COVID-19 Case Count Numbers in Toronto Increasing but the Cases are Becoming Less Severe*

Charles Lu

2022-02-06

Abstract

An observational report on the COVID-19 Cases in Toronto, examining the case count numbers and case severity over infection date, gender, age group and source of infections. Across all four epidemic waves in Toronto, the peak case count numbers significantly increases every wave, but proportionally less cases are hospitalized every wave. Moreover, female and male mostly have equal amount of case numbers, older population have less case numbers but more severe cases and community contact is the most common source of infection. However, the cases provided from the data are likely to be underreported in the context of selective reporting, asymptotic, and Omicron variant.

1 Introduction

In the context of an evolving global pandemic, Toronto Public Health constantly updates data regarding COVID-19 cases to provide demographic, geographic, and severity information for all confirmed and probable cases reported. The case numbers are based on the number of COVID-19 cases reported to local public health departments directly from the laboratory, demonstrating high credibility. However, there might be bias upon the reported data since the data is only collected through reported COVID-19 cases in Toronto. This could be due to factors such as selective reporting, difficulty to collect data for asymptotic cases, and shortage of availability to confirmatory PCR testing driven by increasing COVID-19 cases related to the Omicron outbreak. Thus, the case count in this report represents an underestimate of the true number of population who are infected by COVID-19 in Toronto.

According to this dataset, the reported COVID cases in Toronto are increasing as the peak of case numbers gets higher every epidemic wave. However, the severity of the virus are decreasing over time, possibly due to the implementation of vaccination and the nature of the less deadly Omicron variant.While the overall case numbers between females and males are similar, statistics have shown that there are more females infected than males in the last two months, December 2021 and January 2022, possibly due to the different behavior of omicron. By comparing the severity of COVID-19 cases among different age groups, we can see that infections with older populations are likely to be more severe but have significantly less case numbers than younger populations. Lastly, by sorting the COVID-19 data by sources of infections, we see that community is the most common source of infection, while travel is the least common among the reported cases.

2 Data

This report utilizes analytic programming language R (R Core Team 2020). R packages tidyverse (Wickham et al. 2019), janitor (Firke 2021) and dplyr (Wickham et al. 2021) are used to clean, organize and manipulate

*Code and data are available at: https://github.com/R300G/COVID19_Cases_Data

the data. Furthermore, R packages ggplot2 (Wickham 2016), scales (Wickham and Seidel 2020) and knitr (Xie 2021b) are used to create figures and tables. R packages bookdown (Xie 2021a) and tinytex (Xie 2021c) are used to generate the R markdown Report.

This report conducted an analysis on the data regarding COVID-19 cases in Toronto (Toronto Public Health 2022). The data collection provides demographic, geographic, and severity information for all confirmed and probable cases reported to and managed by Toronto Public Health. The data is taken from the provincial Case and Contact Management System (CCM). Since the first case has been reported in January 2020, the dataset has been made available to the public through the City of Toronto Open Data Portal (Toronto Open Data 2022). The COVID 19 dataset analyzed in this report was downloaded in csv format from the City of Toronto Open Data Portal. The data will be thoroughly updated and rewritten on a weekly basis, with the data being retrieved at 8:30 a.m. on Tuesdays and uploaded on Wednesdays (Toronto Public Health 2022).

The case numbers are based on the number of COVID-19 cases reported to local public health departments directly from the laboratory. All data are primary, and they will constantly be updated when new information became available and quality assurance work is completed (City of Toronto 2022b).

When Toronto Public Health receives reports of COVID-19 cases, the officers first need to determine whether the case meets the provincial case definition. After confirming that the provincial case definition is met, the officers act immediately to conduct a thorough investigation of the case (City of Toronto 2022b). Thus, the credibility of this dataset is high. The data collected is compiled to highlight tendencies.

The population of this data are all cases in Toronto that meet the provincial case definition. This includes confirmed cases, probable cases, and cases of reinfection. The confirmed case includes population who have been tested positive by laboratory-based nucleic acid amplification test (NAAT)-based assay such as real-time PCR or nucleic acid sequencing performed at government-recognized community, hospital, or reference laboratory. The probable case includes population who had symptoms compatible with COVID-19 or had high-risk exposure with a confirmed case of COVID-19 and their NAAT-based assay has not been completed or is inconclusive. Reinfection indicates previously cleared confirmed cases of COVID-19 that have a subsequent confirmed infection with COVID-19(Ontario Ministry of Health 2022). There is no sampling in this data because the purpose of the data is to collect the entire infected population in Toronto.

While the credibility of this data is high, there may exists statistical bias and may inaccurately represents the current pandemic situation in Toronto, since the data only collects reported COVID-19 cases in Toronto. For instance, population who refuse to complete the COVID-19 test or intentionally hide Covid symptoms to Health Officials due to their conflict of interest will not be in the record of the data. According to reports, travelers were faced heavily fine for providing fake COVID-19 pre-departure test results at Toronto Pearson International Airport (Tsekouras 2021). In addition, studies suggest that at least 1 in 3 COVID-19 infections are asymptomatic (Somos 2021). Asymptomatic patients are infected population with no symptom which makes it almost impossible for the city of Toronto to track and record these cases. Moreover, upon its discovery in November 2021, the Omicron variant has quickly become the city's dominant COVID-19 variant. The Omicron variant is substantially more contagious than all other prior forms, and the province is witnessing a considerable increase in cases(Public Health Ontario 2021). The demand for diagnostic testing has risen dramatically as a result of the rising cases. This has made it more difficult for population who have symptoms or have been in close contact with a confirmed case of COVID-19 to get tested. Since data on the confirmed cases is only collected with regards to PCR results, it is likely that it is under-reported due to a shortage of availability to confirmatory PCR testing following a positive rapid antigen test result (Public Health Ontario 2021).

From the raw data set, the attributes that I select which I think are important and interesting to conduct analysis on are , and . I have also selected and decide to not use as the infection date represents a more accurate report of a COVID-19 case in my opinion. I have also created a new attribute by combining three attribute , , as all three of them have yes and no values and all represents the severity of the cases.

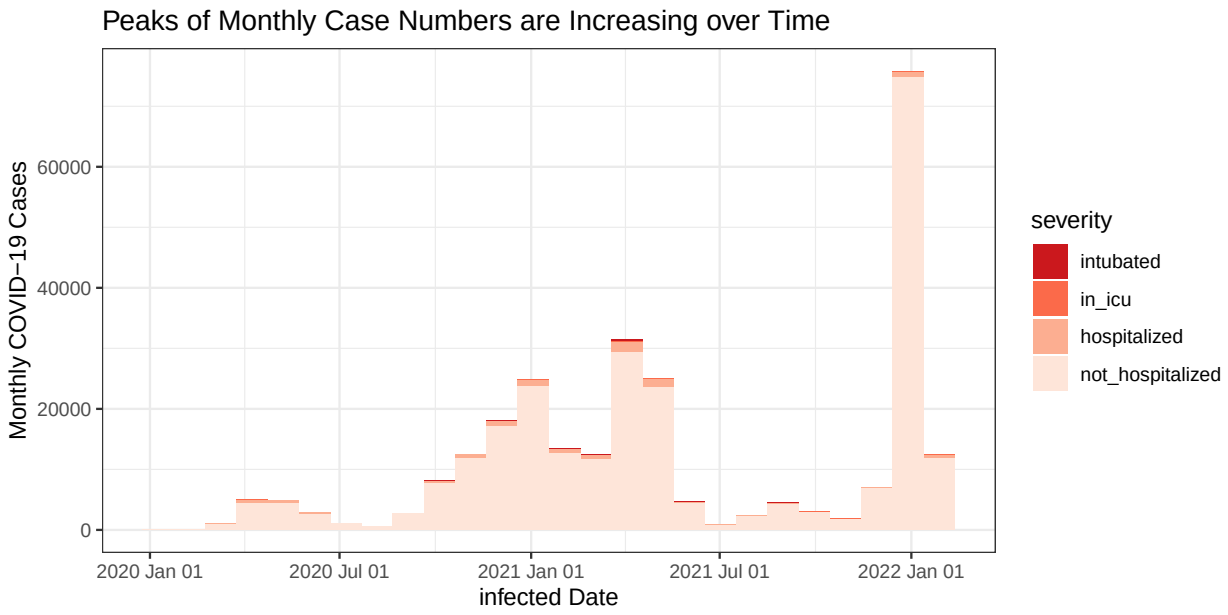

Figure 1: Monthly Number of COVID-19 Cases in Toronto by Severity and over time

Figure 1 shows the monthly numbers of COVID-19 infected cases in Toronto from January 1, 2020 to February 1, 2022 while showcasing the severity with different type of red. The Toronto Public Health officials have split and categorized the pandemic into four epidemic waves (City of Toronto 2022b). The first ever wave of COVID-19 hit Toronto during January 21 to July 31, 2020, Figure 1 showcased the wave with a small peak of 5000 monthly cases during the specified time period. The second wave came during August 1st, 2020 to February 20, 2021. Significantly higher case numbers than the first wave, the second wave peaked at around 25000 cases during January 2021. The third wave came during February 21 to July 10th, 2021, shortly after the second wave peak, this wave experienced an even higher peak at around 31000 cases at April 2021. Lastly, the fourth wave's timeframe was set at July 11th, 2021 to the last day of the available data February 1st 2022. Due to the discovery and outbreaks of highly infectious Omicron variant (Public Health Ontario 2021), the wave reaches a historic high peak of around 76000 cases, three times the size of third wave's peak. Overall, a trend can be easily noticed, the peak of infected numbers get higher every wave. Due to limited portions of infected population were severe enough to be hospitalized, it is hard to read the severity filling in Figure 1 on the scale of monthly case numbers. Thus, Figure 2 shows the monthly proportion of severity on a scale of 1 over time.

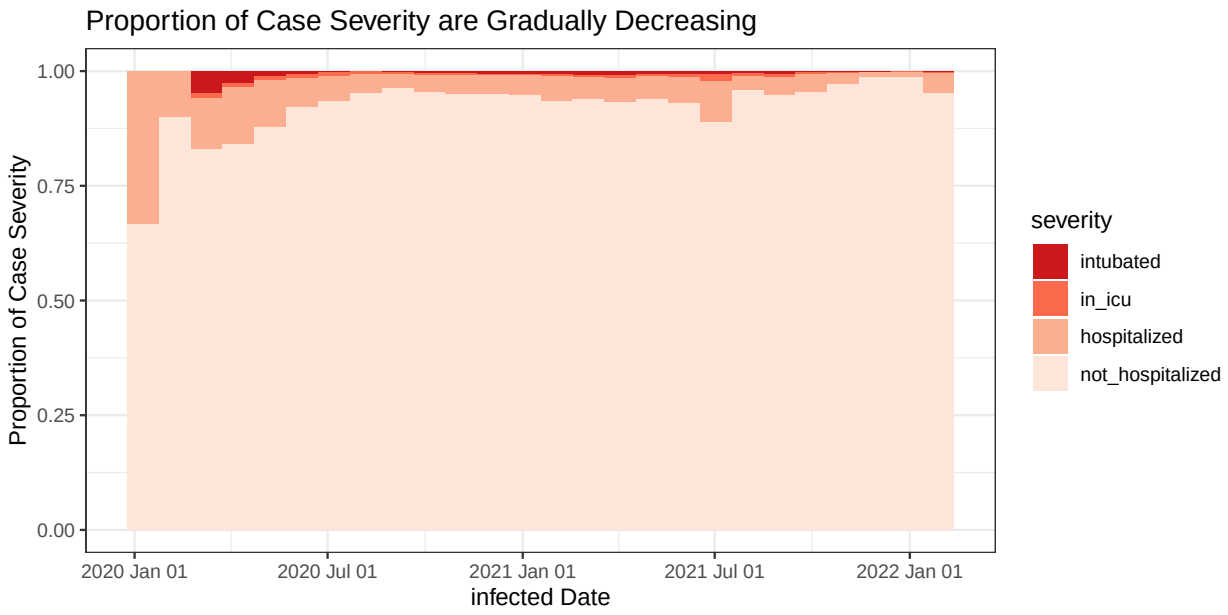

Figure 2: Proportion of Severity of COVID-19 cases in Toronto over time

Although majority of the COVID-19 cases were not hospitalized across the timeline, there were a big portion of hospitalized cases peaked at the start of the pandemic, January 2020. I think this may be due to fact that this is during the very early stage of the pandemic, and it is more likely to be hospitalized as doctors and health officials were still trying to observe and study this new virus. It may also because January 2020 have the lowest case numbers, which indicates the hospitals are more likely to accept COVID-19 patients due to its higher vacancies. As pandemic went on, cases that were severe enough to be hospitalized or more were proportionally less over time. I think this is mainly due to the wide spread implementation of Covid vaccine, which it not only reduces one's possibility of getting infected with the virus, but also greatly reduces the severity of the virus (City of Toronto 2022a). The reduced proportion may also be explained by overwhelmingly higher peak case numbers over time, as hospitals will more likely to be over full capacity during each wave, thus may raise the bar of hospitalizing Covid patients to accept the ones that needed more. Lastly, the Omicron variant of the virus spreading in wave 4, has proven to causes less severe diseases (Goethe University Frankfurt 2022), thus we can see that there are significant less proportion of hospitalized cases towards the ends of the plot.

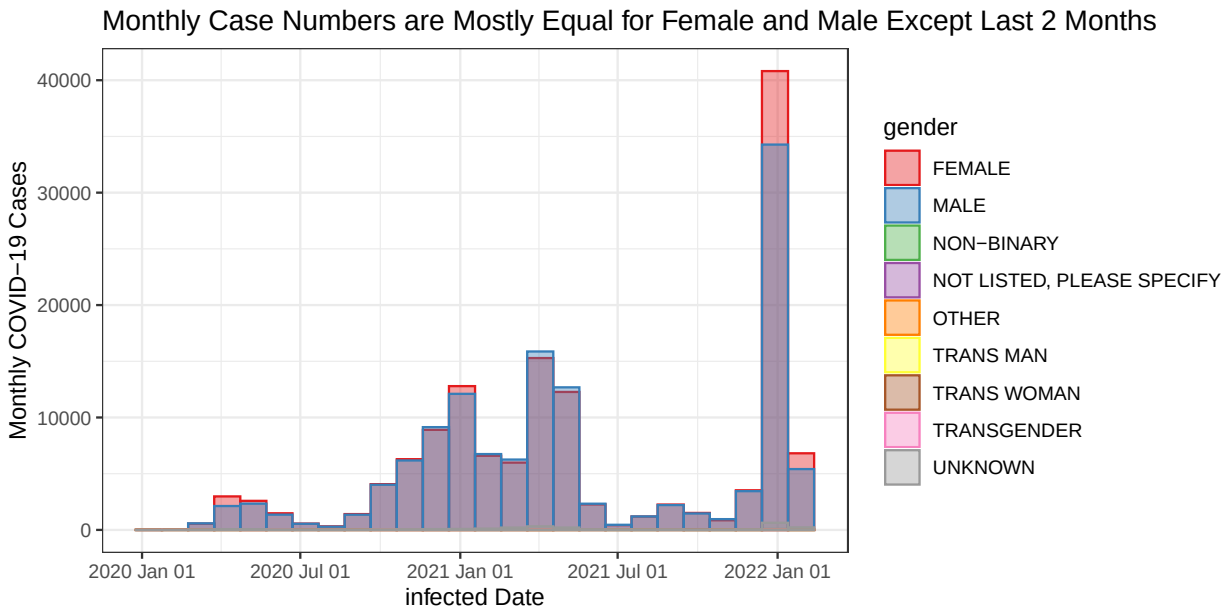

Figure 3: Monthly Number of COVID-19 Cases in Toronto by Gender and over time

Figure 3 shows the monthly numbers of COVID-19 cases in Toronto over time by gender. Since majority of the population in the data are identified as female or male, I will be mainly examining at the infection numbers between these two gender. Red bins displayed in the plot represents monthly case numbers for female, and blue bins for male. Up until December 2021, the case numbers between female and male fairly equal as there seems to be no preference when the virus select its target. However, there seems to an substantial increase in female case numbers when comparing to male case numbers during the last two month of the data. This may be explainable as research have reported that the Omicron variant is also infecting more female and male around the globe (Directorate General of Health Services 2022), although the reason to such behavior remains undiscovered.

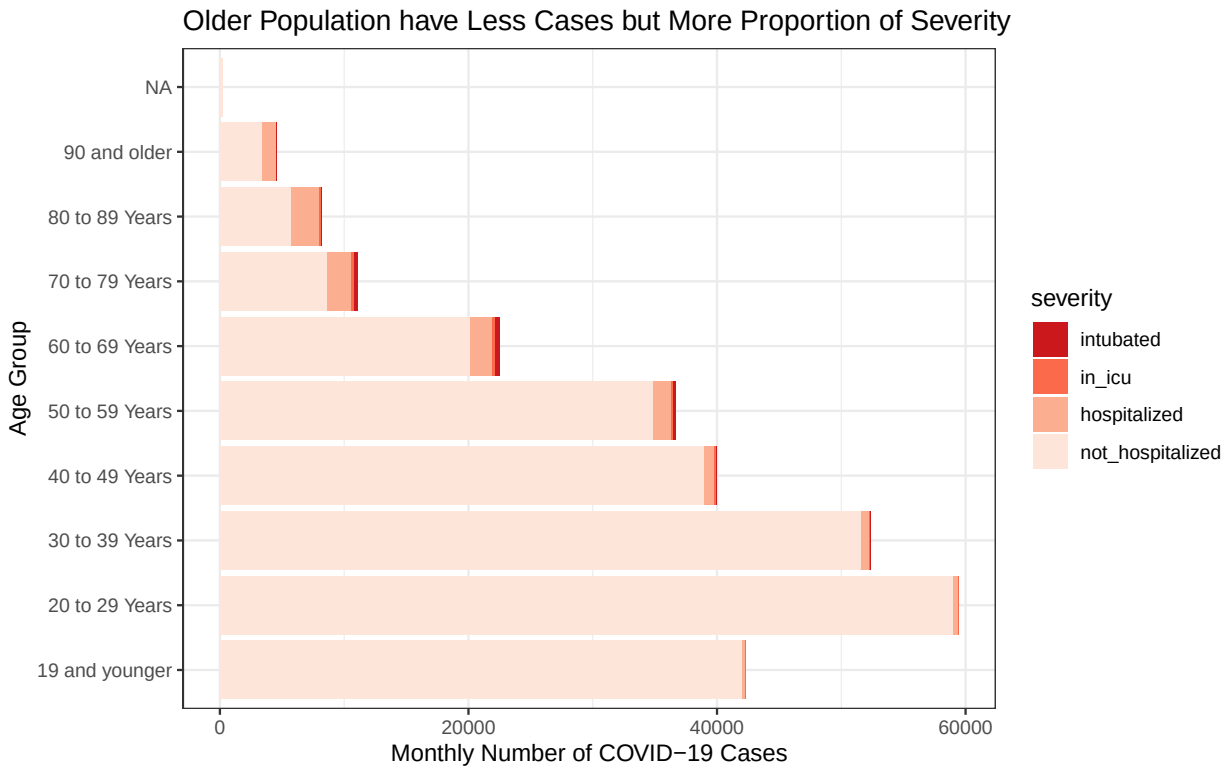

Figure 4: Monthly Numbers of COVID-19 Cases in Toronto by Severity and Age

Figure 4 shows the monthly number of infected cases by age group, with severity once again being the color filling. Younger population(age 59 and lower) have significantly more case number count than older population. Specifically, age 20 to 29 have the most cases of all ages. However, we should not conclude younger population are more likely to get infected with the virus, as other data like Toronto population numbers in each age group should be taken into consideration. Age 20 to 29 years old may have the most case numbers because they have the largest population in Toronto.

Although older population have less case numbers, age 50 years and older have significantly more proportion of severe cases which ended up being hospitalized, or worse, in ICU, or got intubated. Unlike the previous one, we can conclude that among the reported cases, infections with older population are likely to be more severe and be hospitalized for various special treatments.

Table 1: Number Count of COVID-19 Cases in Toronto by Source of Infections

Source	Count
Community	63829
Household Contact	36793
Outbreaks, Healthcare Institutions	15660
Close Contact	15476
Outbreaks, Other Settings	10440
Pending	8187
Outbreaks, Congregate Settings	4117
Travel	3810

Table 1 shows the number count of COVID-19 cases in Toronto sorted by sources of infections. Community

contact seems to be the most common source of infections for COVID-19 and travel seems to be the least common source of infections. The table numbers sounds reasonable as most population interact with their community on a daily basis even in a lockdown situation. For travel, only very small amount of population travel under the pandemic situation and Canadian boarder enforces strict rules to make sure the travelers arriving in Canada are free of the virus(GOverment of Canada 2021), thus making travel the least common source of infections.

References

City of Toronto. 2022a. "COVID-19: Benefits of Being Fully Vaccinated." https://www.toronto.ca/home/covid-19/covid-19-vaccines/covid-19-benefits-of-getting-vaccinated/.

———. 2022b. "COVID-19: Case Counts." https://www.toronto.ca/home/covid-19/covid-19-pandemic-data/covid-19-weekday-status-of-cases-data/.

Directorate General of Health Services. 2022. "Omicron Variant: Infecting More Women Than Men." The Daily Star. https://www.thedailystar.net/health/disease/coronavirus/news/omicron-variant-infecting-more-women-men-2930971.

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

Goethe University Frankfurt. 2022. "Why the Omicron Variant Causes Less Severe Disease: Eight COVID-19 Drugs Remain Active Against Omicron in Cell Culture Study." ScienceDaily. https://www.sciencedaily.com/releases/2022/01/220124203747.htm.

GOverment of Canada. 2021. "COVID-19: Travel, Testing and Borders." https://travel.gc.ca/travel-covid.

Ontario Ministry of Health. 2022. "Case Definition – Coronavirus Disease (COVID-19)." https://www.health.gov.on.ca/en/pro/programs/publichealth/coronavirus/docs/2019_case_definition.pdf.

Public Health Ontario. 2021. "Underreporting of COVID-19 Case Counts in the Context of the Omicron Variant." https://www.publichealthontario.ca/en/about/blog/2021/underreporting-of-covid-19-case-counts-in-the-context-of-the-omicron-variant.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Somos, Christy. 2021. "At Least 1 in 3 COVID-19 Infections Are Asymptomatic, Study Suggests." CTV News. https://www.ctvnews.ca/health/coronavirus/at-least-1-in-3-covid-19-infections-are-asymptomatic-study-suggests-1.5279433.

Toronto Open Data. 2022. *About City of Toronto Open Data.* https://open.toronto.ca/about/.

Toronto Public Health. 2022. "COVID-19 Cases in Toronto." https://open.toronto.ca/dataset/covid-19-cases-in-toronto/.

Tsekouras, Phil. 2021. "Two Travellers Slapped with Hefty Fines for Presenting Fake COVID-19 Test Results at Toronto Airport: Transport Canada." CTV News. https://toronto.ctvnews.ca/two-travellers-slapped-with-hefty-fines-for-presenting-fake-covid-19-test-results-at-toronto-airport-transport-canada-1.5416514.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, and Dana Seidel. 2020. *Scales: Scale Functions for Visualization.* https://CRAN.R-project.org/package=scales.

Xie, Yihui. 2021a. *Bookdown: Authoring Books and Technical Documents with r Markdown.* https://github.com/rstudio/bookdown.

———. 2021b. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.* https://yihui.org/knitr/.

———. 2021c. *Tinytex: Helper Functions to Install and Maintain TeX Live, and Compile LaTeX Documents*. https://github.com/yihui/tinytex.