Senior Torontonians Receive More Treatment for COVID-19 yet Suffer More Fatal Outcomes*

An analysis of age-related treatment and outcomes of COVID-19 cases in Toronto

Adrian Wong

31 January 2021

Abstract

Age is often associated with health and well-being, especially in terms of transmission and infection. Within this paper, analysis of COVID-19 cases in Toronto was compared to various age groups to gauge treatment and outcomes. It was discovered that senior patients received more treatment in terms of hospitalization, admission to intensive care units, and intubation. However, senior Torontonians were among the highest age group to also see fatal outcomes. For example, 23% of cases with patients ages 90 and over received hospitalization, but 30% of these cases ended in fatality. By analyzing this data, results may help guide a more nuanced understanding of how COVID-19 treatment and outcomes differ by a factor of patients' age.

1 Introduction

In 2020, COVID-19 contributed to growing concerns about the health and safety of people nationwide but especially for those who had underlying health conditions and those who were among the elderly. This is because previous research has demonstrated that the elderly were more susceptible to transmission and infection (Davies et al. 2020). Moreover, the World Health Organization has listed COVID-19 to be especially threatening for seniors due to age-related complications and changes in terms of physiology ("WHO Delivers Advice and Support for Older People During Covid-19," n.d.). As a result, data collected by Toronto Public Health and sourced from `opendatatoronto` was analyzed to develop a better understanding of the amount of treatment that elderly Torontonians get, as well as whether this treatment lead to less severe outcomes.

This report is organized in two sections: the data overview and the data analysis. Within the data overview, the source of the data will be discussed, as well as the variables within the dataset and the methodology of analysis. Subsequently, the data analysis will explore the population within the dataset, the trends in cases over time along age groups, the sources of infections, how cases were treated, and the case outcomes.

2 Data Overview

2.1 Source

This dataset was collected by Toronto Public Health in order to better understand COVID-19 in the context of the community, as well as to help inform reactive measures in preventing fatal outcomes in new cases and outbreaks. Toronto Public Health (TPH) was established in 1883 and reports to the Board of Health, overlooking the health and well-being of 2.9 million Torontonians. Their mandate focuses on the preventing the spread of disease, using surveillance to monitor the health status of the population, and developing/implementing public policy and practices (C. of Toronto 2019b).

*Code and data are available at: https://github.com/A1001949/covid19_toronto.git

Rows from this dataset were collected from the provincial communicable disease reporting system (iPHIS), which is an information system that stores case information in regards to reportable communicable diseases in Ontario. Each public health unit is required to report individual cases within the system ("Extracts from the Integrated Public Health Information System (iPHIS) - Ontario Data Catalogue" 2018). This data is also combined with Toronto Public Health Coronavirus Rapid Entry System (CORES) after the iPHIS was deemed to be "outdated" and "not meeting the city's needs" ("Toronto Creates New It System to Track Covid-19" 2020). As is the case, a limitation in this dataset is that a non-response bias was not accounted for - outcomes represented from this package only indicate the case data for respondents who visited a public health unit. As a result, descriptive analysis of this data does not extend to **all** cases of COVID-19 in Toronto, only those cases that were reported to a public health unit since the first reported case in January 2020.

As cases continue to be reported, the dataset is completely refreshed and overwritten by Wednesday of every week. As is the case, the results found in this paper is up to date as of January 31st, 2021 but the results of this analysis is subject to change as new cases are reported.

2.2 Variables

Within this dataset, there were **75268** responses and 18 of the following features:

- `_id` : Unique row identifier for Open Data database
- `Assigned_ID` : A unique ID assigned to cases by Toronto Public Health
- `Outbreak Associated` : Sporadic or Outbreak Associated case
- `Age Group` : Age of case patient (9 groups in decades + unknown group)
- `Neighbourhood Name` : 1 of the 140 geographic neighbourhoods in Toronto, based on the patient's FSA
- `FSA` : Forward sortation area, based on the patient's home address
- `Source of Infection` : Most likely source of infection (8 options)
- `Classification` : Categorization of the case as CONFIRMED or PROBABLE
- `Episode Date` : Best estimate when the disease was acquired
- `Reported Date` : The date on which the case was reported to TPH
- `Client Gender` : The patient's self-reported gender
- `Outcome` : Patient case was Fatal, Resolved, or is currently Active
- `Currently Hospitalized` : Cases that are currently admitted to the hospital
- `Currently in ICU` : Cases that are currently admitted to the intensive care unit
- `Currently Intubated` : Cases that are currently intubated related to their COVID-19 infection
- `Ever Hospitalized` : Cases that are ever admitted to the hospital (includes currently hospitalized)
- `Ever in ICU` : Cases that are ever admitted to the intensive care unit (includes currently ICU)
- `Ever Intubated` : Cases that were ever intubated related to COVID-19 (includes currently intubated)

On top of these variables, TPH has cited evidence to suggest that social and economic factors can play a large role in calculating the risk of COVID-19 through decreased access to healthcare (Ontario 2020). As such, TPH has also collected data in regards to racial group, income, household size, and Indigenous identity since May 20, 2020. Though important, this data is not yet available for analysis, nor is it in the scope of this paper.

2.3 Methodology

To analyze this dataset, `R` statistical programming language (R Core Team 2020) was used to import the respective dataset (Toronto 2021) from the Open Data Toronto Portal using the `opendatatoronto` package (Gelfand 2020). For data manipulation, exploration, and visualization, the `tidyverse` package was used as well (Wickham et al. 2019). In particular, graphs were created using `ggplot2` (Wickham 2016) and a table was constructed using `kable` (from the `knitr` package (Xie 2020)) and `kableExtra` (Zhu 2020). The `scales` package was used to convert raw numbers to percentages (Wickham 2018). Finally, `janitor` was used to clean the data (Firke 2021).

Previous research has indicated that there are age-related differences in the outcomes of COVID-19, potentially due to the age-related differences in terms of transmission and infection. In particular, previous research has demonstrated that susceptibility to infection for individuals ages 20 years and younger are approximately half of that for individuals ages 20 years and older. Moreover, symptoms manifest in 21% of infections for those between 10 and 19 years old, whereas 69% of infections for those over 70 manifest symptoms (Davies et al. 2020). As is the case, analysis in this report relied heavily on the `Age Group` feature within the dataset.

Once the data was imported, an `R` script were used to clean the data. This was done by first formatting the column names such that capitalization and spaces were removed in favour of underscores. Next, rows where the `Classification` of the case was *PROBABLE* as opposed to *CONFIRMED* were dropped. This was done in order to verify that the analysis being preformed was on cases that were with regard to COVID-19 and not a similar infection. Finally, as this analysis focuses on age-related differences in COVID-19, rows where the age-group was unknown were dropped.

Identifying features were unrelated to the analysis and were not used (`_id`, `Assigned ID`). The `Outbreak Associated` feature was not utilized as well, since the `Source of Infection` column already indicates whether the case was caused by an outbreak. The `FSA` column was not used either, since the `Neighbourhood Name` data was a good representation for the geological spread of these cases. Instead of analyzing `Reported Date`, it was more fitting to base analysis on `Episode Date` as it was a estimation of when the patient first contracted the disease (there may be variability for when the patient decided to visit a public health unit). Finally, `Currently Hospitalized`, `Currently in ICU`, and `Currently Intubated` were dropped in favour of `Ever Hospitalized`, `Ever in ICU`, and `Ever Intubated` since the latter is inclusive of the former.

3 Data Analysis

3.1 Population

A stacked bar graph was used to illustrate the distribution of gender per age group within this population (Figure 1). This figure also illustrates that within this dataset, the largest age group was those who were 20 to 29 years old and the smallest age group was those who were 90 and older. However, it is important to note that we can not conclude the likelihood of catching COVID as associated to one's age group as the distribution of age in Toronto is beyond the scope of this dataset. In other words, there may be an over-representation of patients aged 20 to 29 years old within this dataset simply because there are more 20 to 29 year-olds residing in Toronto than there are residents who are 90 years and older.

Additionally, it also appears to be an equal amount of females to males for respondents between 19 and 79 years old. However within the 80 to 89 and 90 and older age groups, there appears to be more female respondents than male respondents. For example, 72% of respondents 90 and older are females while 26% are males. Within all age groups, there appears to be a negligible amount of transgender respondents (n=10). Furthermore, these respondents were distributed across 140 neighbourhoods in Toronto (a jitter plot was used to illustrate this neighbourhood data; see Figure 5 in the Appendix for more details).

3.2 Cases over Time

Spikes in COVID19 cases seem to vary by age groups across time. As illustrated in the line graph in figure 2, there was a recent spike in COVID19 cases for respondents under the age of 59 years old in January 2021. Interestingly, the spike was much more moderate for respondents between the ages of 60 and 69, and almost non-existent for participants ages 70 and older. Instead, it appears that there were several spikes in cases between October and December of 2020 for respondents aged 90 and up.

3.3 Source of Infection

If we look at the source of infection for each of these age groups, we also see that respondents who were 80 years and older tended to be exposed to COVID19 through non-sporadic, outbreak-associated means (see figure 3). On top of this, a higher percentage of respondents within this age range were exposed to COVID19 through institutional means compared to other age ranges. This may point to the many recorded outbreaks seen in retirement homes and other healthcare institutions (C. of Toronto 2019a). This might help explain

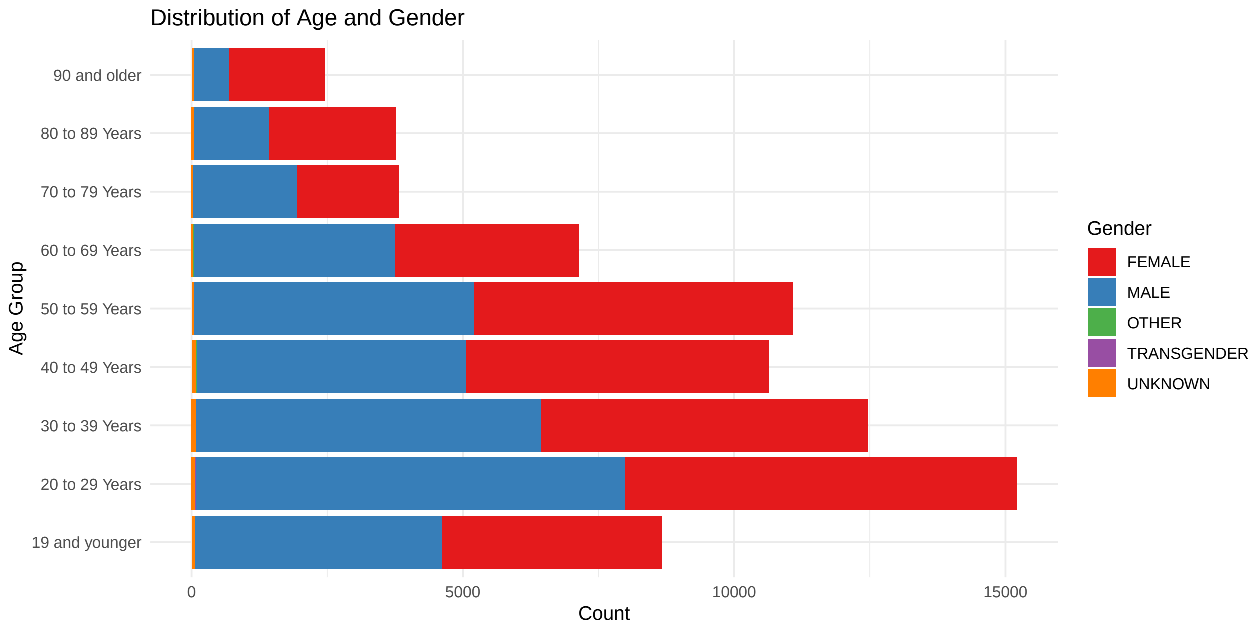

Figure 1: Most respondents within this sample are between the ages of 20 to 29 years old

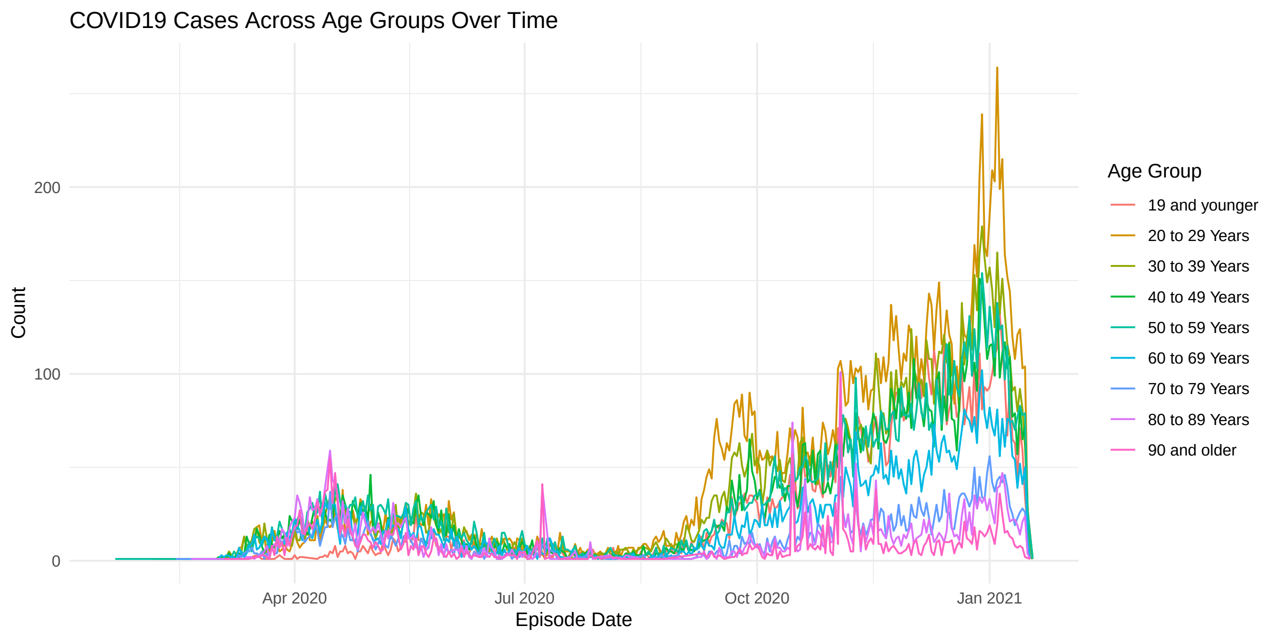

Figure 2: COVID19 cases spiked for Torontonians under the age of 59 in January 2021

the volatility in COVID19 cases among those 90 and older in figure 2 as cases resulting from outbreaks might be reported as comparable episode dates. In contrast, a large percentage of respondents who were younger than 70 years old tended to be exposed to COVID19 through close contact and through the community. In particular, 79% of respondents 90 and up were exposed to COVID19 through outbreak-associated means and 49% of respondents 19 and under were exposed to COVID19 through close contact instead.

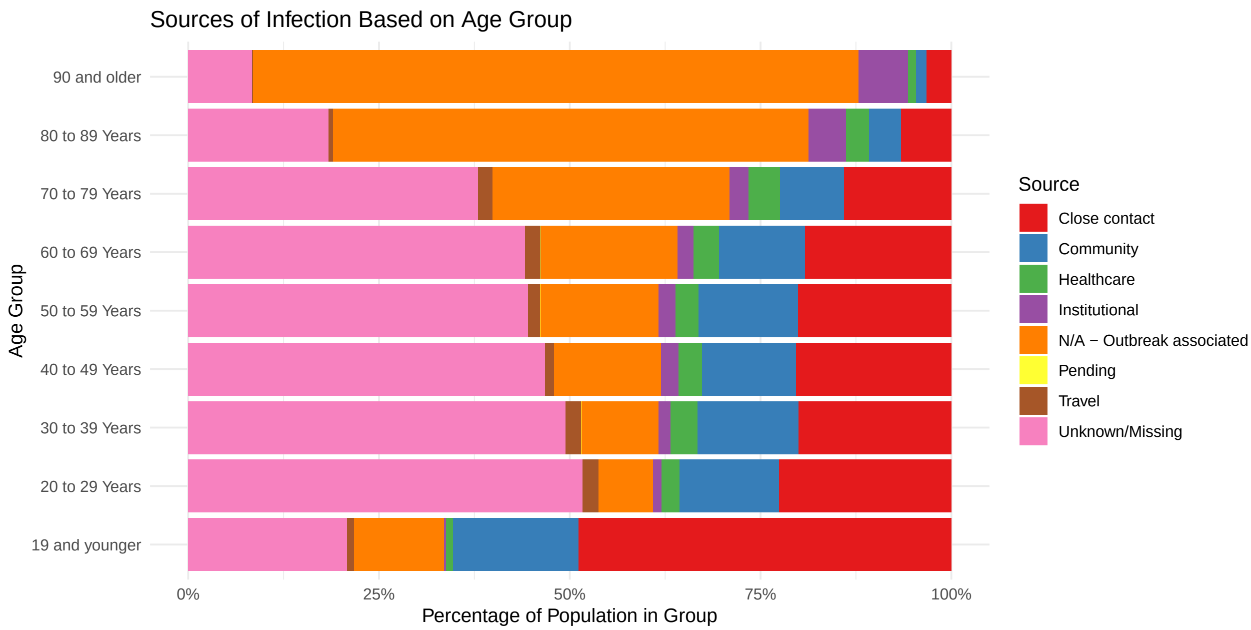

Figure 3: Most COVID19 cases among those 90 and older came from outbreaks

3.4 Treatment and Case Response

As indicated by table 1, seniors appear to receive more treatment for COVID19 than their younger counterparts. The age group with the highest percentage of hospitalization is those who are ages 80 to 89 years old (approximately 30%) and the age group with the highest percentage of intubation is those who are ages 70 to 79 years old (approximately 3.6%).

Table 1: Senior Torontonians receive more treatment for COVID19

Age Group	Hospitalized (%)	ICU (%)	Intubated (%)	Count per Group
19 and younger	0.69	0.18	0.035	8667
20 to 29 Years	0.82	0.13	0.072	15210
30 to 39 Years	1.60	0.18	0.088	12470
40 to 49 Years	3.10	0.72	0.450	10645
50 to 59 Years	5.50	1.60	1.200	11084
60 to 69 Years	13.00	3.90	2.400	7143
70 to 79 Years	25.00	5.60	3.600	3818
80 to 89 Years	30.00	3.30	1.900	3769
90 and older	23.00	1.10	0.370	2462

3.5 Case Outcomes

One might assume that with a higher percentage of senior patients receiving treatment compared to younger patients, their outcomes would fare much better. Unfortunately, this does not seem to be the case. As shown in the stacked bar chart in figure 4, participants who were older had more fatal cases of COVID19 than younger respondents. 30% of respondents aged 90 and older had fatal cases of COVID19. This is contrast to the 1 and 1 fatal cases of COVID19 among respondents ages 20 to 29 and 19 & younger respectively.

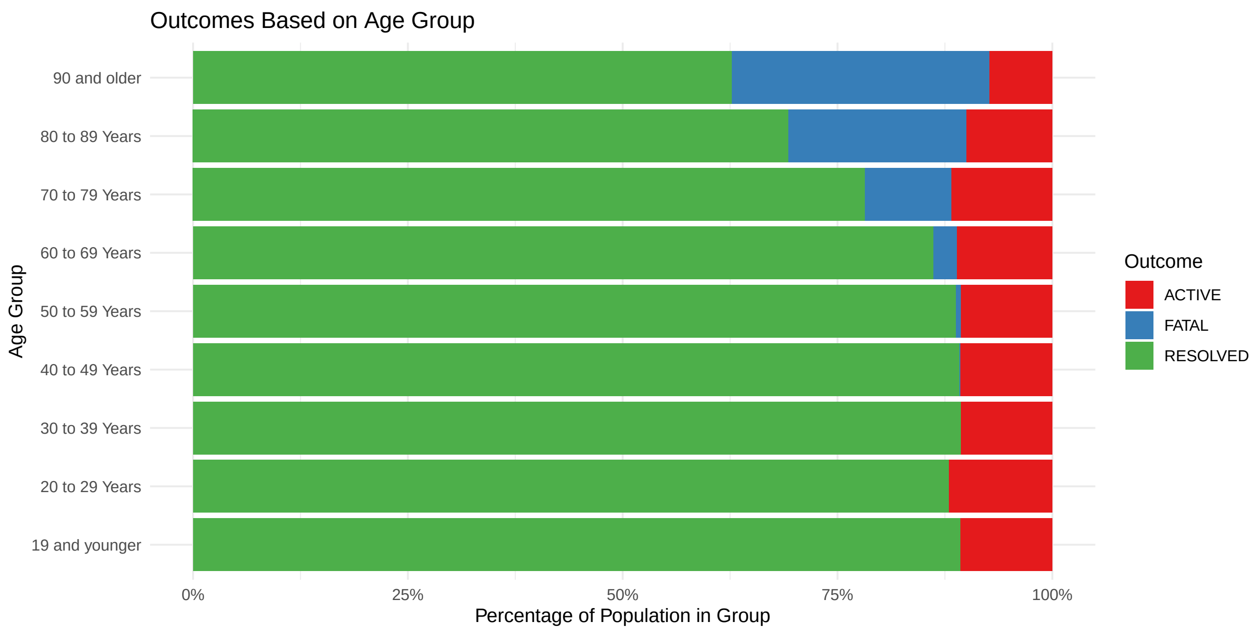

Figure 4: Senior patients have significantly more fatal cases than young patients

4 Conclusion

In conclusion, senior Torontonians receive more treatment than younger Torontonians, but still don't necessarily fare better in terms of outcomes. This mirrors previous research by the World Health Organization that indicated that older individuals are more susceptible to developing higher severity of the illness due to physiological changes as a result of aging and a declining immune system ("WHO Delivers Advice and Support for Older People During Covid-19," n.d.). It is important to note the connection between higher amounts of treatments and case outcomes are not yet known and can not be inferred from this analysis. We can not conclude whether increases in treatment were reactionary to the increasing severity of cases which finally lead to fatal outcomes, or whether treatment acted as a preventative but ineffective measure against the already high fatality rate of cases among the elderly. Moreover, a limitation of this dataset was that it also did not include data in regards to underlying health conditions. If old age correlates with significant underlying health conditions, this may explain the increased fatality rate of COVID-19 cases given other complications and implications of these conditions. Future studies may seek to understand the relationship of old age, treatment, case conditions, and underlying conditions in order to answer these questions and extend this analysis.

Appendix

Respondents were distributed across 140 of Toronto's neighbourhoods (Figure 5).

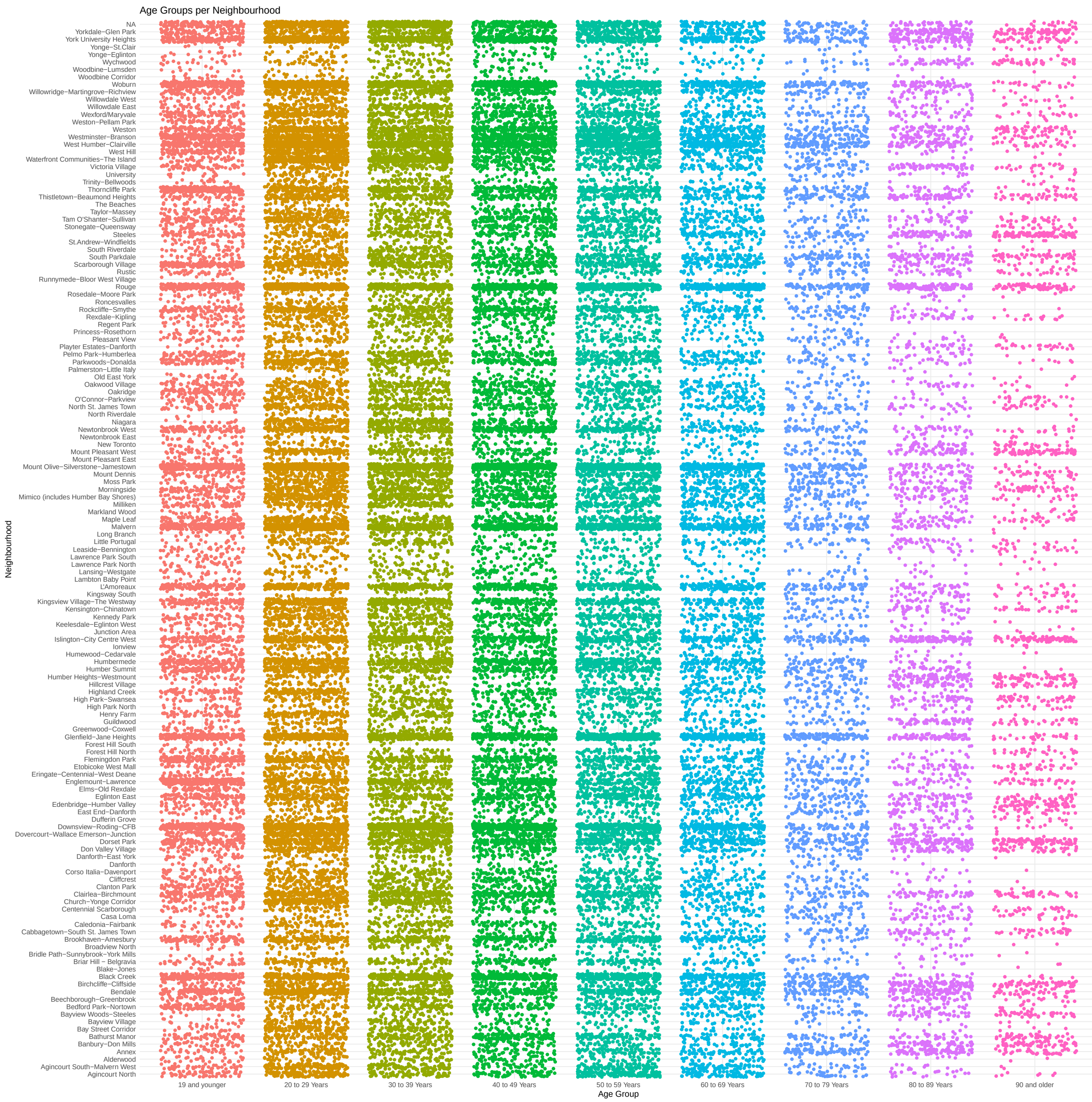

Figure 5: Age groups were distributed relatively evenly across Toronto neighbourhoods

References

Davies, Nicholas G., Petra Klepac, Yang Liu, Kiesha Prem, Mark Jit, and Rosalind M Eggo. 2020. "Age-Dependent Effects in the Transmission and Control of Covid-19 Epidemics." https://doi.org/10.1101/2020.03.24.20043018.

"Extracts from the Integrated Public Health Information System (iPHIS) - Ontario Data Catalogue." 2018. *Extracts from the Integrated Public Health Information System (iPHIS) - Datasets - Ontario Data Catalogue*. https://data.ontario.ca/dataset/extracts-from-the-integrated-public-health-information-system-iphis.

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

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Ontario, Public Health. 2020. "COVID-19 – What We Know so Far About Social Determinants of Health." https://www.publichealthontario.ca/-/media/documents/ncov/covid-wwksf/2020/05/what-we-know-social-determinants-health.pdf?la=en.

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

Toronto, Open Data. 2021. "COVID-19 Cases in Toronto." City of Toronto Open Data Portal. https://open.toronto.ca/dataset/covid-19-cases-in-toronto/.

"Toronto Creates New It System to Track Covid-19." 2020. *Canadian Healthcare Technology*. https://www.canhealth.com/2020/04/22/toronto-creates-new-it-system-to-track-covid-19/.

Toronto, City of. 2019a. "Active Outbreaks in Toronto Healthcare Institutions." *City of Toronto*. https://www.toronto.ca/community-people/health-wellness-care/health-inspections-monitoring/active-outbreaks-in-toronto-healthcare-institutions/.

———. 2019b. "Toronto Public Health." *City of Toronto*. https://www.toronto.ca/city-government/accountability-operations-customer-service/city-administration/staff-directory-divisions-and-customer-service/toronto-public-health/.

"WHO Delivers Advice and Support for Older People During Covid-19." n.d. *World Health Organization*. World Health Organization. https://www.who.int/news-room/feature-stories/detail/who-delivers-advice-and-support-for-older-people-during-covid-19#:~:text=The COVID-19 pandemic,potential underlying health conditions.

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

———. 2018. *Scales: Scale Functions for Visualization*. https://CRAN.R-project.org/package=scales.

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.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.

———. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. https://yihui.org/knitr/.

———. 2020. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. https://yihui.org/knitr/.

Zhu, Hao. 2020. *KableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.