Preserving the Life of the Homeless*

Mohid Sharif

January 23, 2024

How valuable is the life of a homeless person, and what efforts do we need to take to preseve their lives? To answer this question I obtained and analyzed data about causes of homeless deaths by age group and year. I found that drug abuse was the leading cause of death among homeless people followed by sickness and disease. By analyzing my findings I hope to identify the leading causes of death among the homeless so we can work towards changes that will help preserve life.

1 Introduction

Homelessness alone is a huge problem that seems to persist. Realizing that homelessness is a problem that may not be solved in the near future, we need to find a way to solve the issue of homeless deaths. How valuable is the life of a homeless person? What efforts should we make to preserve their lives? What takes more effort, reducing homeless mortality or ending homelessness? To answer these questions I set out to obtain information and see if I could understand why homeless people are dying. By understanding why they are dying I hope to find the best course to action to preserve the lives of the homeless.

I first analysed the data-set to understand how reports of deaths were being recorded and organized. Every report in the data-set records the cause of death, year of death, age group, gender, and count. These are just a handful of variables from a single data-set in a larger package but my research only required the ones I have listed. I then created three graphs to compare the amount of reports per age group, per year and the number of deaths per report, all of these color coded by cause of death. The first graph is a comparison of reports by age group, I found that younger ages (<20 years) were most likely to die due to drug toxicity, while the reports of age groups over 40 had the highest reports of disease/sickness followed by drug toxicity. The second graph shows the deaths per report, I used this graph to illustrate that despite seeming equal in terms of reports both sickness/disease and drug toxicity had alarmingly different death numbers by report. Drug toxicity showed a much higher number of

---

*Code and data are available at: https://github.com/MohidSharif/Causes-of-Homeless-deaths.

deaths per report than any other cause of death. The last graph illustrated the change over the years of death reports, in this graph we can see that despite fluctuations we see no real pattern of increase or decrease in report numbers, however, there was a significant drop in reports in 2023 compared to previous years.

Using my data I believe that we can deduce the main cause of deaths for homeless individuals and start targeting that cause to try and eliminate the threat towards homeless people. By far the biggest threat to homeless life is drug toxicity, knowing this we need to find a way to get drugs off the streets and get homeless people help with staying off of drugs and finding productive alternatives. The second highest cause is sickness/disease which could be due to issues with accessibility or affordability of healthcare. To reduce deaths by sickness/disease we can look towards increasing accessibility and affordability of healthcare for homeless people. Lastly, looking at the previous years we can look at any changes that may have been made in these categories and see if they made any differences. After analyzing the data from previous years I could see that if any changes were made before 2023, these changes made no difference in the data. In 2023 however, there is a big drop in reports which shows that if any changes were made in 2023 they may have affected the homeless deaths situation, although more data will be required to be sure if any changes are productive or not.
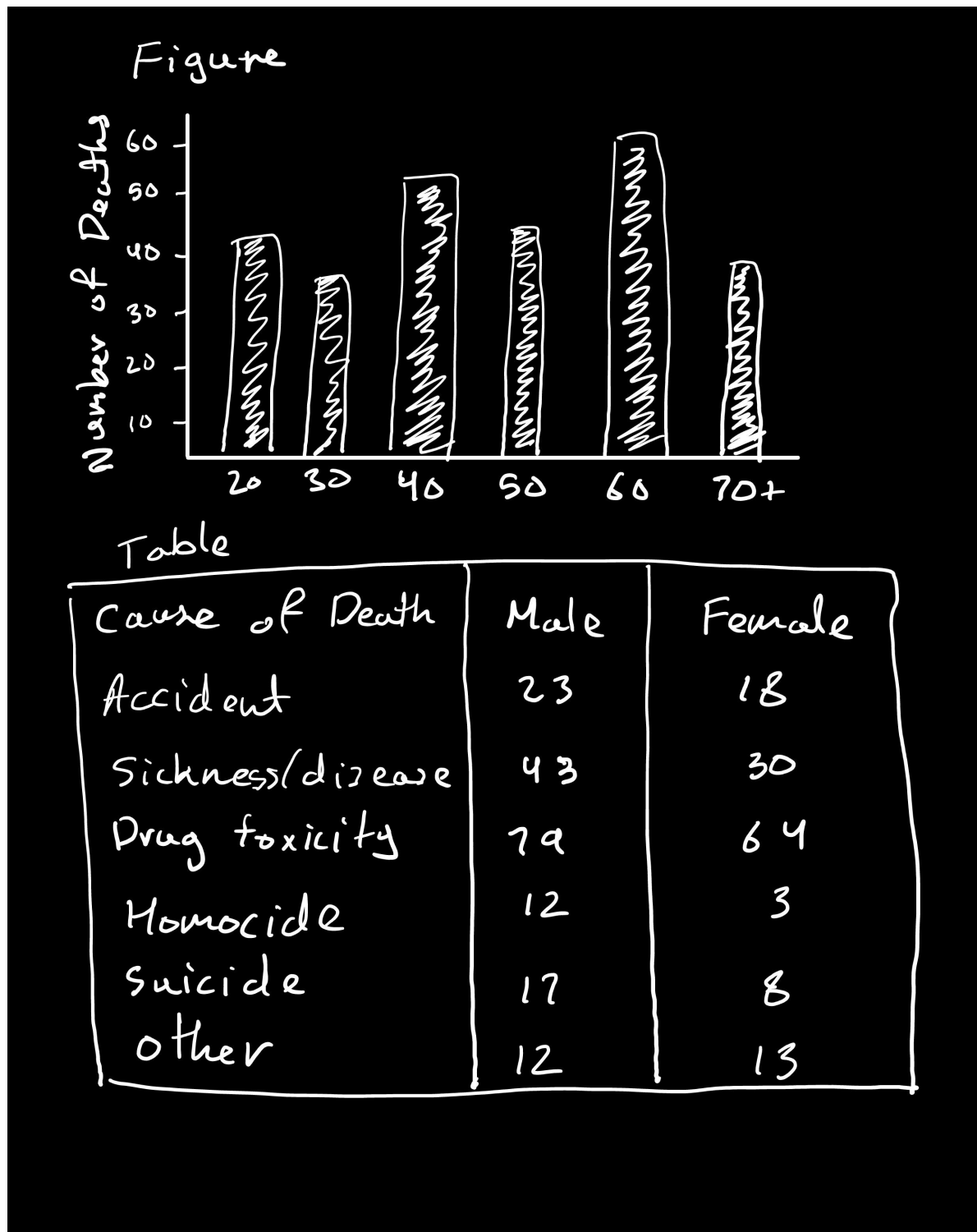
Figure

Number of Deaths

20	30	40	50	60	70+

Table

Cause of Death	Male	Female
Accident	23	18
Sickness/disease	43	30
Drug toxicity	79	64
Homocide	12	3
Suicide	17	8
Other	12	13

Figure 1: Sketch of what our graphs and tables may look like.

The remainder of this paper is: Section 2....

2 Data

I obtained my data from the City of Toronto **opendatatoronto** database portal, using the 'opendatatoronto' package (Gelfand 2020) and the statistical programming language **R** (R Core Team 2020). I used the **tidyverse** package for data manipulation (Wickham et al. 2019) and **kableExtra** for table formatting (Zhu 2021). The header includes two lines of code **"usepackage{float}"** which allows the use of float in R markdown and the line **"float-placement{figure}{H}"** (user9112767 2018) which keeps the tables and figures locked in the specific place where they are written in R markdown.

This data set records reports of homeless deaths across Toronto and records the details of the deceased. The data records the **year of death, cause of death, age group, gender, and number of deaths** for every report. The data classifies cases by age group rather than specific ages, age groups are grouped by 20 year age gaps (E.g. "20 to 39 Years") except for the first group which is **"<20"** and the last group which is **"60+"**. Gender is identified and reported as **"Male"** or **"Female"**. Year of death is simply as the year the deaths were reported, starting from **2017** up until **2023**. Cause of deaths are classified as follows: **"Accident", "Cancer", "Cardiovascular Disease", "Covid-19", "Drug Toxicity", "Homicide", "Pneumonia", "Other", "Suicide", and "Unknown/Pending"**. And the number of deaths is simply the number of deaths provided in that report.

Table 1 shows the data associated with cases reported as "death due to drug toxicity".This data shows us year, age group, gender and count of each reported case of drug toxicity.

Table 1: Deaths Due to Drug Toxicity

Year	Age Group	Gender	Count
2017	40-59	Female	2
2017	20-39	Female	3
2017	Unknown	Male	1
2017	60+	Male	3
2017	20-39	Male	10
2017	40-59	Male	13
2018	60+	Female	1
2018	40-59	Female	2
2018	20-39	Female	8
2018	60+	Male	1

Table 1 line 1 can be read as such; 2 females in the age range 40-49 died in 2017 due to drug toxicity. The table is a good representation of which data is available to us and how we can use it to obtain which cause of deaths are most common by counting all reports that fall under that cause. I specifically filtered this table to only include drug toxicity cases to reduce result size and to showcase the data for a single cause of death. I also filtered all age groups identified as "Unknown" as this information provides no benefit to us.

From here I further cleaned the data for graphing. I cleaned the data by grouping **"Cancer", "Covid-19", "Cardiovascular Disease", and "Pneumonia"** into a single group **"Sickness/Disease"**. I grouped these classifications because they all fall under one solvable service, healthcare. I also filtered out cause of deaths classified as "Unknown/Pending" and "other" as these provide no benefit to our data.

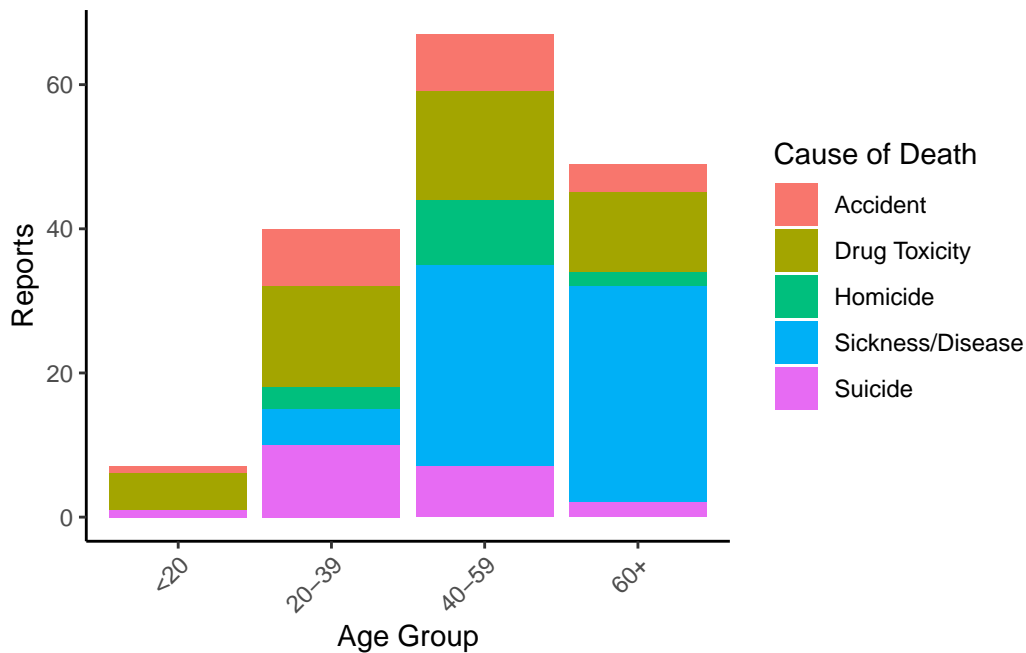Figure 2 shows the comparison of number of reports recorded for each age group.

Figure 2: Death Reports by Age group

Figure 2 shows a visual of the number of reports per age group for each cause of death. The biggest observation here is the difference in reports within age groups. Figure 2 shows us that youth despite having the least number of reports have the highest proportion of drug toxicity reports, demonstrating a big drug problem within the age group of <20. Sickness/disease seems to become a problem after 40 years of age as more health problems occur with age. Figure 2 shows a huge spike in sickness/disease reports for ages from 40+. Other than this we can see that drug toxicity is the number one cause of death for all ages. Understand that Figure 2 only shows reported cases, each reported case contains multiple deaths and as we will see in Figure 3, drug toxicity has the most deaths per-report. Even though Figure 2 shows sickness/disease as the leading cause of death in some age groups that may not necessarily be the case as drug toxicity has the higher proportion of deaths reported per case.

Our biggest concern from seeing this comparison should be to reduce the access to drugs for all ages and find alternative productive solutions to drug use. Secondly, we need to provide better and accessible healthcare for the homeless over 40 years of age. Suicide seems to be a big portion of reports for people in the age range of 20-39, maybe we can look towards mental health support for the homeless and see if we can provide some quality of life changes to homeless shelters and food banks.

Figure 3 shows the comparison between deaths reported for each cause of death.
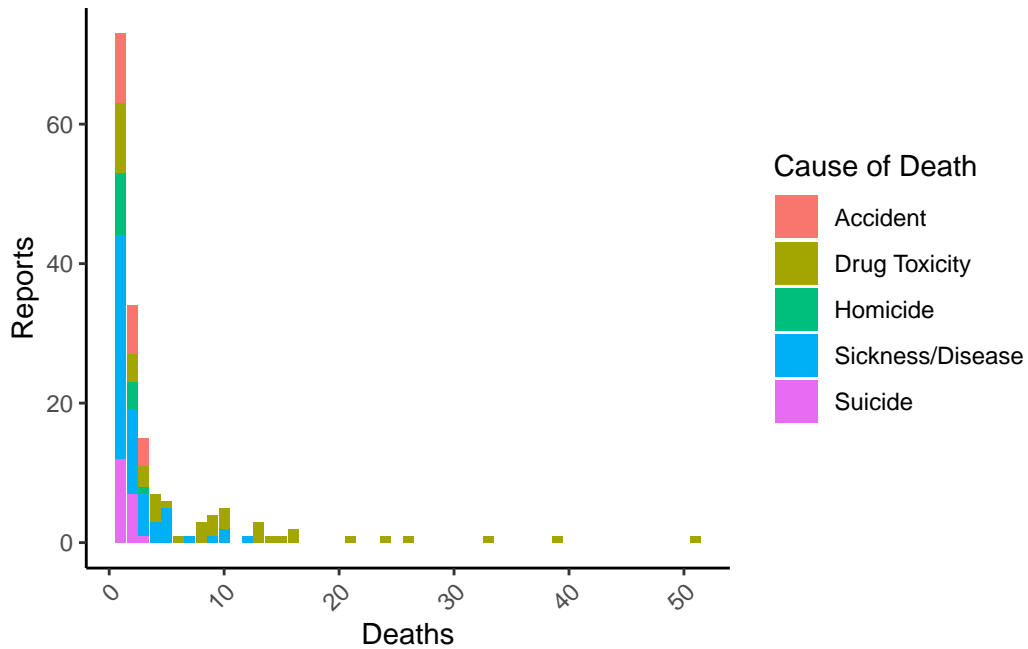

Figure 3: Number of Deaths per Report

The reason I wanted to illustrate this comparison was to show that quantity in reports is not proportional to death toll. As Figure 3 illustrates, most reports are within the 1-2 range, and within this range lies the majority of sickness/disease reports. With this we can see that despite having the most reports in Figure 2 the number of deaths per report is low. Now looking at drug toxicity we can see that the number of deaths are spread out from 0-16 deaths per report while there are some reports that have death tolls as high as 40 or even 50 deaths. Showing that drug toxicity is the leading cause of death and may be the problem that needs to be prioritized.

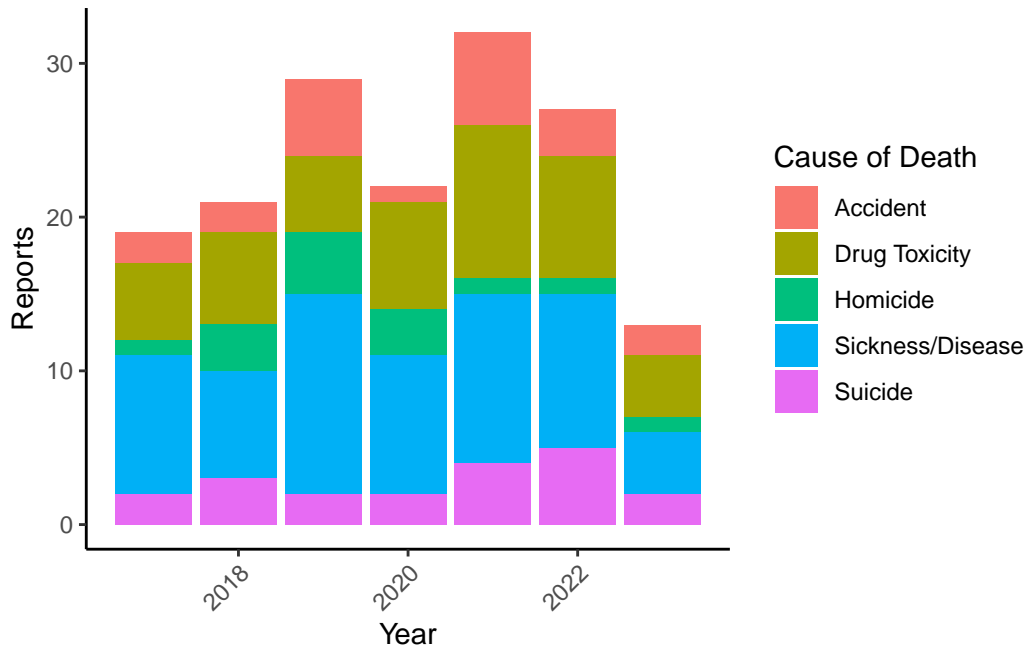Figure 4 shows the change in number of reports over the years, from 2017 up until 2023.

Figure 4: Number of Reports Since 2017

After looking at Figure 4 I was honestly disappointed to see that practically no progress has been made in regards to this issue over the last 7 years. Even though there are changes and fluctuations, there is no decreasing or increasing trend in neither the overall reports number nor reports of individual causes. The only ray of hope in Figure 4, is the number of reports in 2023, which shows an almost 50% decrease from previous years. Going forward we should look at what changes were made between 2022 and 2023 that caused the situation to improve so drastically. Doing this we can apply this to other regions in Canada and preserve the life of the homeless throughout the country.

References

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal.*

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

user9112767. 2018. "How to Hold Figure Position with Figure Caption in PDF Output of Knitr?" *Stack Overflow.* https://stackoverflow.com/questions/29696172/how-to-hold-figure-position-with-figure-caption-in-pdf-output-of-knitr.

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.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax.*