Abstract

COVID-19 also known as Coronavirus has been the top health issue that the whole world faces for the past 12 months. This disease strikes at such an inscrutable tone that prompts health experts to race for a vaccine to put it to an end. This highly contagious disease has infected millions of people around the world, and it has taken the lives of thousands in a short interval of time. Its devastating effects across the globe have made it a very significant pandemic that raised a lot of concerns as to how to contain its spread. This health crisis has forced the whole world to adopt control measures that way they can control the spread of the virus. These measures have significantly impacted the way that humankind conduct businesses throughout the world. As a result, the strict control measures have taking a huge toll on the global economy that significantly stressed it to a point that make economists speculate that we are entering a depression. The COVID-19 pandemic is a serious and global public health concern. It is now well known that COVID-19 cases may result in mild symptoms leading to patient recovery, but parameters like severity of infection, fatality rates, and treatment responses across different countries, age groups, and demographic groups suggest that the nature of infection is diverse, and a timely investigation of the same is needed for evolving sound treatment and preventive strategies. In the early onset of the disease, drug like Hydroxychloroquine had been suggested as a possible treatment for COVID-19. Hydroxychloroquine has been around since the 1940s, and it has proven to be effective to treat or prevent malaria. Further, Hydroxychloroquine is also FDA-approved to treat autoimmune conditions such as chronic discoid lupus erythematosus, systemic lupus erythematosus in adults, and rheumatoid arthritis, but it was not found to be a good candidate to treat COVID-19 infected patients. It was rapidly rejected by the Food and Drug Administration (FDA) for COVID-19 outside of the hospital setting or a clinical trial due to risk of heart rhythm problems. In the race to contain the spread of COVID-19, the FDA approved Remdesivir also known as Veklury in October 2020 as a possible treatment for use in adult and pediatric patients 12 years of age and older and weighing at least 40 kilograms (about 88 pounds) for the treatment of COVID-19 requiring hospitalization. The medicine should only be administered in a hospital or in a healthcare setting capable of providing acute care comparable to inpatient hospital care, and it was the first treatment for COVID-19 to receive FDA approval. Remdesivir is not a very well-known drug. It is a nucleotide analogue prodrug that perturbs viral replication, originally evaluated in clinical trials to thwart the Ebola outbreak in 2014. Studies conducted by numerous virology laboratories demonstrated the ability of remdesivir to inhibit coronavirus replication, including SARS-CoV-2. Remdesivir had been proven to be beneficial to speed the recovery time of patients suffered from COVID-19 virus, but its benefit appeared much more limited in patients who needed mechanical ventilation. This limitation has made it a restricted drug to treat COVID-19. In early December, the FDA has approved a messenger RNA (mRNA)vaccine developed by Pfizer as a primary medicine to treat COVID-19. This vaccine seems to have shown the light at the end of the tunnel, and it appears to be a possible weapon that will end the pandemic.

For this project, I collected a dataset from data.gov that account for the number of COVID-19 cases by age distribution among Maryland residents. The data will be analyzed to investigate what role age plays in contracting COVID-19, which age group is less susceptible to COVID-19, and which age group is more prominent to contract COVID-19.

Introduction

COVID-19 is a disease that affects folks of all age group. Its mode of transmission is not very well understood. Some speculate that COVID-19 is transmitted from person to person when an infected person coughs, sneezes, or talks, droplets or tiny particles called aerosols carry the virus into the air from their nose or mouth. Anyone who is within 6 feet of that person can breathe it into their lungs and becomes infected if the person ingests a great deal of virus. However, the real method of transmission is not very well studied since the disease is new. The COVID-19 pandemics is a major challenge that left health officials around the world stranded to come up with measures that can contain it. Covid-19 does not discriminate against gender, race, age, and geographical position. Thus, it attacks folks of all genders, races, and ages around the globe, and it has already caused close to 76.4 million infections and more than 1.69 million deaths worldwide. This aggressive spread of COVID-19 has rushed scientists around the world to find a treatment and containment measures. To understand and combat such a large-scale outbreak, the treatment and containment strategies must also consider the diversity of infection, and potential differences such as the way a population or an age group responds to the many variants that might exist. Quite often, scientists conduct studies on the genetic diversity of viruses so that they can get insights on their genetic materials and their method of transmission. Since the beginning of the COVID-19 pandemic, several studies on the genetic diversity of these viruses have been performed. As of now, not a lot of information about the diversity of COVID-19 variants and the likely impact on clinical outcomes can be determine by the biomedical community while the transmission dynamics and mutation rates of viral genomes remain two factors of significant importance. At the time of writing this report, the fatality rates in the world range widely from as low as less than 4% to as high as more5%, and The United Stated account for 315 thousands of the 1.69 million deaths worldwide. The real reason behind these dramatically different mortality rates and possible differences in affected age groups are not very well understood. Therefore, it is imperative to collect and present accurate clinical data as well as elaborate epidemiological studies such that both current and future outbreaks like COVID-19 can not only be understood but also controlled. Patient age has been the center of many COVID-19 demographic studies as the key factors that can predict clinical outcomes. These clinical outcomes are included and not limited to fatality rate and severity of symptoms. Nevertheless, the mechanism in which a given viral variant like COVID-19 variant might impact different patient age groups has not been very well investigated. This makes it very hard to contain this horrible virus. Therefore, a detailed impact analysis of COVID-19 requires sequencing of thousands of samples across various ethnicities and age groups, which is yet to be attempted. In this project, an alternative approach to understand positive COVID-19 infections and their relationships to the age of patients in term of correlation is presented.

Methods

As of December 20, 2020, the coronavirus disease 2019 (COVID-19) pandemic had resulted in more than 17.6 million reported U.S. cases and more than 315 thousand associated deaths. COVID-19 incidence was highest among older adults early in the pandemic. During the months of May through August, The Centers for Disease Control (CDC) examined the changing age distribution of the COVID-19 pandemic in the United States by assessing indicators like COVID-19–like illness-related emergency department (ED) visits, positive reverse transcription–polymerase chain reaction (RT-PCR) test results for SARS-CoV-2, the virus that causes COVID-19, and confirmed COVID-19 cases. Reverse transcription-polymerase chain reaction is a laboratory method used to make many copies of a specific genetic sequence for analysis. It uses an enzyme called reverse transcriptase to change a specific piece of Ribonucleic Acid (RNA) into a matching piece of Deoxyribonucleic Acid (DNA). This piece of DNA is then amplified (made in large numbers) by another enzyme called DNA polymerase. The amplified DNA copies help tell whether a specific messenger RNA (mRNA) molecule is being made by a gene. This highly advance Ribonucleic Acid (RNA) technique has very useful purposes in activation of certain genes, which may help diagnose a disease, such as Coronavirus and cancer. It may also be used to study the RNA of certain viruses, such as the human immunodeficiency virus (HIV) and the hepatitis C virus to help diagnose and monitor an infection. This technique is also called RT-PCR. In COVID-19 diagnostic testing, RT-PCR is used to detect the presence of the virus in the patient's body. In a single-stranded RNA, the genome of SARS-CoV-2 is encoded. Therefore, an appropriate nucleic acid preparation method must be used so that all RNA in the patient's swab sample can be extracted. Once the RNA has been extracted, the remaining RNA molecules that have been left isolated are then converted to complementary DNA (cDNA) using the reverse transcriptase enzyme.

Besides investigating the changing age distribution of the COVID-19 pandemic in the United States, it is also critical to examine how this pandemic affects different age groups. It is not a secret that some specific age groups are more vulnerable to COVID-19 compared to other ones. The main reason laid in the difference in the immune system. For instance, a health 20-year-old person immune system is stronger compared to a 60-year-old. The risk of developing severe COVID-19 complications or dying increases dramatically with age. About 79% of COVID-19 infections reported by August 29, 2020 occurred at ages 65+, and the population infection rate (the risk of infection from COVID-19 among the general population) dramatically increased by age. Infection rates are always the highest among older people, across countries with varying levels of COVID-19 mortality and infection. This appears to result from the fact that higher age is linked to more underlying health conditions and weaker immune systems. Thus, how should older people respond to the COVID-19 pandemic? Older folks are encouraged to engage in more preventive behaviors and avoid more risky behaviors. Some earlier research relevant to epidemics has suggested that being older is related to greater compliance and participation in practicing precautionary behaviors, not all results are consistent in finding a strong age difference.

This project is based on crowd-sourced data on COVID-19 patients in Maryland compile at https://catalog.data.gov/dataset/md-covid-19-cases-by-age-distribution. It is a collection of positive COVID-19 test results that have been reported each day by the local health department

via the ESSENCE system. The collection provides positive COVID-19 cases among Maryland residents by age: 0-9; 10-19; 20-29; 30-39; 40-49; 50-59; 60-69; 70-79; 80+; Unknown. The resource which is a compilation of daily updates provided by the Government officials from throughout Maryland attempts to present crucial information in a unified format; most importantly different patient age groups or age distribution assigned by the data administrators. It is a csv raw data file that contains data for about 252 patients combined into 10 different age groups, and it is the source of data used in this project. This data published on July 31, 2020 was updated on December 16, 2020, and it is likely to change again prior to the completion of this project since COVID-19 data fluctuate at a very high rate.

Results

Table (1) provides a first assessment of the patient age distributions within each group of patients. Table (1) shows a clear difference between the overall distribution of patients from all the different age groups of positive COVID-19 cases. Table (2) shows a more detailed distribution, which suggests that COVID-19 does not discriminate against age group when infections is taking into perspective.

Figure (1) is a clustering over age histograms. It provides a good, detailed distribution, which suggest that even though the average age values place age group Age_40_to_49 closer to Age_20_to_29 and Age_80plus, individual histogram bins suggest the case of Age_40_to_49 to be more complex, and a rigorous clustering over age histograms suggest that Age_40_to_49 clusters better with Age_60_to_69 groups. In figure (2) I explored a detailed correlation between the age groups displayed in the previous histograms. The figure shows how the patients from different age group infection correlate in terms of their age histograms. Clearly, the Age_40_to_49 group, which looked like Age_60_to_69 in terms of average patient age, tended to correlate closer to Age_20_to_29, which suggests that younger people are being infected just like older folks are. Thus, Coronavirus infections do not discriminate against age group.

Table 1

	OBJECTID	DATE	Age_0_to_9	Age_10_to_19	Age_20_to_29	\
0	1	3/29/2020 10:00	4	21	165	
1	2	3/30/2020 10:00	5	27	188	
2	3	3/31/2020 10:00	6	32	215	
3	4	4/1/2020 10:00	6	39	252	
4	5	4/2/2020 10:00	9	46	286	

	Age_30_to_39	Age_40_to_49	Age_50_to_59	Age_60_to_69	Age_70_to_79	\
0	203	241	243	191	117	
1	247	262	274	219	139	
2	290	304	335	260	157	
3	345	372	396	309	187	
4	410	437	471	357	217	

	Age_80plus	Age_Unknown
0	54	NaN
1	53	NaN
2	61	NaN
3	79	NaN
4	98	NaN

Table 2

Describe Data

	OBJECTID	Age_0_to_9	Age_10_to_19	Age_20_to_29	Age_30_to_39
count	251.000000	251.000000	251.000000	251.000000	251.000000
mean	126.000000	3150.760956	6567.358566	15732.609562	16291.123506
std	72.601653	2272.537924	5118.562940	10484.511429	9291.781336
min	1.000000	4.000000	21.000000	165.000000	203.000000
25%	63.500000	1263.500000	2291.500000	7249.500000	9743.000000
50%	126.000000	2992.000000	5536.000000	15427.000000	16837.000000
75%	188.500000	4737.000000	10819.500000	23946.000000	22814.500000
max	251.000000	8785.000000	18587.000000	39694.000000	37254.000000

	Age_40_to_49	Age_50_to_59	Age_60_to_69	Age_70_to_79	Age_80plus
count	251.000000	251.000000	251.000000	251.000000	251.000000
mean	14705.876494	13418.685259	9194.601594	5580.756972	4630.358566
std	7960.260718	7324.048222	4768.866145	2693.351309	2002.324882
min	241.000000	243.000000	191.000000	117.000000	53.000000
25%	9454.000000	8540.000000	6159.500000	3959.000000	3736.500000
50%	15095.000000	13371.000000	9281.000000	5748.000000	5078.000000
75%	20151.000000	18454.500000	12340.000000	7358.000000	5900.000000
max	32739.000000	31000.000000	20549.000000	11919.000000	8664.000000

	Age_Unknown
count	235.0
mean	0.0
std	0.0
min	0.0
25%	0.0
50%	0.0
75%	0.0
max	0.0

Summarized Data

	DATE
count	251
unique	251
top	8/5/2020 10:00
freq	1

Figure 1

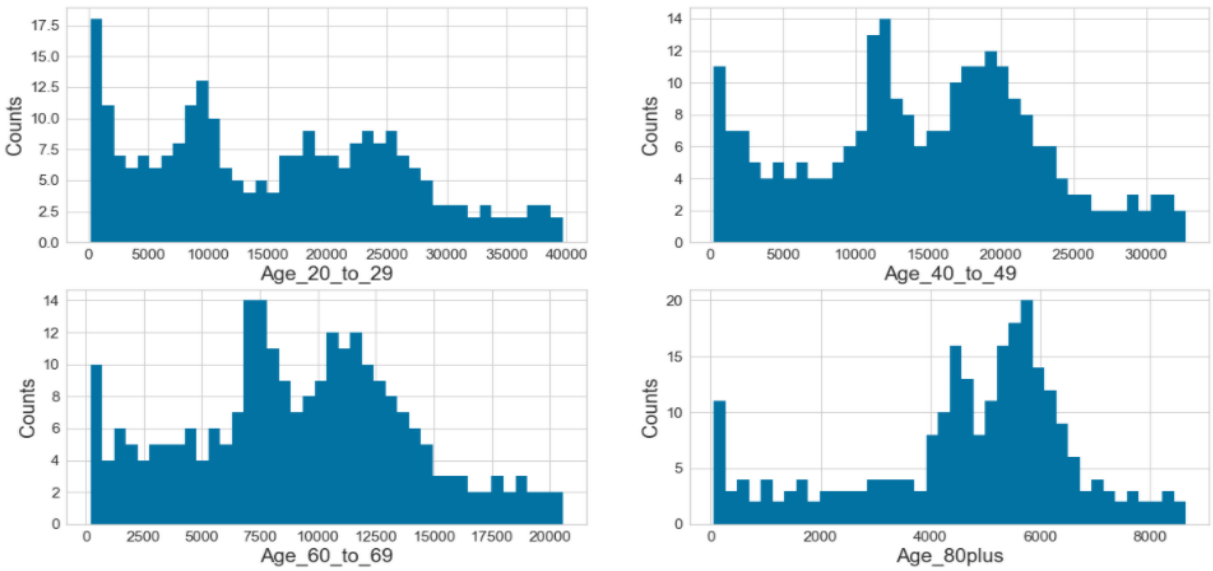
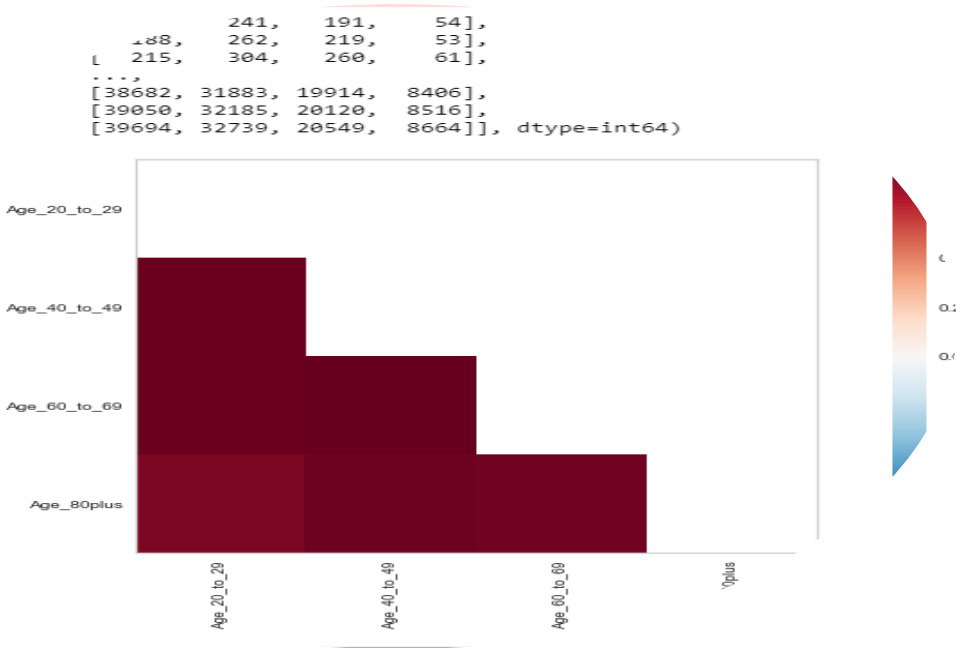


Figure 2



Discussion

There are many COVID-19 related compilations of data in the public domain emerging from agencies as varied as World Health Organization (WHO) and state governments. Infection origin studies like this could have been carried out on group specific data where all infections are more or less of the same origin. However, in such a scenario, the age profile differences observed between different conditions would be further confounded by additional factors such as different age distribution in the general population age variations, and the time of testing.

Research have shown that the COVID-19 pandemic in the United States affected a larger proportion of younger persons during the months of June-August compared to January-May 2020, and all four US census regions have seen the shift towards younger ages. This shift occurred regardless of changes in incidence during this period, and it was reflected in COVID-19–like illness-related emergency departments (ED) visits, positive SARS-CoV-2 RT-PCR test results and confirmed COVID-19 cases. During January-May, Europe experienced a similar age shift in which the median age of COVID-19 cases declined from 54 years to 39 years compared to June–July. In this shift, persons aged 20–29 years constituted the largest proportion of cases. Case and laboratory surveillance are based on consistent availability of diagnostic testing to all segments of the population. It is often observed that changes in testing across age groups have some significant effects on the age distribution of positive SARS-CoV-2 test results and confirmed cases. It is unlikely that the observed age shift resulted solely from changes in testing availability since testing availability has varied by place, time, and test provider.

Conclusion

The study established that focus on correlation of positive COVID-19 infections by age distribution has shown a close correlation pattern among patient age profiles. Out of the 8 of

such group, 4 clusters were targeted, and their correlation patterns were studied. All 4 clusters have a strong positive correlation coefficient. This indicates that they are closely related to each other. Thus, the COVID-19 virus infects all age group in such a way that it does not discriminate against age groups.

Questions

1-What is the mode of transmission of COVID-19?

2-How does COVID-19 impact different age group?

3-Does COVID -19 affect older folks at a higher rate compared to younger ones?

4-What role does age play in contracting COVID-19?

5-What age group is more vulnerable to Coronavirus infections?

6-Why younger adults are less likely to be severely ill compared to older persons?

7-What role does the immune system of a person play in combatting the Coronavirus?

8-Why COVID-19 infections are higher in certain age groups compared to others?

9-Why does COVID-19 affect children at a lower rate?

10-Which age group does COVID-19 target the most?

Appendix 1: COVID-19 Cases by Age Distribution for the state of Maryland

This appendix describes various groups of COVID-19 cases by age distribution for the state of Maryland, including the process the Centers for Disease Control (CDC) uses to examine the changing age distribution of the COVID-19 pandemic in the United States, and then gives details of the crown-sourced data that will be used to complete the project. See the main text above for a description of the methods, analysis, and results.

References

1-Jung Ki Kim, Eileen M. Crimmins: How does age affect personal and social reactions to COVID-19: Results from the national Understanding America Study, November 10, 2020

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0241950#:~:text=About%2079%25%20of%20COVID%2D19,55%E2%80%9364%2C%200.02%25%20for

2-Matthew Herper: Covid-19 study details benefits of treatment with remdesivir, and also its limitations, May 22, 2020

https://www.statnews.com/2020/05/22/covid-19-study-details-benefits-of-treatment-with-remdesivir-and-also-its-limitations/

3-Mahase E. COVID-19: death rate is 0.66% and increases with age, study estimates. BMJ. 2020

4-Richard T. Eastman, Jacob S. Roth, Kyle R. Brimacombe, Anton Simeonov, Min Shen, Samarjit Patnaik, and Matthew D. Hall - Remdesivir: A Review of Its Discovery and Development Leading to Emergency Use Authorization for Treatment of COVID-19 - ACS Cent Sci., May 27, 2020

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7202249/

5-FDA News Release: FDA Approves First Treatment for COVID-19, October 22, 2020

https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19

6-US Food and Drug Administration: FDA's approval of Veklury (remdesivir) for the treatment of COVID-19—The Science of Safety and Effectiveness, October 22, 2020

https://www.fda.gov/drugs/drug-safety-and-availability/fdas-approval-veklury-remdesivir-treatment-covid-19-science-safety-and-effectiveness

7-Wang C., Liu Z., Chen Z., Huang X., Xu M., He T. The establishment of reference sequence for SARS-CoV-2 and variation analysis. J Med Virol. 2020 [PMC free article] [PubMed]

8-https://coronavirus.app/map?mode=infected

9- https://www.who.int/emergencies/diseases/novel-coronavirus-2019