

Reducing the Risk of COVID-19 Infections in Vulnerable ages*

Mohid Sharif

04 February 2022

Abstract

COVID-19 has taken over the world for more than two years now and it is important to study and understand in which circumstances cases rise and fall. I studied and analyzed data obtained about COVID-19 cases from various communities in Toronto specifying age, gender, and source of infection of the patient. I found that younger age groups are more likely to be infected in their household while the elderly are more likely to be infected in community settings. My findings can be used to identify circumstances in which outbreaks are less likely for certain age groups.

1 Introduction

COVID-19 has taken over the world for more than two years now. Thousands of cases have been recorded in this data set and we can use this to our advantage. Analyzing the data set we can compare variables such as age groups and gender to isolate sources of infection which are more likely to infect certain groups and find changes to those circumstances to reduce the likeliness of infection in those groups. Using this technique we can keep making adjustments to what should and should not be allowed during the pandemic to ensure cases don't rise uncontrollably.

I first analysed the data set to understand how cases had been recorded. Every case in the data set records the source of infection, age group, and gender of the patient. These are only a few of the variables but they are also the most important and they are the ones I mainly looked at in this data study. I then created two graphs to compare variables in the data set. The first graph is a comparison of source of infection between age groups, I used these graphs to visualize which age groups were more likely to be infected in certain settings. I found that younger ages (0-19 years) were more likely to be infected in households while elderly people are more likely to be infected in community settings or healthcare institutes. The second graph compares the source of infection between gender groups, I used these graphs to analyze which gender groups were more likely to be infected in certain settings. I found that both males and females had almost identical data for how they were being infected, however there were twice as many females being infected in healthcare institutes.

I can use my data to isolate situations in which infections are more likely for certain groups and how we can change that to reduce infections. Firstly younger ages are more likely to get infected in their homes, in this case they are likely infected by a family member. Family members should in this case sanitize regularly while entering and exiting the house and reduce outdoor activity to necessities. Secondly elderly are more likely to be infected in healthcare institutes because of regular visits for their health. For this healthcare institutes should do their best to isolate the non-infected elderly from the infected reducing their risk for infection. Lastly females are more likely than men to be infected in healthcare institutes due to the fact that women dominate nursing, otherwise both males and females are just as likely to be infected in any setting.

The remainder of this paper is: Section 2.

*Code and data are available at: https://github.com/MohidSharif/COVID19_data_paper

2 Data

This data set records COVID-19 cases across Toronto and records the details of the infected. The data records the age group, gender, neighborhood, source of infection, and if the infection was outbreak associated or sporadic. The data classifies cases by age group rather than specifics, age groups are grouped into 10 year age groups (E.g. “20 to 29 Years”) except for the first group which is “19 and younger” and the last group which is “90 and older.” Gender is self-reported and sorted into 7 categories, from which we will mainly look at males and females.

Source of infection is determined by three factors:

- A public health investigator’s assessment of the most likely source of infection
- Being associated with a confirmed COVID-19 outbreak
- Reported risk factors such as contact with a known case or travel

The classifications are very self explanatory categorized as household contact, close contact, community, outbreak associated (this is then further classified into 3 categories, congregate settings, healthcare institutions, and other settings), and travel.

(Table 1) shows the data recorded from the neighborhood of Agincourt North. Most cases in this data set are classified as “No Information” or “Pending” under source of infection, these cases need to be discluded to see cases which give us more valuable information.

Table 1: Outbreaks in Agincourt North

Source of Infection	Age Group	Gender
Travel	30 to 39 Years	FEMALE
Outbreaks, Other Settings	50 to 59 Years	FEMALE
Outbreaks, Congregate Settings	20 to 29 Years	MALE
Community	50 to 59 Years	MALE
Outbreaks, Other Settings	30 to 39 Years	MALE
Outbreaks, Healthcare Institutions	30 to 39 Years	FEMALE
Close Contact	30 to 39 Years	FEMALE
Outbreaks, Healthcare Institutions	19 and younger	MALE
Outbreaks, Healthcare Institutions	50 to 59 Years	FEMALE
Outbreaks, Healthcare Institutions	20 to 29 Years	FEMALE

(Figure 1) shows the relationship between age and source of infection.

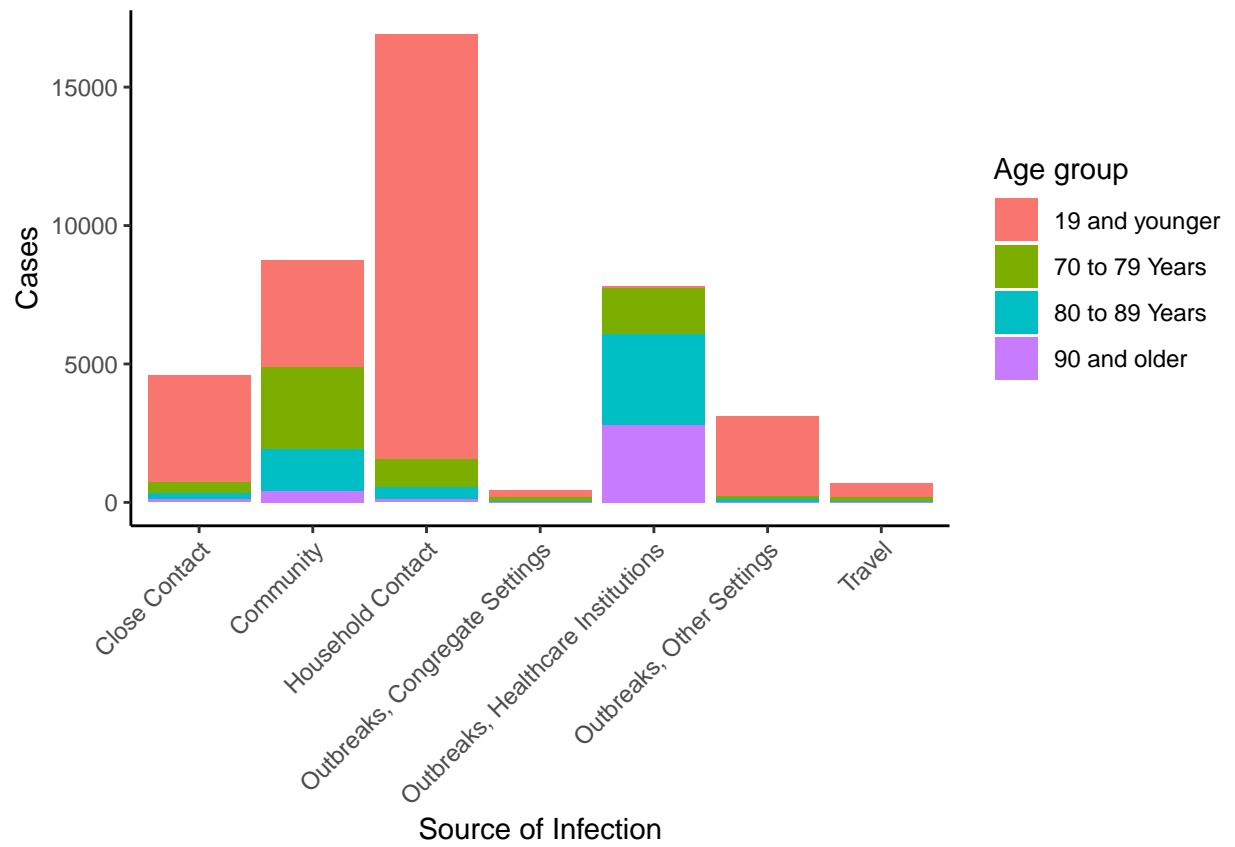


Figure 1: Source of Infections by Age group

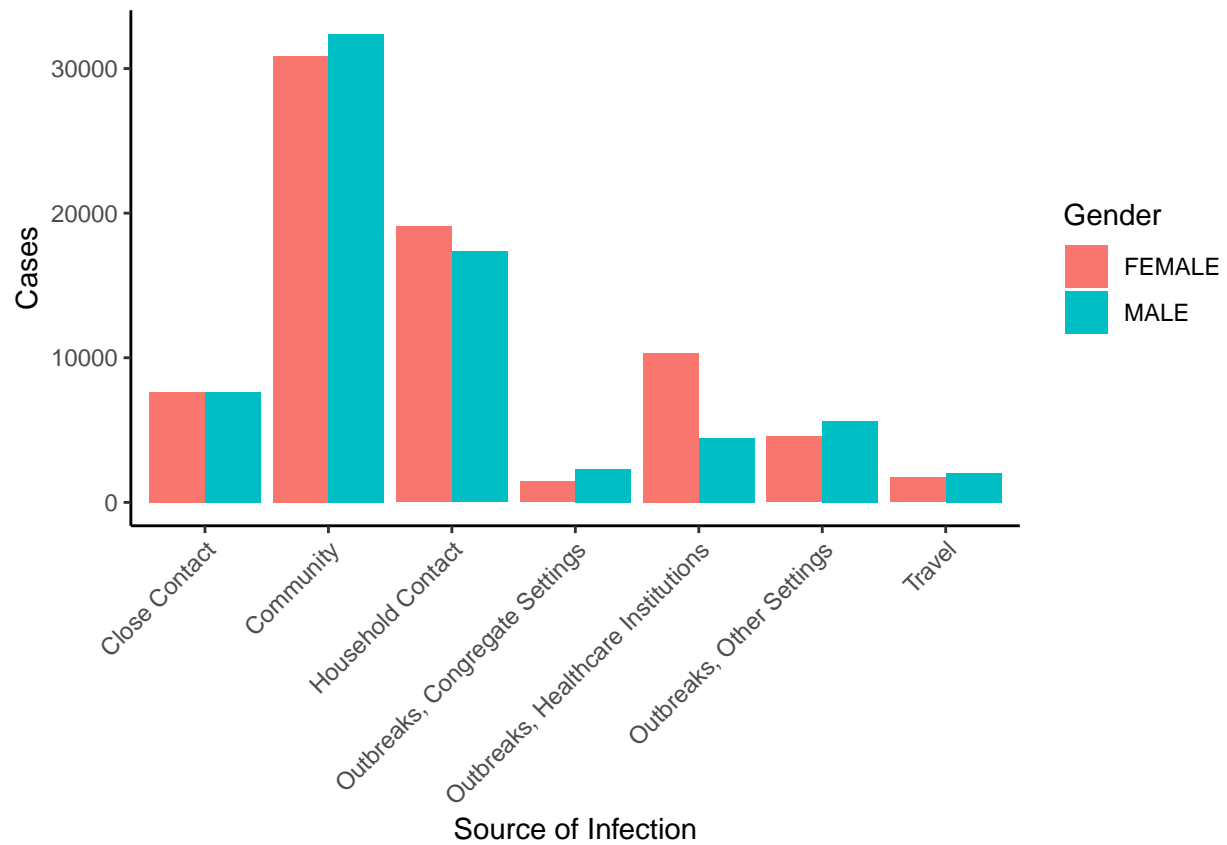


Figure 2: Source of Infections by Gender

Talk more about it.

3 References