Modeling Mortality Trends in Alberta*

A Quantitative Approach to Public Health Dynamics

Ziheng Zhong

March 14, 2024

Abstract

Table of contents

1 Introduction

Alberta, a province known for its vibrant culture and dynamic economy, faces ongoing challenges in public health, with mortality rates reflecting a complex interplay of socioeconomic, environmental, and healthcare factors. This paper aims to delve into the intricacies of mortality rates in Alberta, exploring the trends, causes, and determinants that shape the province's

---

*Code and data are available at: https://github.com/iJustinn/Alberta_Mortality_Rate.git

public health landscape. By leveraging data on causes of death, including chronic diseases, accidents, and emerging health threats, we provide a detailed analysis of mortality patterns and their implications for Alberta's healthcare system and policy-making.

2 Data

Data used in this paper was cleaned and processed with the programming language R (R Core Team 2022). Also with support of additional packages in R: `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `janitor` (Firke 2023), `readr` (Wickham, Hester, and Bryan 2023), `knitr` (Xie 2014), `rstanarm` (Goodrich et al. 2023), `modelsummary` (Arel-Bundock 2023), `tidybayes` (Kay 2023).

2.1 Source

In the pursuit of understanding and modeling mortality trends within Alberta, our study primarily utilizes the "Leading Causes of Death" dataset made available by the Alberta Open Government Portal. This comprehensive data source offers an extensive overview of mortality statistics across the province, categorized by various causes of death. The Alberta Open Government Portal, a key initiative for promoting transparency and public engagement, hosts this dataset at open.alberta.ca (Alberta Open Government Portal 2023), providing open access to valuable health-related data.

The "Leading Causes of Death" dataset encompasses a detailed compilation of death counts and mortality rates spanning multiple years. It includes categorizations by age, sex, and specific causes of death, ranging from chronic diseases to accidents and emergent health threats. This granularity allows for a nuanced analysis of mortality trends, facilitating the identification of prevalent health issues and potential public health interventions.

The selection of this data source was driven by its credibility, comprehensiveness, and relevance to our research objectives. Being an official government resource, the dataset ensures reliability and accuracy, crucial for the validity of our analytical models. The dataset's extensive coverage of demographic and cause-specific mortality information enables a thorough exploration of the factors influencing mortality rates in Alberta. Furthermore, the data's organization and accessibility significantly aid in the efficient processing and modeling of mortality trends.

2.2 Method

Our analysis commenced with the essential task of refining the raw dataset. The initial phase involved cleaning and standardizing the dataset's variable names to enhance readability and analytical efficiency. We also aggregated the data to quantify the prevalence of each cause of death, thereby streamlining the dataset for more focused analysis. A critical aspect of our

Table 1: top 8 major causes in 2019

Year	Cause	Ranking	Deaths	Years
2019	Organic dementia	1	1,997	22
2019	All other forms of chronic ...	2	1,886	22
2019	Malignant neoplasms of trac...	3	1,523	22
2019	Other chronic obstructive p...	4	1,159	22
2019	Acute myocardial infarction	5	1,061	22
2019	Atherosclerotic cardiovascu...	6	678	22
2019	Accidental poisoning by and...	7	677	10
2019	Stroke, not specified as he...	8	602	22

methodology was the isolation of data from the year 2019. This strategic choice was made to examine mortality trends immediately before the onset of the COVID-19 pandemic. By focusing on this specific year, we aimed to establish a baseline understanding of mortality patterns unaffected by the pandemic's impact. Within this scope, as showing in Table 1, we identified and prioritized the top 8 causes of mortality, based on their significance and prevalence.

Building on the refined dataset, our analysis further narrowed to investigate the top 5 causes of mortality in 2019. This deliberate focus allowed us to concentrate our efforts on the most impactful health challenges faced by Alberta before the pandemic. We applied statistical models to fit these mortality trends, seeking to identify underlying patterns and determinants. This approach was aimed at uncovering insights that could inform public health strategies and interventions, contributing to a more nuanced understanding of mortality dynamics in the province.

3 Results

3.1 Data Trend

Results in the Figure 1a displays the annual number of deaths in Alberta for the top five causes in 2019 unveils distinct trends for each cause over approximately two decades. The data suggest a relatively stable trend in deaths from acute myocardial infarction, with a small rise that may indicate a steady prevalence of risk factors or an equilibrium between advancing medical interventions and demographic shifts. Conversely, the increasing trajectory seen in deaths due to chronic diseases points to a potential escalation in lifestyle-related health issues or could reflect the demographic realities of an aging population.

Particularly notable is the sharp increase in deaths attributed to organic dementia, surging upwards more markedly in the latter half of the time span. This trend could signify heightened

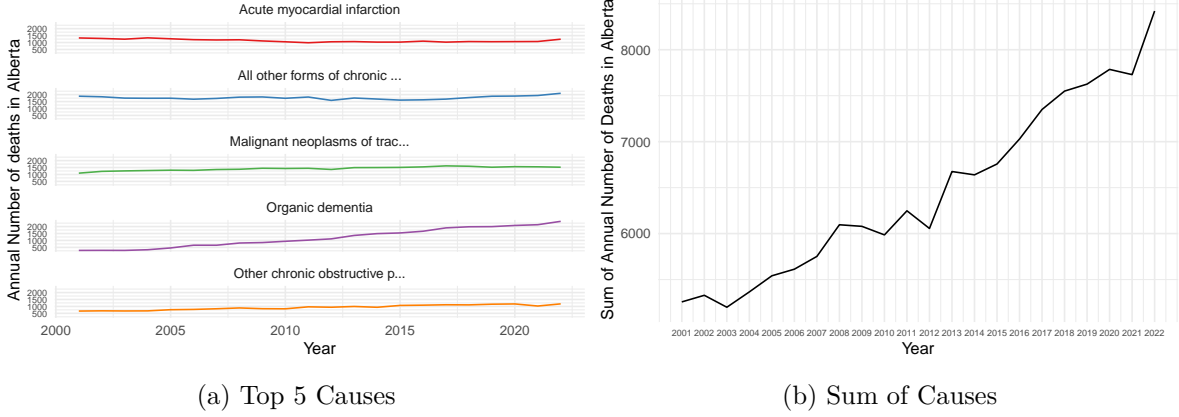(a) Top 5 Causes   (b) Sum of Causes

Figure 1: Trend for Alberta Mortality

awareness and improved diagnostic capabilities, alongside the impact of an aging society. Meanwhile, deaths from malignant neoplasms related to the respiratory system display minor fluctuations, hinting at persistent challenges in cancer prevention, detection, and treatment. The upward trend in chronic obstructive pulmonary disease-related mortality could be linked to environmental factors or historic smoking patterns, underscoring the pressing need for robust public health policies and effective healthcare strategies.

The graph presented in Figure 1b illustrates a clear upward trajectory in the total number of deaths from 2001 to 2022. This trend, while indicative of a rising mortality rate, necessitates a nuanced interpretation. The increasing death toll may not necessarily be attributed to the lethality or prevalence of specific causes alone. Instead, it potentially reflects broader demographic changes, such as the growth of the overall population, particularly within older age groups who are statistically more vulnerable to various health-related fatalities.

3.2 Modeling

Then the Table 2 summarizes the coefficients of both models built to fit the mortality trend in Alberta, offering a comparison of their estimations for the top causes of mortality. In both models, "other chronic" conditions and "neoplasms" exhibit positive coefficients, suggesting these causes of death have an increasing trend over time. Meanwhile, "obstructive pulmonary" diseases show a negative relationship in both models, indicating a decreasing trend in mortality rates. The consistency across both models reinforces the reliability of our findings.

For the Poisson model, the expected log count is defined as:

$$\log(\mu_i) = \beta_0 + 0.447 \times (\text{other chronic})_i + 0.223 \times (\text{neoplasms})_i + 0.046 \times (\text{dementia})_i -$$

4

Table 2: summary of modeling

	Poisson	Negative binomial
other chronic	0.447	0.449
		(0.102)
neoplasms	0.223	0.226
		(0.100)
dementia	0.046	0.048
		(0.101)
obstructive pulmonary	-0.205	-0.203
		(0.104)
Num.Obs.	110	110
Log.Lik.	-5718.182	-810.934
ELPD	-5926.8	-815.4
ELPD s.e.	1216.0	10.6
LOOIC	11 853.7	1630.8
LOOIC s.e.	2432.0	21.2
WAIC	11 954.1	1630.8
RMSE	325.38	325.39

$$0.205 \times (\text{obstructive pulmonary})_i$$

For the Negative Binomial model, the expected log count is defined as:

$$\log(\mu_i) = \beta_0 + 0.449 \times (\text{other chronic})_i + 0.226 \times (\text{neoplasms})_i + 0.048 \times (\text{dementia})_i -$$

$$0.203 \times (\text{obstructive pulmonary})_i$$

Then to see how good two models fits the original data. We can reffer to the values in the lower half of Table 2, but it's always more clear to see the visual, as depicted by the graphs for the Poisson (Figure 2a) and negative binomial models (Figure 2b), the diagnostic plots further elucidate the model fits. The negative binomial model, appears to capture the dispersion in the data more effectively than the Poisson model, which is evident in two plots. The spread of the simulated lines (y_rep) around the observed data (y) suggests that the negative binomial model may accommodate the overdispersion present in the data, which is a common characteristic in count data such as mortality rates.

The presence of multiple peaks in the second plot's simulated data distributions indicates potential subgroups within the data or non-linearity in the mortality trends that the Poisson model may not be accounting for. Such patterns underscore the importance of considering different model structures when dealing with complex datasets. These graphical diagnostics
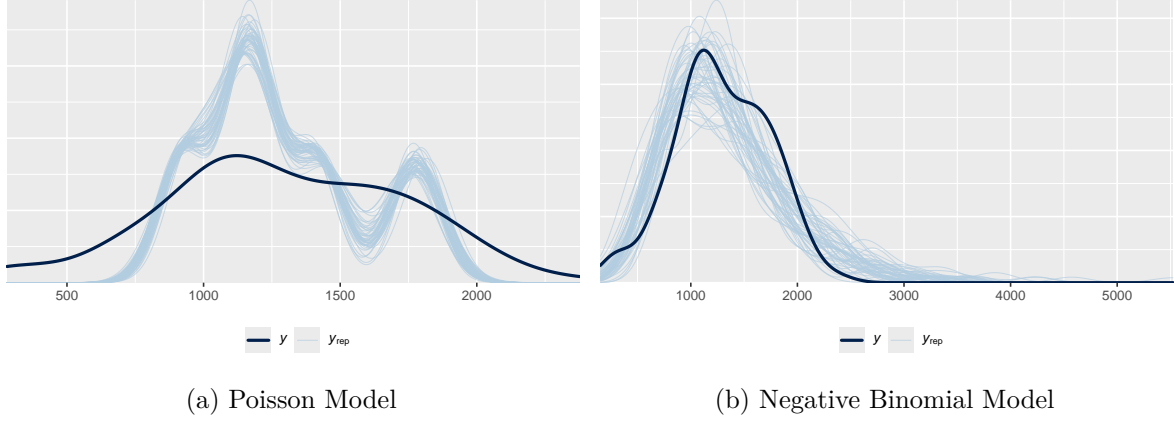
(a) Poisson Model (b) Negative Binomial Model

Figure 2: Model Results

not only affirm the adequacy of the negative binomial model over the Poisson model for this particular dataset but also highlight the necessity of thorough model checking in the process of statistical analysis.

4 Discussion

The preceding sections have laid the groundwork for a comprehensive analysis of mortality trends in Alberta, paving the way for this discussion. The application of Poisson and Negative Binomial models to the data has illuminated key trends and has provided quantitative support to the qualitative observations made earlier. The positive coefficients for "other chronic" and "neoplasms" in both models suggest that deaths from these causes have steadily increased, raising questions about the broader factors at play beyond the healthcare system's direct influence. Simultaneously, the negative coefficient for "obstructive pulmonary" diseases suggests a decline in mortality from these conditions, potentially reflecting successful public health initiatives and changing social habits, such as reduced smoking rates.

However, the increasing overall mortality trend signals a need to consider demographic shifts, such as population growth and aging, as significant contributors to the observed patterns. The spike in deaths from 2020 onwards aligns with the onset of the COVID-19 pandemic, underscoring the substantial impact of infectious disease outbreaks on mortality statistics. This outlier emphasizes the importance of having robust health surveillance systems that can adapt to sudden changes in public health landscapes.

Furthermore, the diagnostic plots from both models underscore the necessity for careful selection of appropriate statistical methods when modeling complex health data. The better fit of the Negative Binomial model compared to the Poisson model, as indicated by the diagnostic

checks, suggests that overdispersion—a common feature in count data—must be accounted for to avoid underestimating the variability in the data.

The limitations of the current study also warrant discussion. While the models applied provide valuable insights, they do not capture all the nuances of the data. Future research could explore the inclusion of additional variables, such as economic indicators, healthcare access measures, and lifestyle factors, to better understand the determinants of mortality. Moreover, the investigation of interactions between causes of death and demographic variables could yield a more granular understanding of the dynamics at play.

Ultimately, this study serves as a call to action for ongoing surveillance, research, and a proactive approach to public health planning. By continuing to refine our models and expand our datasets, we can better anticipate and mitigate the factors contributing to mortality, striving for a healthier future for all Albertans.

5 Conclusion

In conclusion, our exploration of mortality trends in Alberta has provided a statistical narration of the province's public health status over the past two decades. We have identified critical causes of death and trends that warrant attention from both healthcare providers and policymakers. The analysis has also emphasized the significance of employing appropriate statistical models to accurately represent and understand complex health data. As we move forward, it is imperative to integrate these findings into a strategic framework that enhances Alberta's health system responsiveness and effectiveness in the face of demographic and epidemiological shifts. Through vigilance and informed action, we can aim to influence these mortality trends positively, enhancing the health and well-being of the Alberta population.

References

Alberta Open Government Portal. 2023. "Leading Causes of Death." https://open.alberta.ca/opendata/leading-causes-of-death.

Arel-Bundock, Vincent. 2023. *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready.* https://vincentarelbundock.github.io/modelsummary/.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

Goodrich, Ben, Jonah Gabry, Imad Ali, Sam Brilleman, and and others. 2023. *Rstanarm: Bayesian Applied Regression Modeling via Stan.* https://mc-stan.org/rstanarm.

Kay, Matthew. 2023. *Tidybayes: Tidy Data and Geoms for Bayesian Models.* https://mjskay.github.io/tidybayes/.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

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, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data.* https://CRAN.R-project.org/package=readr.

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