

Full Report

Shan

11/23/2020

Contents

Introduction	1
Methods	1
Data background	1
Data analysis	2
Result	2
Conclusion	5

Introduction

It is obvious that the Covid-19 is becoming a pandemic issue for the world, especially for the U.S. Also, the States is one of the most racially and ethnically diverse nations in the world. Therefore, I'd like to look into the association between the Covid-19 and races in the U.S. This report is based on the key question: Among population with known race category across the U.S, what races were being affected most by Covid-19? Three common epidemiology indicators were being analyzed respect to the key question in this report.

Methods

Data background

The primary data source is the COVID Racial Data Tracker. They collect, cross-check, and publish COVID-19 data from 56 US states and territories in three main areas: testing, patient outcomes, and, via The COVID Racial Data Tracker, racial and ethnic demographic information. Data is being updated twice per week. The KFF data: Population Distribution by Race/Ethnicity (2019) was being utilized, too. The KFF.org data is a non-profit foundation that does research, journalism, and communications programs. In addition, the AIAN is short for American Indian and Alaska Native, and the NHPI is short for Native Hawaiians and Pacific Islanders.

Table 1: Overall table

	cases_number	deaths_number	total_population	case_fatality_rate(%)	prevalence(%)	death_rate
AIAN	84581	2031	2125700	2.40125	3.97897	0.95545
Asian	169969	8125	18007200	4.78028	0.94389	0.45121
Black	1030625	44797	38872200	4.34659	2.65132	1.15242
LatinX	1312430	42774	8856100	3.25915	14.81950	4.82989
Multiracial	45976	557	59102900	1.21150	0.07779	0.00942
NHPI	24943	332	528200	1.33103	4.72226	0.62855
White	3334324	118636	191757100	3.55802	1.73883	0.61868

Data analysis

Given the data above, First we removing the NA's. Also, excluded the "other" category in COVID data since there is no correspond variable in the population data. Then We summed up the case number and death number of each known races by the latest updated numbers. Prevalence is the proportion of a population who have a specific characteristic in a given time period. Therefore, we use case number divided by the total number of population to get the prevalence for each known races. Death rate is a measure of the number of deaths (in general, or due to a specific cause) in a particular population, scaled to the size of that population, per unit of time. We use death number divided by the total number of population to get the death rate for each know race. Case fatality rate (CFR) is the proportion of deaths from a certain disease compared to the total number of people diagnosed with the disease for a particular period. So we use death number divided by case number for each known race. All rate are multiplied by 100. The time period is from 04/12/2020 to 08/11/2020 for all the calculated rates.

Result

The percentage of reported COVID cases and deaths for each known race

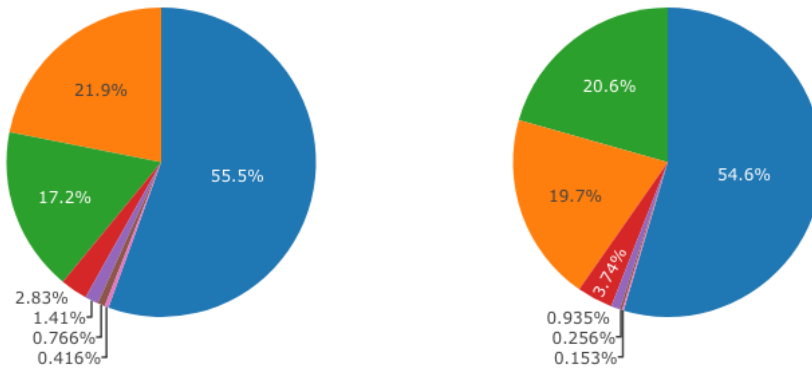


Figure 1: Pie chart of absolute number for case and death

These pie charts were built based on absolute number(fig.1). We can see that for both case and death number, white contributed more than a half to the total number of them. Differently, for the case number,

the proportion of latinX is slightly larger than the black, which is opposite on the death count. For both case and death, the white, black and latinoX contributed most to the total number.

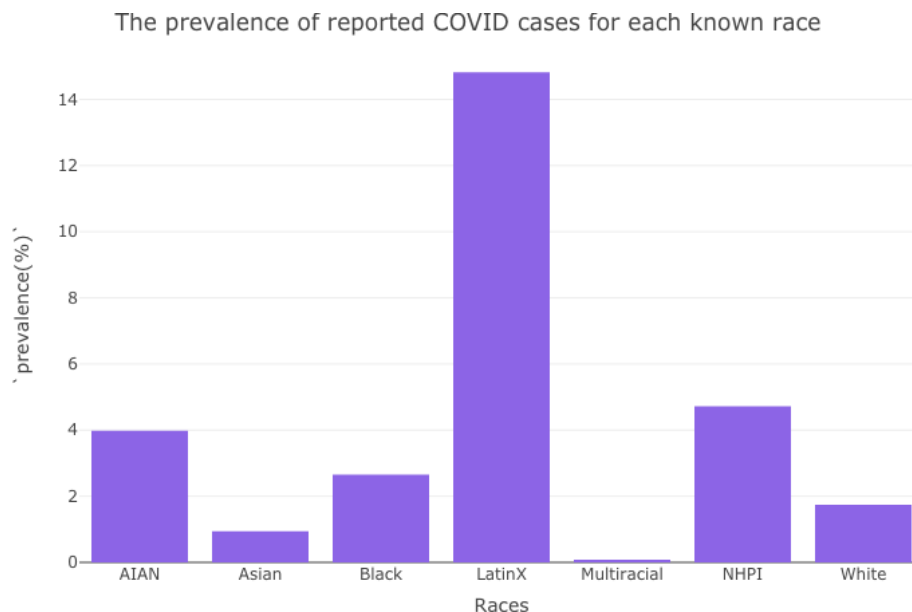


Figure 2: Bar chart for prevalence

From the prevalence bar plot(fig.2), we can tell that the latinX has the highest prevalence: 14.8%, which is more than twice of the second high race: the NHPI(4.7%). The AIAN race has the third high prevalence(3.9%). Those are the three races which have relative high prevalence comparing with other races.

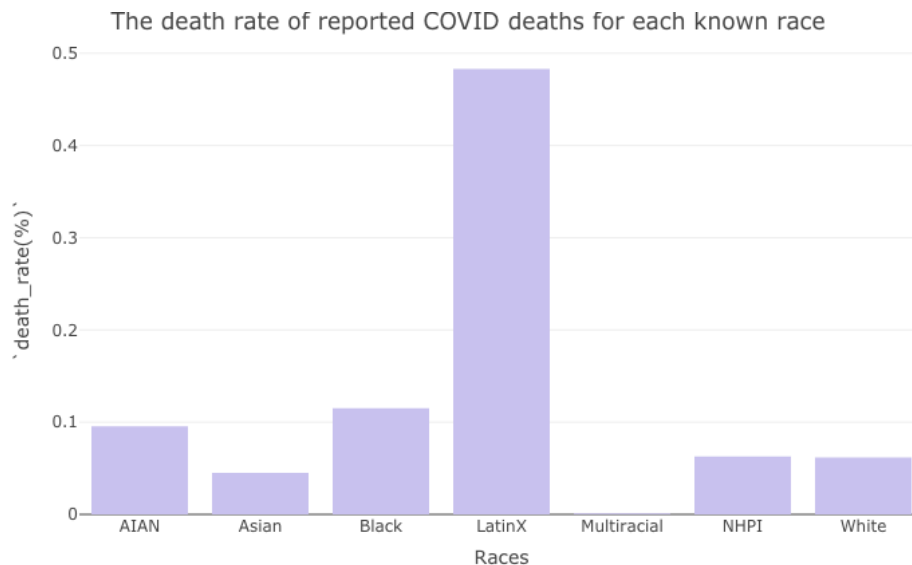


Figure 3: Bar chart for death rate

From the death rate bar plot(fig.3), we can see that as same as the prevalence, the latinX also has the highest death rate(0.48%), it is triple of the second high death rate: the black(0.11%). The AIAN race has the third high death rate(0.09%). The NHPI and white has very similar death rate(0.06%).

Interestingly, in the case fatality rate(fig.4), the top three high CFR are related with asian(4.7%), black(4.3%) and white(3.5%). This order is quite different with any previous ones. Take latiX for example, though it

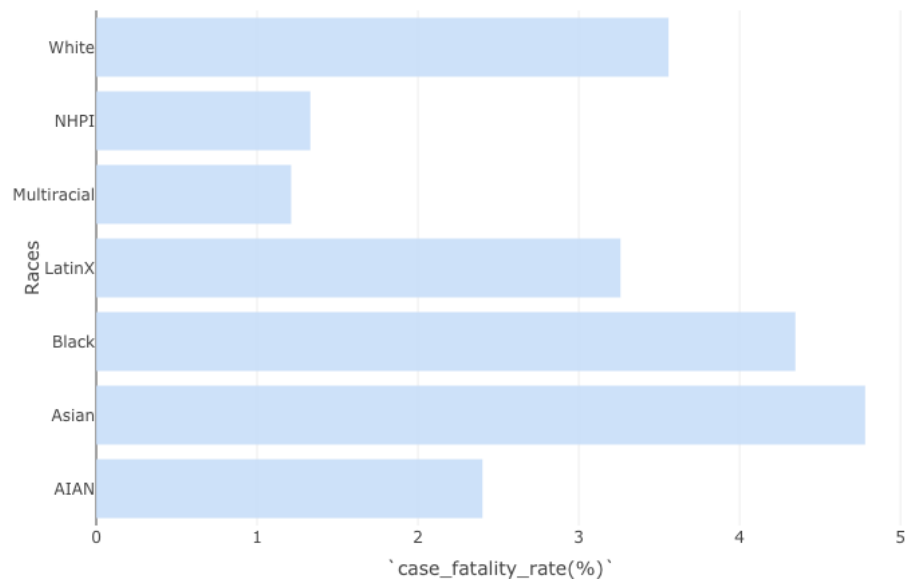


Figure 4: Horizontal bar chart for case fatality rate

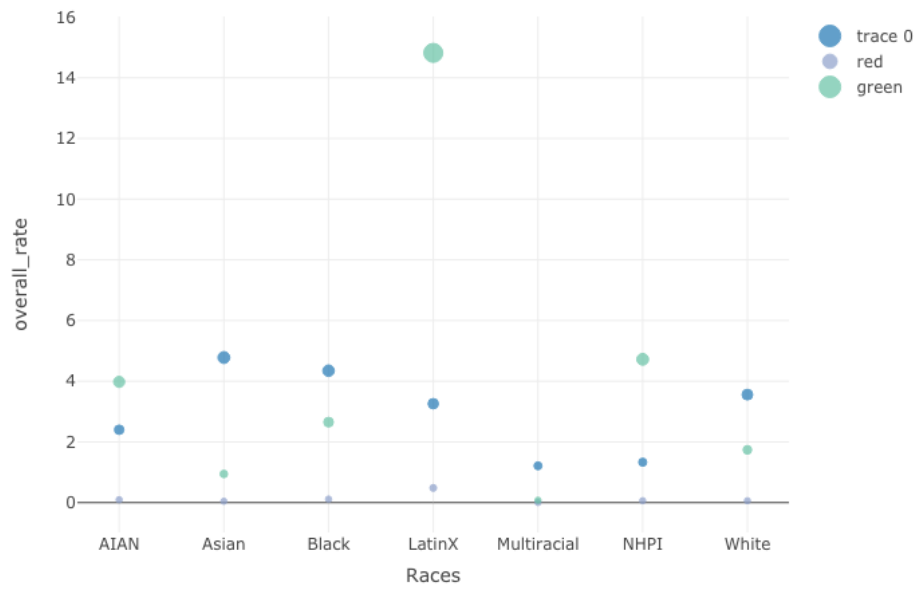


Figure 5: Scatter plot for three epidemiology indicators

rank first in both prevalence and death rate, it only rank forth in the case fatality rate(3.2%). I think it could have something to do with people of different race's immune system. It might be a good research question for health care scholars. To make the compare more obvious, here is a scatter plot for those three rates. The size of the scatter indicate the rate of it.

Conclusion

Given the prevalence and the death rate results, there are four races that been affected most by COVID: LatinX, Black, AIAN and NHPI, which are minorities in the States that need more taken care of.

Data Souces:

2019 races number data:<https://www.kff.org/other/state-indicator/distribution-by-raceethnicity/?dataView=1¤tTimeframe=0&selectedRows=%7B%22wrapups%22:%7B%22united-states%22:%7B%7D%7D%7D&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D> Covid data:<https://covidtracking.com/race/about#download-the-data>