led to more conclusive results
- Research presented here is aimed to identify the pre-existing medical conditions most associated with the SARS-CoV-2 virus
  - Hopefully inform the general public about the influence of these medical conditions in NJ, and aid in treatment delivery



# Citations

Armitage, H. (2021, July 21). Genetics could explain why some people get severe COVID-19. Stanford Medicine.

Retrieved October 13, 2021, from

<https://scopeblog.stanford.edu/2021/07/21/genetics-could-explain-why-some-people-get-severe-covid-19/>

Collins, F. (2021, July 20). More Genetic Clues to COVID-19 Susceptibility and Severity. NIH Director's Blog.

<https://directorsblog.nih.gov/tag/the-covid-19-host-genetics-initiative/>

A Complete Guide to Stepwise Regression in R. (2019, April 27). Statology.

<https://www.statology.org/stepwise-regression-r/>

Dhar et al. (2021, June 3). Genomic characterization and Epidemiology of an emerging SARS-CoV-2 variant in Delhi,

India. JohnsHopkins Bloomberg School of Public Health. Retrieved August 25, 2021,

from <https://ncrc.jhsph.edu/research/genomic-characterization-and-epidemiology-of-an-emerging-sars-cov-2-variant-in-delhi-india/>

<https://ncrc.jhsph.edu/research/genomic-characterization-and-epidemiology-of-an-emerging-sars-cov-2-variant-in-delhi-india/>

# Citations (Continued)

Katara, P. Role of bioinformatics and pharmacogenomics in drug discovery and development process. *Netw Model Anal*

*Health Inform Bioinforma* 2, 225–230 (2013). <https://doi.org/10.1007/s13721-013-0039-5>

Mayo Clinic Staff. (2021, October 7). COVID-19: Who's at higher risk of serious symptoms? Mayo Clinic.

National Center for Immunization and Respiratory Diseases. (2021, August 21). People with Certain Medical Conditions.

Centers for Disease Control and Prevention.

Sauer, L. M. (2021, May 19). What Is Coronavirus? Johns Hopkins Medicine. Retrieved August 23, 2021,

from <https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus>

Williamson, E.J., Walker, A.J., Bhaskaran, K. et al. Factors associated with COVID-19-related death using OpenSAFELY.

*Nature* 584, 430–436 (2020). <https://doi.org/10.1038/s41586-020-2521-4>

Zeberg, H., Pääbo, S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature* 587,

610–612 (2020). <https://doi.org/10.1038/s41586-020-2818-3>



**Thank You**