Homelessness and Drug Toxicity*

Gavin Crook Samarth Rajani Mohid Sharif

March 16, 2024

In this study, we investigated fatalities due to drug toxicity among the homeless population over the years, examining trends across age and gender. Fatalities increased steadily from 2017 onwards across all ages and genders until 2021, where we observed a sharp spike in deaths of males aged between 20 and 59 years of age, which returned back to expected pre pandemic levels in 2023. The study highlights the neglected state of the homeless in Toronto, which require targeted interventions to bolster their welfare especially during a public health crisis.

1 Introduction

You can and should cross-reference sections and sub-sections. We use R Core Team (2023) and Wickham et al. (2019a).

For many years, Toronto has struggled with homelessness. Amidst an affordable housing shortage and escalating living expenses, escaping the cycle of homelessness has grown increasingly challenging. Being homeless brings along with it a string of issues - the most obvious ones being lack of food, being caught out in adverse weather and exposure to drugs and other substances. Often, survivability for part most depends on consumption of numbing drugs, especially in the harsh winter. As a result, overdosing becomes a common risk factor among homeless populations, accentuating the urgency for comprehensive research into drug toxicity dynamics within this demographic. This paper aims to delve into the intricate relationship between drug toxicity and mortality rates among homeless individuals in Toronto, with a focus on how these dynamics vary across different age groups and genders over time.

It is important to study the effects of drug toxicity among the homeless for several reasons. First, it gives insight into the challenges of homelessness beyond just shelter and sustenance. Moreover, it highlights the shortcomings of local institutions in both aiding the homeless and restricting availability of non-legal drugs. Finally, it provides insights for policy makers to devise targeted intervention to help solve problems on both accounts.

*Code and data are available at: https://github.com/MohidSharif/Drug-Toxicity-and-Homelessness..

By examining data over several years, this paper will help identify trends across different demographics in Toronto. It will also illustrate any changes in drug lethality due to the pandemic, which may or may not affect proclivity for drug consumption. This will help draw inference about any disproportionate affects on any demographic, and thereby aid in tackling the problem at hand.

The remainder of this paper is structured as follows. Section 2….

2 Data

We obtained our data from the City of Toronto **opendatatoronto** database portal, using the 'opendatatoronto' package (Gelfand 2020) and the statistical programming language **R** (R Core Team 2023). We used the **tidyverse** package for data manipulation (Wickham et al. 2019b) and **kableExtra** for table formatting (Zhu 2021). The header includes two lines of code **"usepackage{float}"** which allow the use of float in R markdown and the line **"floatplacement{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.

Since we are only interested in deaths due to drug toxicity, we simplified the data and created two other data sets. First the data was cleaned to remove any **"Unknown"** or **empty** values. Then causes of death due to **Drug Toxicity** were isolated and all others were removed from the data set. We then created two data sets, one highlighting the number of deaths per year for each gender and one for the number of deaths per year for each age group. Since each report in the data set can contain multiple deaths, therefore all counts from each report had to be added to their respective grouping to create a new death count variable for the two data sets. Using these new data sets we can now compare and analyze the death trend over the years for each gender and age group.

Table 1 shows the data associated with cases reported as "death due to drug toxicity". This data sample shows the number of deaths reported for the years 2017 and 2018 for each age group.

Table 1: Age Group Deaths Due to Drug Toxicity

Year	Age Group	Deaths
2017	<20	0
2017	20-39	13
2017	40-59	15
2017	60+	3
2018	<20	0
2018	20-39	16
2018	40-59	15
2018	60+	2

Line one of Table 1 can be read as follows: In the year 2017 there were no deaths due to intoxication for homeless individuals aged less than 20 years. Similarly we can read the rest of the data set. Just looking at this sample table we can see that ages 20-39 and 40-59 seem to have drastically more deaths due to intoxication compared to the other age groups. We can use this when visualizing our data to determine if our data and graphs match up as expected.

This data was of interest to us because we wanted to study which age groups were most prone to drug intoxication. By isolating age groups we can determine which age group is in need of the most support against the drug problem. We also wanted to observe how the death toll for each age group changed over the course of our 5 year data set. By doing this we can determine why a rise or dip in deaths occurred for that age group in that year.

Table 2 also shows the data associated with cases reported as "death due to drug toxicity". This data however, shows the number of deaths reported for the years 2017 and 2018 for each gender rather than age group.

Table 2: Gender Deaths Due to Drug Toxicity

Year	Gender	Deaths
2017	Male	26
2017	Female	5
2018	Male	22
2018	Female	11
2019	Male	26
2019	Female	13
2020	Male	57
2020	Female	17

Line one of Table 2 can be read as follows: In the year 2017 there were 26 deaths caused by intoxication for Male homeless individuals. Similarly we can read the rest of the data set. From this table we can observe that the Male homeless population seems to consistently suffer more deaths due to drug intoxication every year. Again we can use this trend to determine if our figures match up as expected with our tables.

We choose gender as our second variable of interest to compare the effect of drugs on homeless males and females. Using this approach we can determine which gender population was more in need of support and pair this with our age group data set to further narrow down which demographic was suffering the most. Looking at just Table 1 and Table 2 we can already speculate that males between the ages of 30-49 were most prone to drug intoxicated deaths.

Using our two data sets we hope to find ways to improve the living situation of the homeless population. We hope that our findings can help determine what changes need to be made and which homeless demographic should be targeted in efforts to reduce homeless mortality.

3 Model

The goal of our modelling strategy is twofold. Firstly, we use a poisson distribution and a negative binomial distribution as we have count data.Both distribution can be use on count data, and by using both we are able to compare results. Secondly, we use a multilevel or Bayesian model. We use this multilevel model to account for the nested nature of the data. All counts of deaths are nested within a year and then age group and gender. This always the model to control variance within each nested data point.

I DO NOTKNOW WHAT TO PUT HERE Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define $count_i$ as the count of deaths that resulted from drug toxicity in Toronto. Then $Gender_i$ is the gender for that count of deaths and $Agegroup_i$ is the range of ages for the count of deaths deaths.

$$
\begin{aligned}
\text{Count}_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (1) \\
\mu_i &= \alpha + \text{Gender}_i + \text{Age\_group}_i && (2) \\
\alpha &\sim \text{Normal}(0, 2.5) && (3) \\
\text{Gender} &\sim \text{Normal}(0, 2.5) && (4) \\
\text{Age\_group} &\sim \text{Normal}(0, 2.5) && (5) \\
\sigma &\sim \text{Exponential}(1) && (6)
\end{aligned}
$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a their to be a larger relationship between males and the number of deaths. This is because there are more males within the homeless population and more males homeless as a result of addiction to drugs. These reason make it clear that that the male relationship is larger than the female.

Secondly, we expect the relationship to be larger between younger individuals and the number of deaths. We expect this result as the older individuals who die when homeless, will die of illness like Covid-19 or cancer. This limits the number of deaths occurring as a result of drugs.

This may result in a negative relationship for the older age groups and a positive relationship for the younger age groups.

Thirdly, we expect changes within years as the death count increased during the pandemic seen in (**age-graph?**). The years pre and post Covid-19 are extremely similar. There was a clear increase during the pandemic but not the topic of this paper, therefor not studied further.

4 Results

(**age-graph?**) visualizes our first data of interest, showing deaths per year for each age group. Looking at this graph teaches us a few things. Firstly, homeless aged 40-59 are the most prone to death due to drug intoxication, followed by age group 20-39. Secondly, deaths by intoxication peaked during the years 2020-2022, these years are when the pandemic was most prevalent. Lastly, there seems to be no improvement in the homeless drug problem over this time period.

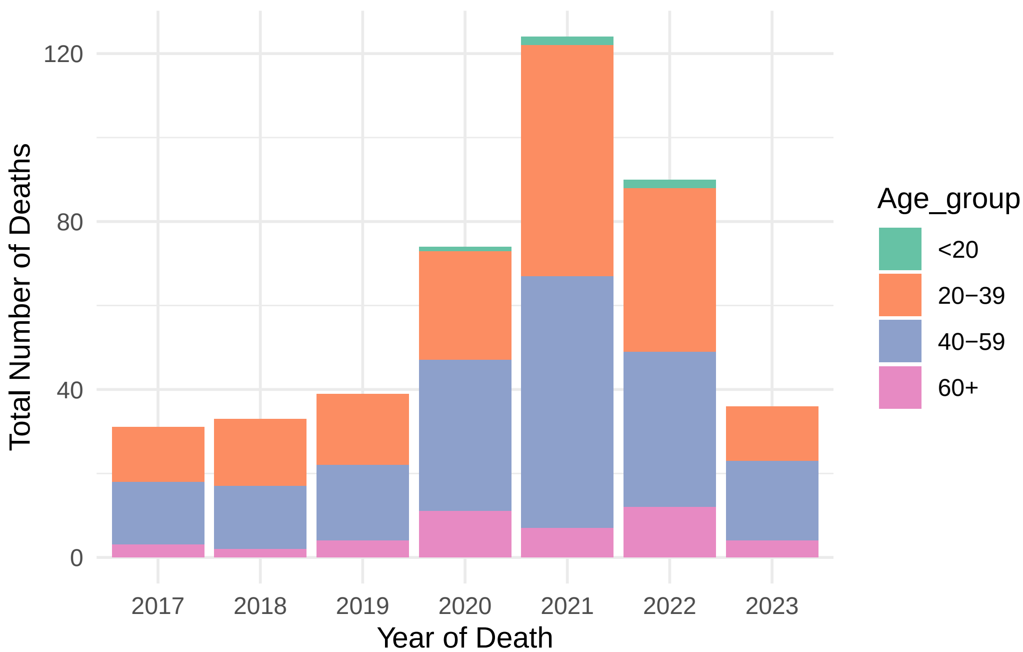


Figure 1: Deaths Per Year for Age Group

Observing (**age-graph?**) we can see that during the years 2017-2019 there are no significant changes in the death tolls reported for each age group. However, the trend does suggest that the death toll is on a steady incline. Once we reach 2020, we see a significant rise in deaths reported for almost every age group. This increasing trend continues for the next two years until 2023 where it stabilizes to how it was before 2020. From this we can incur that something irregular and unexpected occurred during the years 2020-2022 that caused our death reports to increase drastically, diverging from our regular trend. This becomes evident when we notice that our death reports fall back to our regular trend once we reach 2023.

(**age-graph?**) also shows us the trends for each age group. Overall, the death ratio of each age group compared to the total death toll seems to be consistent. Some major observations are as follows: homeless aged 40-59 years of age are the most prone to drug intoxicated death, followed by homeless aged 20-39, and finally individuals <20 years of age being least affected.

(**gender-graph?**) visualizes our second data set of interest, showing number of deaths for males and females from 2017-2023. We can learn a few more things on top of what (**age-graph?**) teaches us. From (**gender-graph?**) we can see that men see more deaths due to intoxication than women. We also see a much drastic increase in deaths by drug intoxication for men during the pandemic years compared to women. We also see a large decrease in deaths of women for the year 2023. Overall, we see no improvement in deaths for men, while we only see an improvement for women in the year 2023.

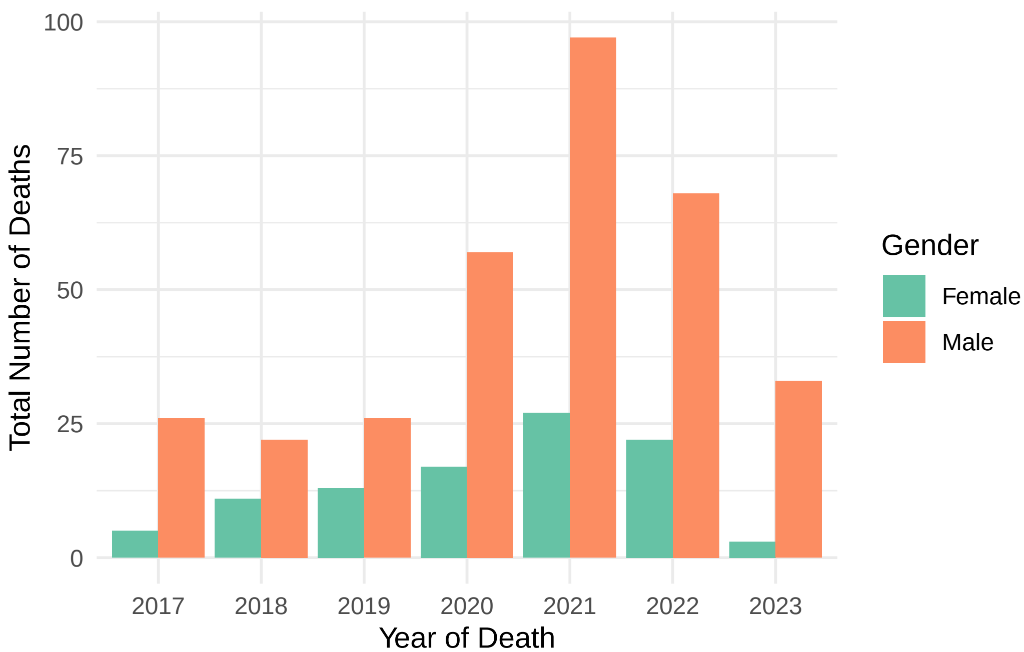


Figure 2: Deaths Per Year per Gender

By observing (**gender-graph?**) we can see that male mortality rate is much higher than female mortality due to drug toxicity. We also see that total number of male deaths peaks at a much higher ratio compared to female deaths. From this we can hypothesize that the effects of the pandemic had a greater affect on the male population subect to drug abuse than the female population. We also see that the female population experiences a much lower dip than the male population after the pandemic, surprisingly the lowest count from previous years.

Table 3: Explanatory models of flight time based on wing width and wing length

	Gender Only	Age Only	Age and Gender
(Intercept)	1.64	−0.12	−0.89
	(0.10)	(0.44)	(0.47)
GenderMale	0.98		1.10
	(0.11)		(0.11)
Age_group20-39		2.74	2.77
		(0.45)	(0.46)
Age_group40-59		2.77	2.85
		(0.45)	(0.46)
Age_group60+		1.47	1.41
		(0.49)	(0.49)
Num.Obs.	43	43	43
Log.Lik.	−240.304	−202.787	−149.961
ELPD	−249.3	−216.2	−162.5
ELPD s.e.	39.2	39.2	22.5
LOOIC	498.5	432.5	325.0
LOOIC s.e.	78.3	78.4	45.0
WAIC	498.9	432.8	324.3
RMSE	9.93	9.35	7.49

(**gender-graph?**) suggests that males occupy a majority of the homeless population as well as being more prone to drug abuse. This observation explains why we see a much greater number of male deaths compared to females. However, this does not explain why the pandemic affected the male population more greatly than the female population, and why did the death count for the female population drop so drastically for the year 2023? Looking at the (**age-graph?**) we can see that during 2023, the age group 20-39 saw a greatest dip pin death count. Coupling this with (**gender-graph?**) could suggest that a large number of females from the ages 20-39 saw improvements in deaths due to drug toxicity.

Our results are summarized in Table 3.

this table is for poisson

This table is for negative binomial model

can discuss difference in results between them and overall coefficients be care when intrepting coefficients in each model

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Limitations

The analysis presented in this study is subject to certain limitations, many of which stem from the method of data acquisition itself. The primary concern is sampling bias - according the opendatatoronto portal, only the deaths reported to the Toronto Public Health (TPH) by specific sources are included in the dataset. Consequently, any instances where deaths are not reported to the TPH could potentially underestimate the prevalence of drug related fatalities among the homeless. Furthermore, about 25% of cases from the raw dataset have either an unknown or pending cause of death, reducing the power of our model.The lack of clarity regarding the cause of death introduces uncertainty.

Furthermore, any death counts for transgender people were withheld due to the extremely small numbers. This decision discounts the impact of drug toxicity on this demographic, and illustrates the difficulty in assessing impacts on this specific subgroup. One can hypothesize that homeless gender-diverse individuals could be polarizingly more or less likely to overdose on drugs, but that can never be verified.

Finally, the portal reports that the data is updated regularly as new cases are reported and previous cases are updated with new information. This means that every time the dataset is updated our analysis becomes less accurate, and must be re-done after some period of time. While not exactly a shortcoming, it represents the dynamic nature of the data and its need to be revised consistently.

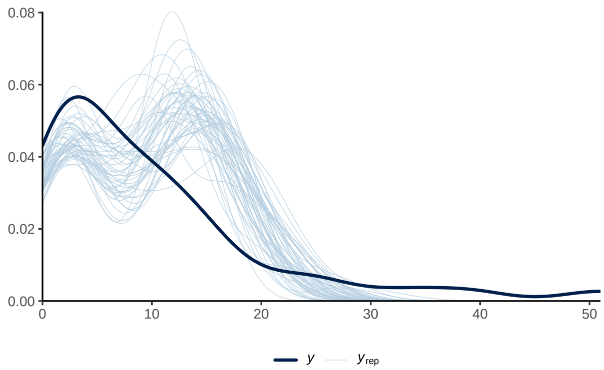
Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In Figure 3 we implement a posterior predictive check. This shows...

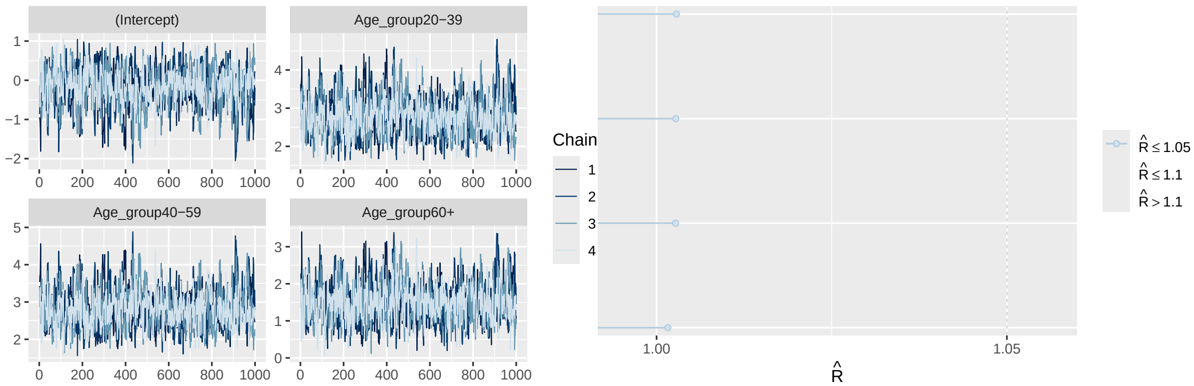


(a) Posterior prediction check

Figure 3: Examining how the model fits, and is affected by, the data

B.2 Diagnostics

Figure 4 is a trace and rhat plot. It shows... This suggests...



(a) Trace plot

(b) Rhat plot

Figure 4: Checking the convergence of the MCMC algorithm

References

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

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *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. 2019b. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

———, et al. 2019a. "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*.