

2023 Ontario Monthly Communicable Disease Specified Cases Study*

HIV, Salmonellosis, and Influenza cases Analysis

Ruiyang Wang

September 24, 2024

This paper examines the monthly trends of three communicable diseases—HIV, Salmonellosis, and Influenza(sporadic) - in 2023 of Ontario. We use data from the City of Toronto’s Open Data Portal, and conduct an analysis to identify seasonal patterns and peaks for each disease. For HIV, the data shows a notable peak, suggesting potential influences on inappropriate social activities in the certain season. Salmonellosis shows a significant rise of case number during the summer, highlighting the importance of the role of environmental and food safety problems. Influenza cases, increased dramatically in December, marking the onset of the flu season. These trends underscore the need for targeted public health interventions, such as enhanced awareness campaigns during identified peak periods and improved vaccination coverage for influenza. Understanding these patterns is crucial for optimizing resource allocation, and improving healthcare systems decisions in Ontario.

1 Introduction

In 2023 Ontario, Canada, experienced distinct trends in the prevalence of communicable diseases such as HIV, Salmonellosis, and Sporadic Influenza. Monitoring and analyzing the seasonal patterns of these diseases is crucial for public health.

HIV, a chronic and life-threatening condition caused by the human immunodeficiency virus, has shown relatively stable case numbers throughout the year.

Salmonellosis is an infection caused by the Salmonella bacteria, typically contracted through the consumption of contaminated food or water. It commonly results in gastrointestinal symptoms such as diarrhea, fever, and abdominal cramps.

*Code and data are available at: <https://github.com/Ruiyang-Wang/STA304-Paper-1.git>

Sporadic Influenza, commonly known as the flu, is a highly contagious respiratory illness caused by influenza viruses.

This paper aims to analyze the statistical trends of these three diseases in Ontario throughout 2023, providing a general overview of their seasonal patterns. We also identify potential drivers of these trends and discuss their implications for public health policy. The findings will contribute to a better understanding of disease dynamics in the region and then support the development of prevention and strategies.

We use data and codes from (R Core Team 2023), (Gelfand 2022), and (Wickham et al. 2019a). The exact data is from (Toronto 2023).

The remainder of this paper is structured as follows.

- Section 1 Introduction
- Section 2 Data Summary
- Section 3 Disease Analysis
- Section 3.1 HIV
- Section 3.2 Salmonellosis
- Section 3.3 Sporadic Influenza
- Section 3.4 Yearly Pattern
- Section 3.5 Weakness and Next Step
- Section A Appendix
- Section A References

2 Data Summary

Our data is from (Gelfand 2022), named “Monthly Communicable Disease Surveillance”. The time range of the whole dataset is from 2016 to 2024. We only extract 3 classic diseases in this dataset for simplification, and set the year of analysis to 2023, which is the most recent full-year data and most unaffected data by COVID-19 after the 2019 Pandemic.

Table 1: Monthly Cases of Diseases (At a Glance)

Disease	2023 YTD Cases	2023 YTD Rate1	January	February	March
HIV	732	24.70	56	54	55
Salmonellosis	456	15.40	40	34	34
Influenza - sporadic	2892	97.50	143	78	203

Figure 1: Diseases Data Table

2.1 Variable Explanation

“Disease”: This variable represents the specific communicable diseases being analyzed in the dataset. “January - December” : These variables indicate the number of reported cases of each disease for each month in the year 2023. They provide a month-by-month breakdown of disease prevalence.

2.2 Data Cleaning

I have cleaned the data by eliminating all unused diseases, and cancelled out all NA values or labels. I also changed the data type of numbers from string to double, in order to make sure the graphing process is accurate.

3 Disease Analysis

3.1 HIV

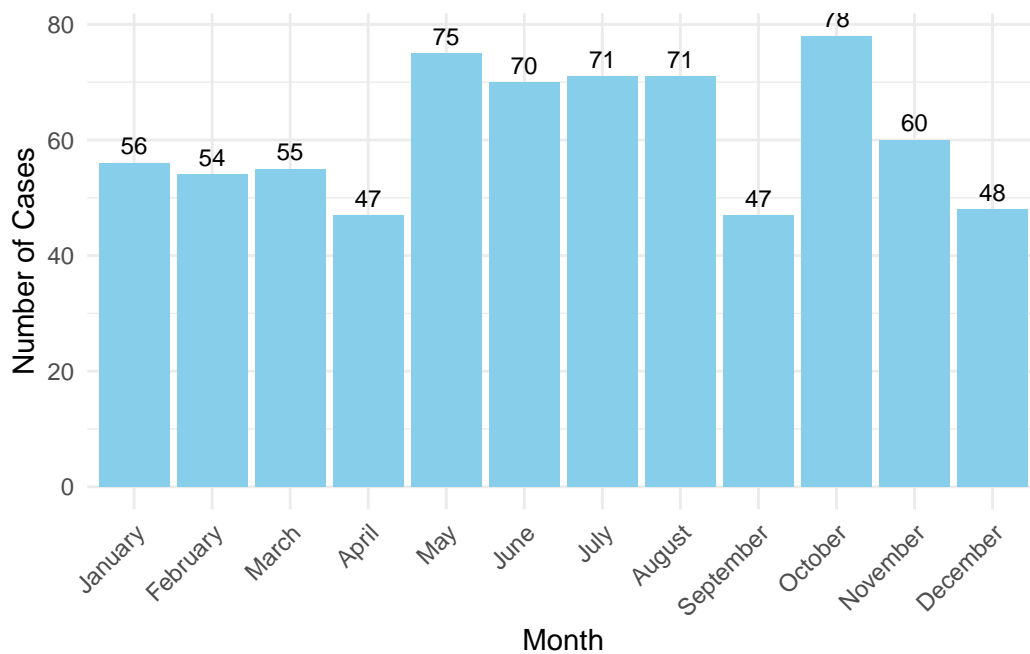


Figure 2: Monthly cases of HIV

According to Section 3.1, the number of HIV cases ranges from around 47 to 78 per month. There are noticeable fluctuations throughout the year, with peaks in May and October and a

significant drop in April and September. The case count is relatively stable in other months, hovering between 50 and 75 cases.

Overall, HIV cases in Ontario displayed moderate monthly fluctuations, with notable peaks in May and October. These variations could be influenced by changes in testing rates or social behaviors.

3.2 Salmonellosis

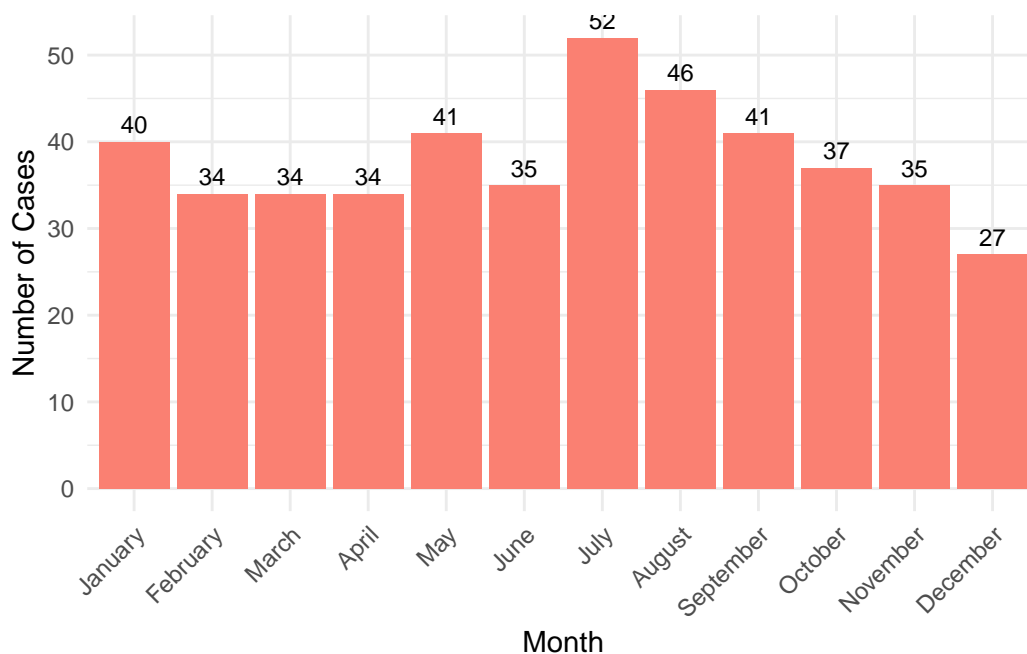


Figure 3: Monthly cases of Salmonellosis

According to Section 3.2, the number of cases ranges from 27 to 52 per month. There is a significant peak in July, with a steady increase from April, followed by a decline from August to December. Salmonellosis cases are highest in summer months, which is consistent with the bacteria's transmission patterns.

Overall, Salmonellosis cases peaked in the summer months, particularly in July, likely reflecting increased risk factors such as outdoor food consumption and higher temperatures.

3.3 Sporadic Influenza

According to Section 3.3, there is a sharp increase in cases in December, reaching a peak of 1560 cases. The rest of the year shows relatively low numbers. A smaller peak in March and a gradual increase from October to November indicate sporadic activity.

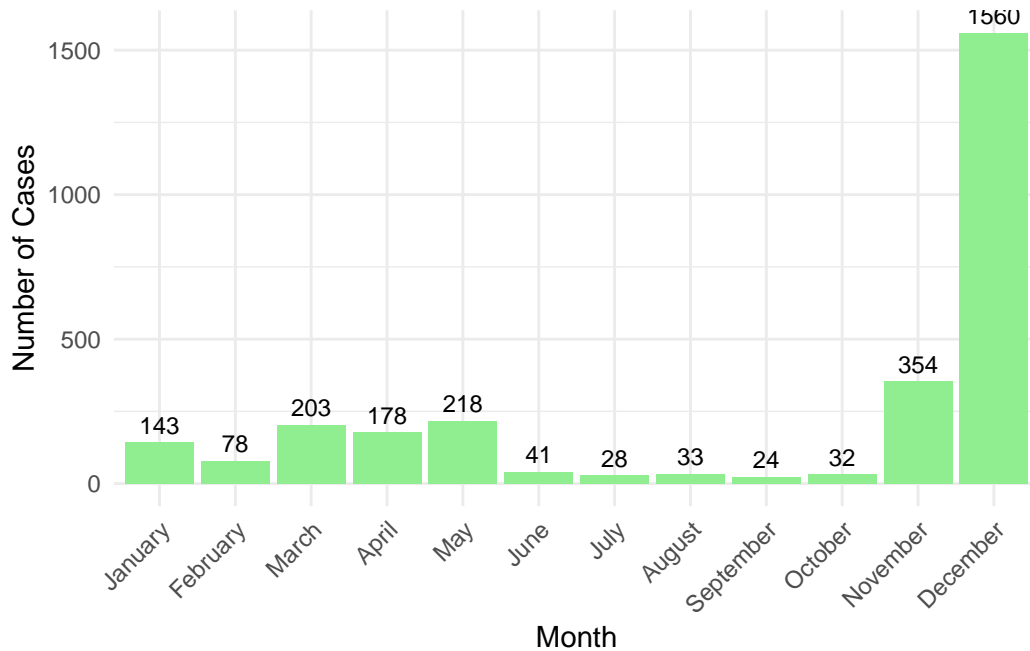


Figure 4: Monthly cases of Sporadic Influenza

Overall, Sporadic Influenza cases spiked sharply in December, marking the onset of the flu season. A smaller peak in March suggests occasional outbreaks earlier in the year.

3.4 Yearly Pattern

Section 3.4 illustrates the yearly pattern of HIV, Salmonellosis, and Influenza in Ontario for the year 2023. The most prominent feature is the dramatic increase in Influenza cases observed in December, reaching a peak of over 1500 cases. This surge matches the flu season in Canada. There are several factors due to the spike in December, including lower temperatures, increased indoor gathering, etc. The Ontario Health Plan for an Influenza Pandemic emphasizes infection prevention and control measures, particularly in occupational health and safety settings. This plan outlines protocols for reducing the transmission of influenza among healthcare workers and vulnerable populations during an outbreak (“Ontario Health Plan for an Influenza Pandemic - Chapter 5: Occupational Health & Safety and Infection Prevention & Control” 2023). We might expect the influenza cases decrease as time passes.

In contrast, HIV cases show a relatively stable trend throughout the whole 2023. The pattern include a few fluctuations observed during May and October. These peaks could be possibly linked to public health observances such as HIV Testing Week in May and World AIDS Day in December. The consistency in HIV cases also suggests that while the disease

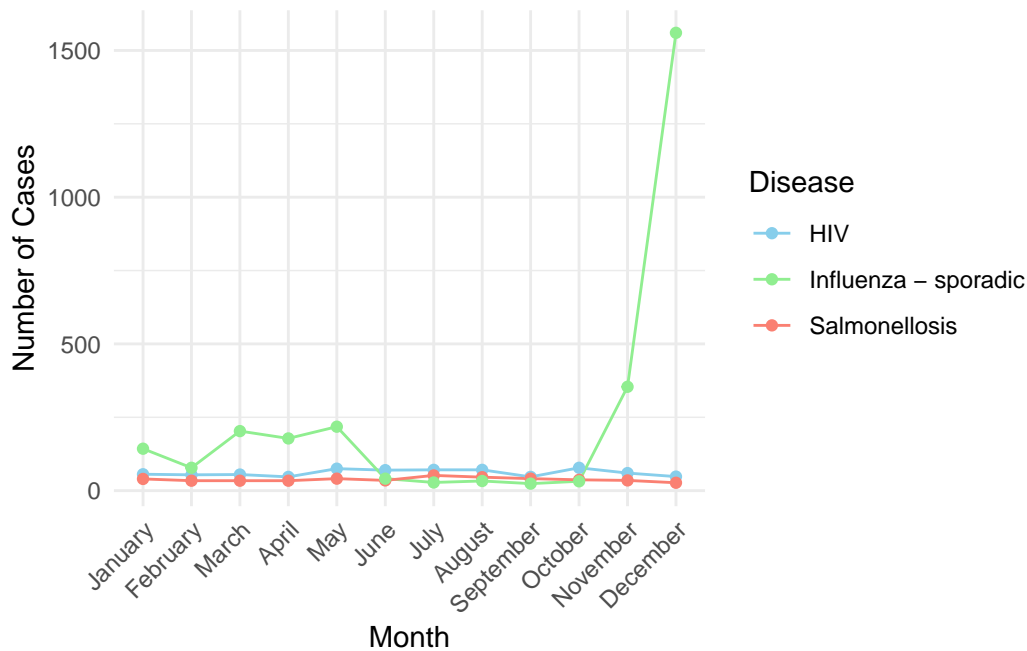


Figure 5: Yearly Pattern of three diseases

remains a public health concern, preventions and medical treatments might help controlling the spread.

Salmonellosis cases remained relatively low and stable across all months, with no significant seasonal spikes. This pattern is consistent with a well-regulated food safety environment and public health measures in Canada. One notable initiative is the Ontario Food and Nutrition Strategy (OFNS), which integrates the food, agriculture, and nutrition sectors to improve food safety, access, and literacy across the province. This strategy outlines key strategic directions, including healthy food access, food literacy, and the promotion of resilient food systems (“Ontario Food and Nutrition Strategy: A Comprehensive Evidence-Informed Plan for Healthy Food and Food Systems in Ontario” 2016). Salmonellosis is often linked to the consumption of contaminated food, and the lack of significant spikes might indicate effective food safety regulations and awareness among the population.

3.5 Weaknesses and Next Steps

3.5.1 Limitations and Weaknesses of the Data Processing Process

First, the analysis is limited to the year 2023, providing a short view without the ability to observe longer-term trends or compare with previous years. Data Figure 1 might be affected by the COVID-19 symptoms due to pharmacological reasons.

Then, the data does not include contextual information such as socioeconomic factors, healthcare access, or specific public health interventions, which might be crucial for understanding the causes of disease trends.

Thirdly, the use of basic descriptive statistics and visualizations limits the depth of insights that can be derived, lacking more advanced statistical or predictive modeling techniques.

3.5.2 Further Analysis for Each Disease

1. HIV

First, we need to conduct a more detailed analysis of potential seasonal factors affecting HIV cases, such as changes in social behaviors, festivals, or events, which might influence the transmission rates. Then, we should evaluate the impact of specific public health initiatives, such as awareness campaigns, increased testing, or outreach programs, to determine their effectiveness in reducing HIV transmission during peak months. At last, we can extend the analysis to include data from previous years to identify long-term trends and changes in HIV incidence over time, adjusting for any policy or healthcare changes.

2. Salmonellosis

First, we need to perform an investigation of foodborne outbreaks, and link the reported cases with specific food sources and outbreak reports from health agencies. Then, we should use some environmental data and behavioral data (e.g., outdoor gatherings, food festivals) to build a model to explain seasonal fluctuations for cases of Salmonellosis.

3. Sporadic Influenza

First, we can use additional data of vaccination and climates, develop a predictive model for influenza cases, in order to forecast future outbreaks and improve Ontario health systems. Then, we may analyze the impact of the winter peak on healthcare resources, like hospitalizations and ICU admissions, to evaluate the preparedness and resilience of the healthcare system during the flu season.

A Appendix

In this analysis, we utilized several data sources and tools. The dataset for communicable disease trends was obtained from the City of Toronto’s Open Data Portal (Toronto 2023). Data processing and visualization were performed using the `opendatatoronto` package (Gelfand 2022) in the R environment (R Core Team 2023).

The `tidyverse` package was a crucial tool for data wrangling and visualization (Wickham et al. 2019b). Additionally, the `Lahman` package was used for data manipulation (Friendly et al. 2020), and Bayesian regression models were fitted using `rstanarm` (Goodrich et al. 2022).

The analysis of the communicable disease trends showed significant seasonal variations in the data (Gebru et al. 2021).

Future work should focus on integrating more comprehensive datasets to provide a better understanding of the disease dynamics (Wickham et al. 2019a).

References

- Friendly, Michael, Chris Dalzell, Martin Monkman, and Dennis Murphy. 2020. *Lahman: Sean “Lahman” Baseball Database*. <https://CRAN.R-project.org/package=Lahman>.
- Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. “Datasheets for Datasets.” *Communications of the ACM* 64 (12): 86–92.
- Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. “Rstanarm: Bayesian Applied Regression Modeling via Stan.” <https://mc-stan.org/rstanarm/>.
- “Ontario Food and Nutrition Strategy: A Comprehensive Evidence-Informed Plan for Healthy Food and Food Systems in Ontario.” 2016. <https://sustainontario.com/custom/uploads/2012/04/OFNS-Final-Report-v3.1-April-8-2016.pdf>.
- “Ontario Health Plan for an Influenza Pandemic - Chapter 5: Occupational Health & Safety and Infection Prevention & Control.” 2023. <https://www.publications.gov.on.ca/ontario-health-plan-for-an-influenza-pandemic-chapter-5-occupational-health-safety-and-infection-prevention-control>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Toronto, City of. 2023. “Monthly Communicable Disease Surveillance Data.” <https://open.toronto.ca/dataset/monthly-communicable-disease-surveillance-data/>.
- 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>.