Connection between COVID-19 Variants (Gamma and Delta Variant) and Demographics Using Python

Sneha Nangunoori[1], Rajagopal Appavu[1#] and Jothsna Kethar[1#]

[1]Gifted Gabber
[2]Advisor

ABSTRACT

From schools shifting to virtual learning and offices promoting work from home, COVID-19 transformed the way the world functions. Like any other virus, the coronavirus has many variants. This research paper discusses the connection between two prominent variants: the Gamma variant and the Delta variant and certain demographic features like gender, age, and location. The method used in this research paper includes finding data from credible sites and other evidence and using python to extract the needed data to support theories. The theories stated in this research paper are not completely valid due to the lack of strong evidence. So, instead of concluding with a strong thesis, this research paper aims to motivate other researchers to delve even deeper into this topic. Basically, this study focuses on the importance of future research and acts as a steppingstone for other researchers who are interested in learning more about COVID-19 variants and their connection with different demographic features.

Introduction

The COVID-19 pandemic has had a huge impact on many lives all over the world. The three countries that have the highest number of COVID-19 cases in the world are the United States (33.9M), India (30.9M), and Brazil (19.1M). Why do these countries have the most cases? The population is relevant to the number of cases but, this is where the word 'variants' comes into play. Some variants have more severe effects and spread more quickly in certain places. For example, the Lambda variant has spread across South America. The Lambda variant that was discovered on 14th June 2021 has a stronger presence in South America. On the other hand, the Delta variant (B.1.617.2. variant) that originated in India is more transmissible than any other variant. This raises many questions. Are there regions in the world that are prone to making a virus even more deadly and contagious? However, are there any other variables that constitute the connection between geographical location and the prominence of variants? To answer these questions, it is crucial to gather information about the common COVID-19 variants and analyze their connection with demographic features such as gender, age, and location.
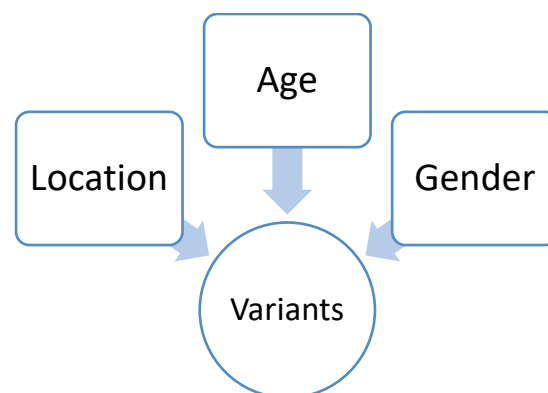

Figure 1. Visual representation of the demographics factors that were discussed and variants

Method

The approach to finding data was done by the following procedure. First, a resource that had the number of a specific variant's cases was used to show which countries had the greatest number of cases. Next, by analyzing data from a credible website, the percentage of the number of female and male cases of a country's total population as of a specific date was used to support theories on the connection between that specific variant and sex. As the credible website allows its users to download the data as a CSV file, Python programming language was used to extract the information that was essential for finding the connection between a certain variant and demographic factor. Using the data from the CSV file, theories (that still need stronger evidence) were developed.

```python
import csv
data=[]
with open('''file name''') as csv_file:
    reader=csv.reader(csv_file, delimiter=',')
    for row in reader:
        data.append(row)
name='''variable name'''
column = [x[0] for x in data]
if name in column:
    for x in range(0,len(data)):
        if name==data[x][0]:
            print (data[x])
else:
    print("Data not found")
```

Figure 2. Python code that was used to extract the essential data from a CSV file.

Variants

During replication, a virus often undergoes genetic changes that may create what are called variants. There are numerous COVID-19 variants globally. However, this research paper focuses on two well-known variants: the Gamma variant and the Delta variant.

Gamma Variant (Lineage P.1)

According to the World Health Organization, the Gamma Variant is a 'Variant of Concern'. This variant has a high attack rate even in fully vaccinated persons and caused the third wave in French Guiana (July 2021). Its notable mutations N501Y, K417N, and E484K can trick the human ACE2 receptor (Vignier, 2021). The Gamma Variant originated in Brazil and later it was designated as a Variant of Concern. The Gamma variant is common in the United States of America (41.0%), Brazil (29.0%), Canada (10.0%), Belgium (4.0%), and Chile (4.0%).
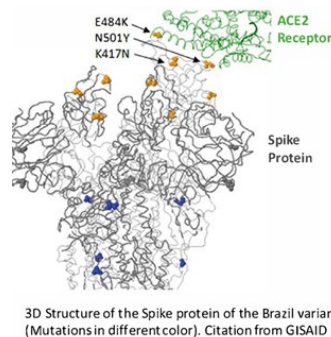

3D Structure of the Spike protein of the Brazil variant
(Mutations in different color). Citation from GISAID

Figure 3. 3D representation of the Spike Protein of the Gamma variant

Delta Variant (Lineage B.1.617.2)

According to the World Health Organization, the Delta Variant is a 'Variant of Concern'. The Delta variant is more contagious than any other variant and it originated in India. This variant was first detected in December 2020 in India, and it has made its way to 80 other countries such as the United States of America and Great Britain. However, there is still a lot more to learn about this variant, like getting to know whether it will make one sicker than the original virus: the severity of this variant is unknown (Katella, 2022). The Delta variant is common in the United Kingdom (66.0%), the United States of America (8.0%), India (6.0%), Denmark (4.0%), and Germany (2.0%).

Location

Gamma Variant

Since the United States of America and Brazil have the highest number of Gamma Variant cases in the world, it is crucial to understand the demographic features that affected the spread of the Gamma Variant in these two countries. As the Gamma variant was first detected in the United States on January 25th, 2021, it is important to examine the number of cases in each demographic feature reported after the time when it was first detected in the USA. Similarly, as the Gamma Variant was first found in Brazil in July 2020 it is vital to gather data on the number of cases in each demographic feature reported after the time it was first detected in Brazil.

Delta Variant

Even though the virus originated in India, India doesn't seem to have a huge number of Delta variant cases. As the United Kingdom has the highest number of Delta variant cases in the world and as the delta variant originated in India, becoming aware of the demographics that might contribute to the high number of Delta variant cases and reasons for its origin in these two countries is crucial. As the Delta variant was first detected in India in October 2020, it is important to examine the number of cases in each demographic feature reported after the time when it was first detected in India. Similarly, as the Delta Variant was first found in the United Kingdom in mid-April 2021 it is vital to gather data on the number of cases in each demographic feature reported after the time it was first detected in the United Kingdom.

Gender

Generally, males are less resistant to infections than females. The same fact applies to COVID-19. As this virus is a new disease there is still a lot to learn about its relationship with gender. Like any other virus, COVID-19 has a different impact on females and males. Each COVID-19 variant may have different impacts, too.

Gamma Variant

The United States of America has 22,113 Gamma variant cases and Brazil has 14,810 Gamma variant cases as of July 23, 2021 (Statista, 2022b). Unfortunately, there are no resources on the number of Gamma variant cases with relation to gender in the two countries. However, there is an alternative to find a very rough estimate. The following pie charts visualize the data retrieved from the CSV file:
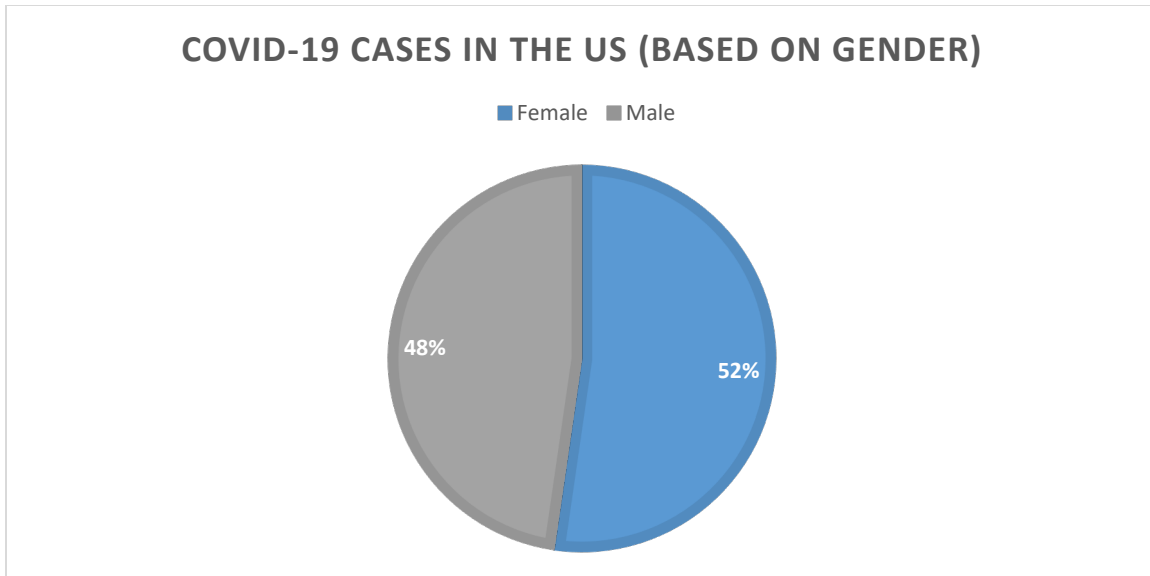
COVID-19 CASES IN THE US (BASED ON GENDER)

Figure 4. Percentage of male and female COVID-19 cases in the United States as of 17th February 2021(more than two weeks after 25th January 2021). This is a rough estimate as it shows the percentages of the total number of COVID-19 cases in the US and not of the total number of Gamma variant cases in the US.
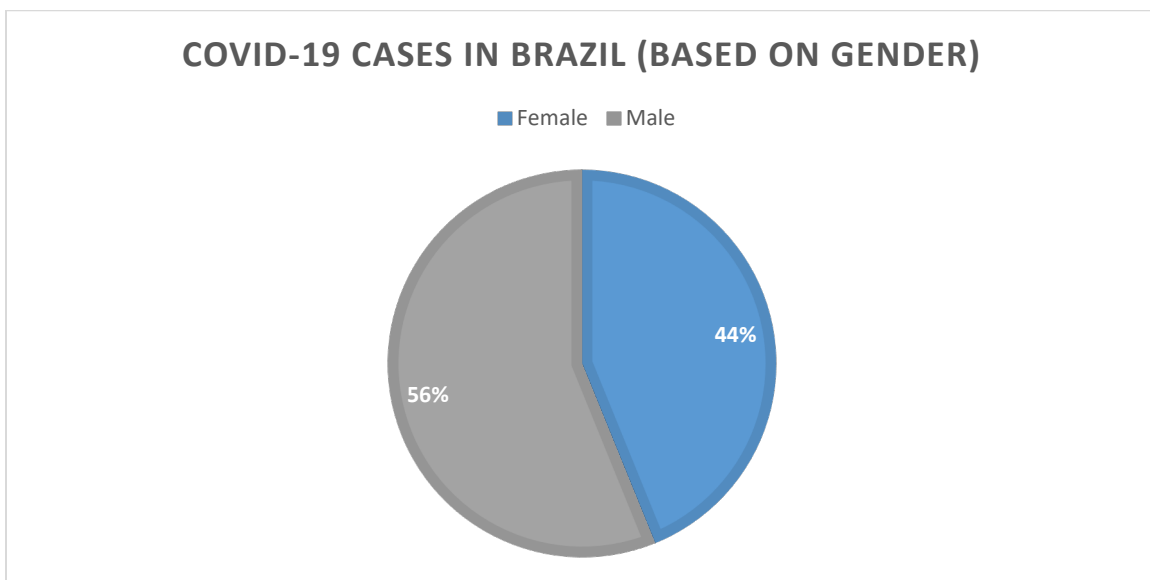

COVID-19 CASES IN BRAZIL (BASED ON GENDER)

Figure 5. This pie chart shows the percentage of male and female COVID-19 cases in Brazil as of 19th December (months after July 2020). This is a rough estimate as it shows the percentages of the total number of COVID-19 cases in Brazil and not of the total number of Gamma variant cases in Brazil.

5.2. Delta Variant

The United Kingdom has 167,856 Delta variant cases and India has 15,474 Delta variant cases as of July 27, 2021 (Statista, 2022). The following pie charts visualize the data retrieved from the CSV file.
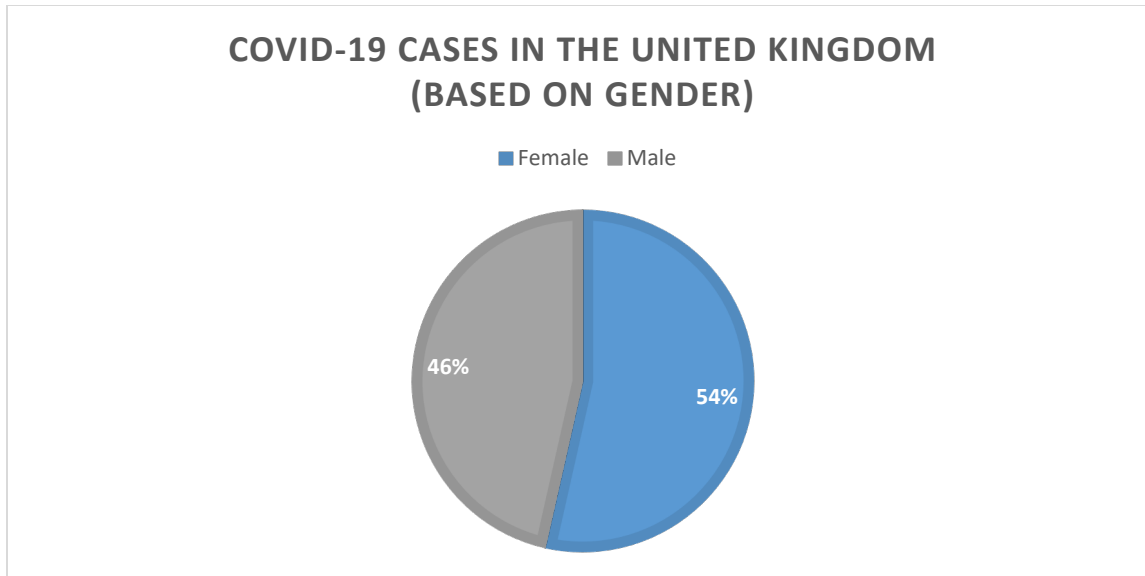
Figure 5. This pie chart shows the percentage of male and female COVID-19 cases in the United Kingdom as of 16th May 2021(more than two and a half weeks after mid-April 2021). This is a rough estimate as it shows the percentages of the total number of COVID-19 cases in the United Kingdom and not of the total number of Delta variant cases in the United Kingdom.
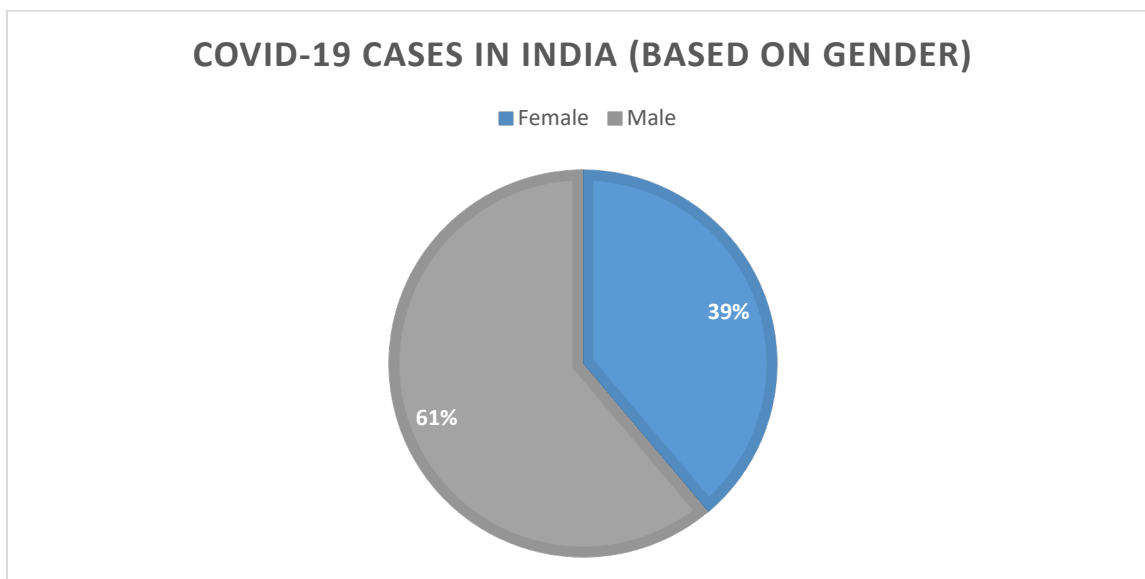

Figure 6. This pie chart shows the percentage of male and female COVID-19 cases in India as of 18th May 2021 (months after October 2020). This is a rough estimate as it shows the percentages of the total number of COVID-19 cases in Brazil and not of the total number of Delta variant cases in India.

Gender and Age

People of all ages can get affected by COVID-19. However, the severity of the symptoms of this virus is high as one's age increases. But there are chances of this theory not being valid for some COVID-19 cases.

Gamma Variant

This part of the research is going to analyze if there is any connection between Gamma variant cases and sex and age in Brazil and the USA. In the USA, females who are 20-29 years old are getting infected by COVID-19 more than any other age group and sex (10756.29 cases per 100,000 cases). This data corresponds with the previously mentioned data in the pie chart: 52% of the cases in the USA were of the female sex (more females are contracting COVID-19 than males). However, there are no resources to state the specific date about when the data on females who are 20-29 years old are getting infected by COVID-19 more than other age groups and sex was collected. This date is a very important piece of information as the Gamma variant was first discovered in the USA on January 25th, 2021. Only the data collected after January 25th, 2021, would be valid for supporting any theory that is to be developed about which sex and age group is more susceptible to getting contracted by the Gamma variant. In Brazil, unfortunately, data on the number of COVID-19 cases based on sex and age wasn't available.

Delta Variant

This part of the research is going to analyze if there is any connection between Delta variant cases and sex and age in the United Kingdom and India. In the United Kingdom, females who are 20-29 years old are getting infected by COVID-19 more than any other age group and sex (11057 cases per 100,000 cases). This data corresponds with the data that was previously mentioned in the pie chart: 54% of the cases in the United Kingdom were of the female sex (more females are contracting COVID-19 than males). However, there are no resources to state the specific date about when the data on females who are 20-29 years old are getting infected by COVID-19 more than other age groups and sex was collected. This date is a very important piece of information as the first Delta variant case was discovered in the United Kingdom in mid-April 2021. Only the data collected after that date would be valid for supporting any theory that is to be developed about which sex and age group is more susceptible to getting contracted by the Delta variant in the United Kingdom.

Conclusion

In conclusion, this research paper aims to inspire other researchers to conduct deeper studies on the connection between different COVID-19 variants and demographics. The demographic features that were discussed in this paper were location, gender, and age. However, there are so many other factors like underlying medical conditions, race, ethnicity, etc. Examining the relationship between the other demographic factors and COVID-19 variants will be highly beneficial for the world. Prospective research on this might even improve vaccine efficacy (to fight against a variety of variants). There are so many other questions to be answered of COVID-19 and a vast number of discoveries to be made. Scientists are still learning about this new virus and its characteristics in different environments. As this is a relatively new virus, there is not much data to utilize for supporting theories. Many countries all over the world don't have published datasets about the demographics of their COVID-19 cases. This proves that there is still a lot of research to be done. This research paper is just the start of many potential studies on variants and demographic features. Hopefully, one day, the theories mentioned in this research paper can be further improved with strong and factual evidence.

Limitations

Due to the lack of specific data required for this research, no rigid conclusions can be made. However, theories can be developed. In the USA, the Gamma variant seems to be affecting the female population slightly more than the male population and in Brazil, this variant appears to be common in the male population more than the female population

by a substantial amount and in the United Kingdom, the Delta variant seems to be affecting the male population slightly more than the female population and in India, this variant appears to be common in male population more than the female population by a massive amount. Considering the lack of information, no theory can be developed on which sex and age groups are more prone to contracting the Delta variant in India and United Kingdom.

References

3D representation of the Spike Protein of the Gamma variant. (2021). [Illustration]. https://gvn.org/covid-19/gamma-p-1/

Gender and COVID-19. (2021, July 15). Home >. Gender & Covid-19. https://www.genderandcovid-19.org/

Katella, K. (2021, July 27). 5 Things To Know About the Delta Variant. Yale Medicine. https://www.yalemedicine.org/news/5-things-to-know-delta-variant-covid

O'Toole, Hill, V., Pybus, O. G., Watts, A., Bogoch, I. I., Khan, K., Messina, J. P., Consortium, G. C. T. U. K., Network for Genomic Surveillance in South Africa (NGS-SA), Brazil-UK CADDE Genomic Network, Tegally, H., Lessells, R. R., Giandhari, J., Pillay, S., Tumedi, K. A., Nyepetsi, G., Kebabonye, M., Matsheka, M., Mine, M., . . . Kraemer, M. U. G. (2021, May 19). Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2. Welcome Open Research. https://wellcomeopenresearch.org/articles/6-121/v1

The Sex, Gender and COVID-19 Project | Global Health 50/50. (2021, July 2). Global Health 5050. https://globalhealth5050.org/the-sex-gender-and-covid-19-project/

Statista. (2021a, July 27). Number of SARS-CoV-2 Delta variant cases worldwide as of July 2021, by country. https://www.statista.com/statistics/1245971/number-delta-variant-worldwide-by-country/

Statista. (2021b, July 27). Number of SARS-CoV-2 Gamma variant cases worldwide as of July 2021, by country. https://www.statista.com/statistics/1246259/number-gamma-variant-worldwide-by-country/

Tracking SARS-CoV-2 variants. (2021, May 31). WHO. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/

Vignier, N. (2021, October). Breakthrough Infections of SARS-CoV-2 Gamma Variant in Fully Vaccinated Gold Miners, French Guiana, 2021. Emerging Infectious Diseases Journal. https://wwwnc.cdc.gov/eid/article/27/10/21-1427_article