The Evolving Landscape of Acute Hepatitis C in the United States: A Demographic Breakdown

Dana Gonzalez

Introduction

Hepatitis C virus (HCV) infection is the most common blood-borne infection in the United States. National surveillance data has shown a significant rise in acute HCV cases (those in the first 6 months of infection) in the last decade, with case counts having nearly doubled. The majority of acute cases (>50%) will develop into chronic hepatitis C, which can lead to complications like liver disease, liver cancer, liver failure, or even death (Centers for Disease Control and Prevention [CDC], 2024).

In lieu of a functional vaccine, noninvasive direct-acting antivirals (DAAs) have proven to be significantly effective treatments for HCV, curing nearly 95% of all cases in which they are utilized (Jeong et al., 2024). However, the asymptomatic nature of HCV, risk-associated stigmas, and other financial and geographical barriers prevent disadvantaged groups from accessing appropriate diagnostic and therapeutic care.

While the "baby boomer" generation (those born between 1946 and 1964) has been identified as a key population due to their likelihood of receiving a blood transfusion before universal screening guidelines, an increasing number of new cases have been reported in younger individuals (Rose et al., 2019).

This shift in the disease landscape is largely driven by the United States' ongoing opioid crisis. Opioid use is largely associated with injection drug use (IDU), which is now being identified as the primary risk factor for HCV and other blood-borne pathogens like HIV (CDC, 2024).

Due to the asymptomatic nature of acute HCV, the true burden of this disease is grossly underestimated. For this reason, it is critical that current at-risk populations are promptly identified to more effectively target public health interventions to connect at-risk and diseased individuals with appropriate preventative, diagnostic, and therapeutic care.

For this reason, I will work to visualize trends in acute HCV incidence across the following demographic groups for the period 2000-2022 and determine which group(s) within each demographic factor have seen the greatest rise in incidence for the period of interest.

I will evaluate rises in incidence by fitting linear models to plots of each group and examine each subgroup's slope-intercept equation. I will interpret the group(s) with the greatest positive slope as having the greatest rise in incidence for the period of interest.

The demographic groups included in my analysis are as follows:

- **Racial/Ethnic Groups:**
  - American Indian/Alaska Native
  - Asian/Pacific Islander
  - Black/African American
  - Hispanic/Latino
  - White
  - Other
  - Unknown
- **Age Groups:**
  - 0-19 years
  - 20-29 years
  - 30-39 years
  - 40-49 years
  - 50-59 years
  - 60 years
- **Sex:**
  - Male
  - Female

For additional reference (will not be included in my analysis), I will visualize trends in incidence rates across the following to gain greater insight into the disease landscape:

- **Geographic Location (2022):**
  - 50 U.S. States (**Visualization only included on website due to incompatibility of plotly with PDF outputs**)
- **Urbanicity (2018-2022):**
  - Urban
  - Rural

Methods

All data for this project was retrieved from the Centers for Disease Control and Prevention's (CDC) National Center for HIV, Viral Hepatitis, STD, and Tuberculosis Prevention (NCHHSTP).

Incidence data for racial/ethnic, age, and sex groups was retrieved via the official NCHHSTP site. This center's site provides access to AtlasPlus, an interactive tool that allows users to download disease surveillance data from the last 20 years. Using this tool, I downloaded national incidence rates for the above-mentioned groups from 2000 to 2022.

State and urbanicity data was collected from the NCHHSTP's 2022 Viral Surveillance Report. From 2007 onwards, the NCHHSTP has published an annual comprehensive report of hepatitis A, B, and C incidence and mortality rates (as well as case counts) across various demographic and geographic factors. These reports are accompanied by annual National Progress Reports, which describe these measures in reference to the CDC's goals for new viral hepatitis infections and viral hepatitis-related deaths.

Data was downloaded from their respective sources and loaded into R as CSVs. After all datasets had been properly loaded and renamed (i.e., "Estimated_Vs_Reported_Data" into "estimated"), I began the process of cleaning and wrangling the data:

- **Racial/Ethnic, Age, and Sex Group Data**:
  - Remove "Indicator" variable (consistent for all observations)
  - Remove "Cases" variable (analysis limited to incidence rates)
  - Exclude data for "Other" and "Unknown" racial/ethnic groups (analysis limited to known groups)
  - Remaining variables renamed for easier recall (i.e., "Rate.Per.100000" to "incidence")
  - Change "group" variable from a character to a factor variable type
  - Change "year" and "incidence" variables from character to numeric variable types
- **State Data**:
  - Remove "Range" variable (analysis limited to incidence rates)
  - Remove "District of Columbia" data (for easier analysis using "plotly" package)
  - Remaining variables renamed for easier recall (i.e., "Cases.100000.Population" to "incidence")
  - Change "state" variable from a character to a factor variable type
  - Change "incidence" variable from a character to a numeric variable type
- **Urbanicity Data**:
  - Rename variables for consistency across datasets (i.e., "Urban" to "urban")
  - Change "area" variable from a character to a factor variable type

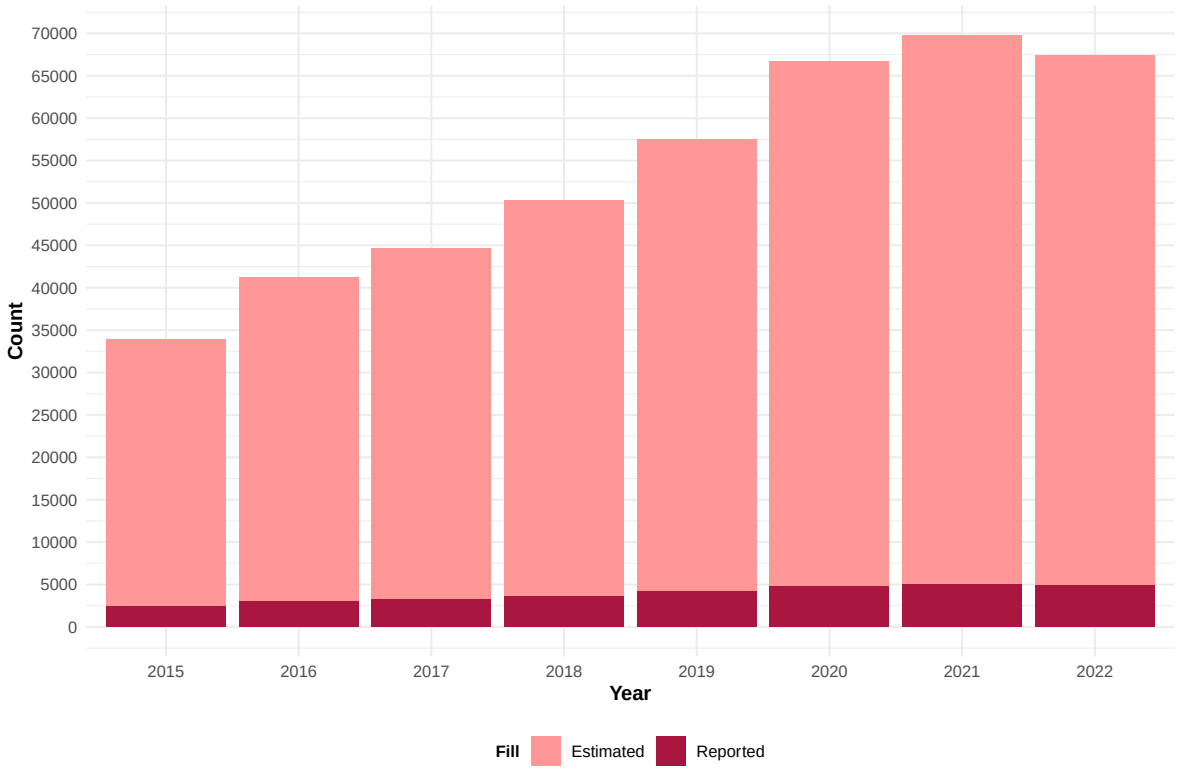
– Change “year” and incidence” variables from character to numeric variable types

Finally, I manually inspected each dataset by using R’s str() function to confirm my edits were successful.

Linear models were fitted using R’s “lm()” function. Slope-intercept equations were then displayed using the “kable()” function.

Results

Figure 1. Stacked Barchart of Reported vs. Estimated Cases of Acute HCV (2015-2022)



Although I am not directly utilizing the following data for this project, I wanted to briefly include it to emphasize the need to identify at-risk groups for HCV infection. As previously stated, true HCV incidence is severely underreported due to the asymptomatic nature of this disease. Identifying key at-risk groups will allow public health officials to implement more effective public health interventions and screening efforts for better disease surveillance.

As seen in Figure 1 above, there is a clear and stark difference in the volume of estimated and reported cases of acute HCV for each year. According to recent modeling, the CDC estimates

that 12.3 cases of HCV occur for every case reported in national surveillance (Klevens et al., 2014).

Figure 2a. Lineplot of Yearly Incidence (per 100,000) by Race/Ethnicity (2000-2022)

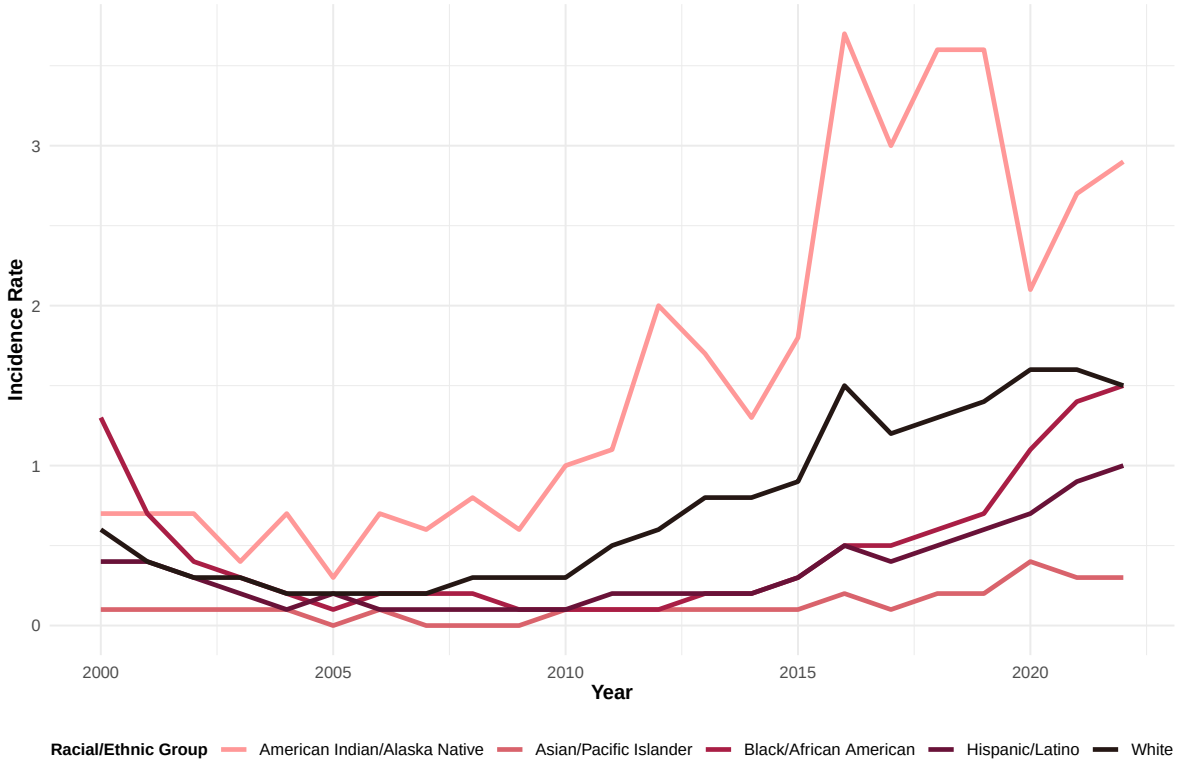


Figure 2b. Faceted Lineplot of Yearly Incidence (per 100,000) by Race/Ethnicity (2000-2022)

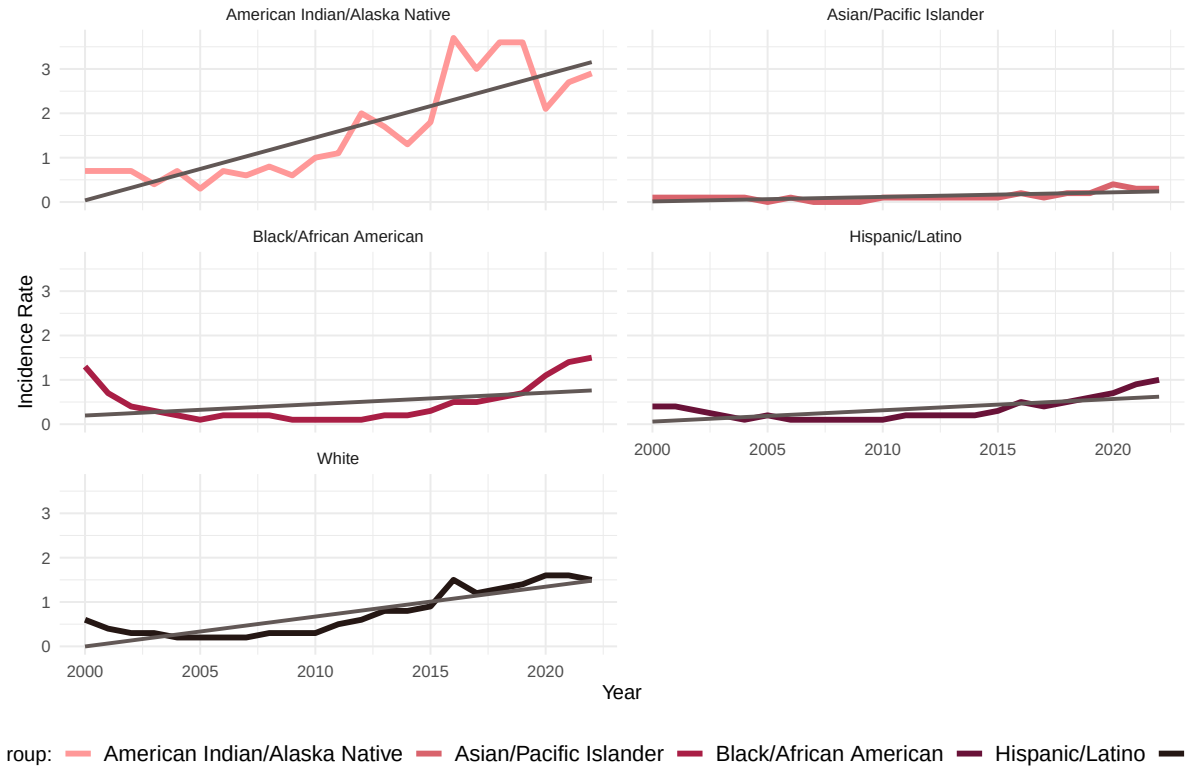


Table 1. Linear Regression Equations for Incidence Rates (per 100,000) by Race/Ethnicity (2000-2022)

Racial/Ethnic Group	Equation (y = b + m*x)
American Indian/Alaska Native	y = -283.36 + 0.14 * x
Asian/Pacific Islander	y = -20.54 + 0.01 * x
Black/African American	y = -51.19 + 0.03 * x
Hispanic/Latino	y = -50.93 + 0.03 * x
White	y = -134.78 + 0.07 * x

Figure 3a. Lineplot of Yearly Incidence (per 100,000) by Age (2000-2022)

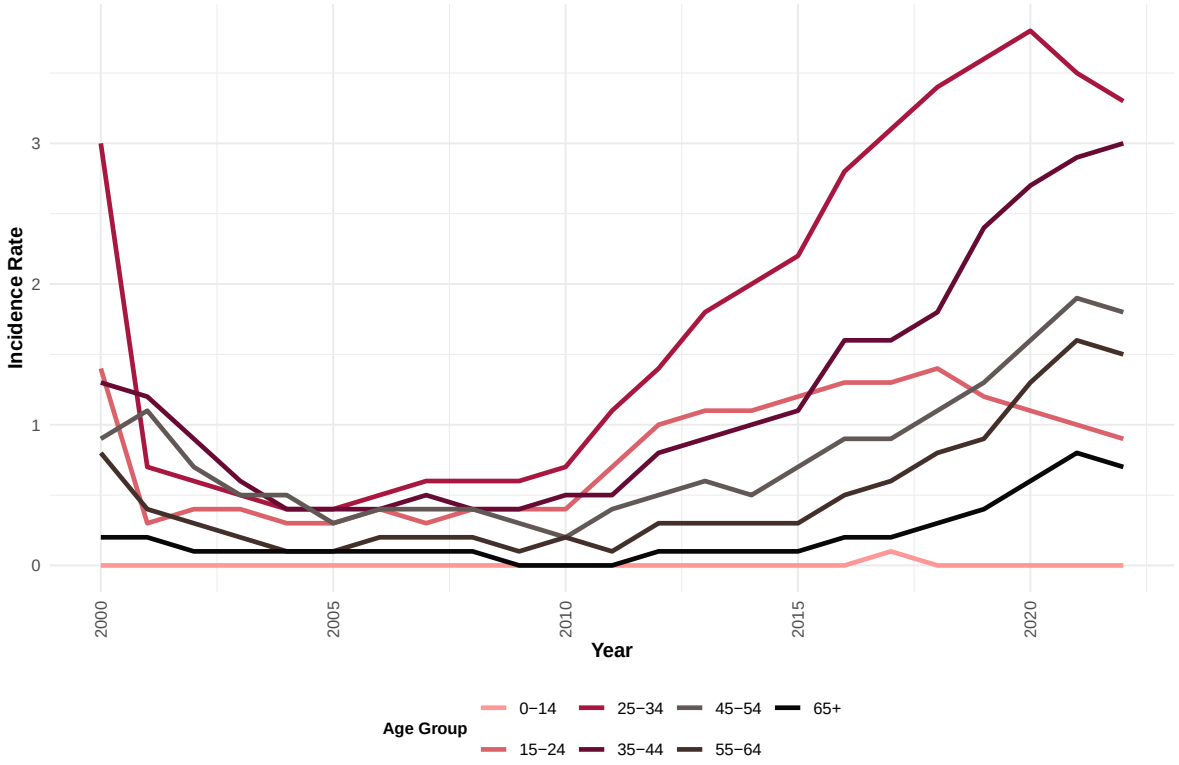


Figure 3b. Faceted Lineplot of Yearly Incidence (per 100,000) by Age (2000-2022)

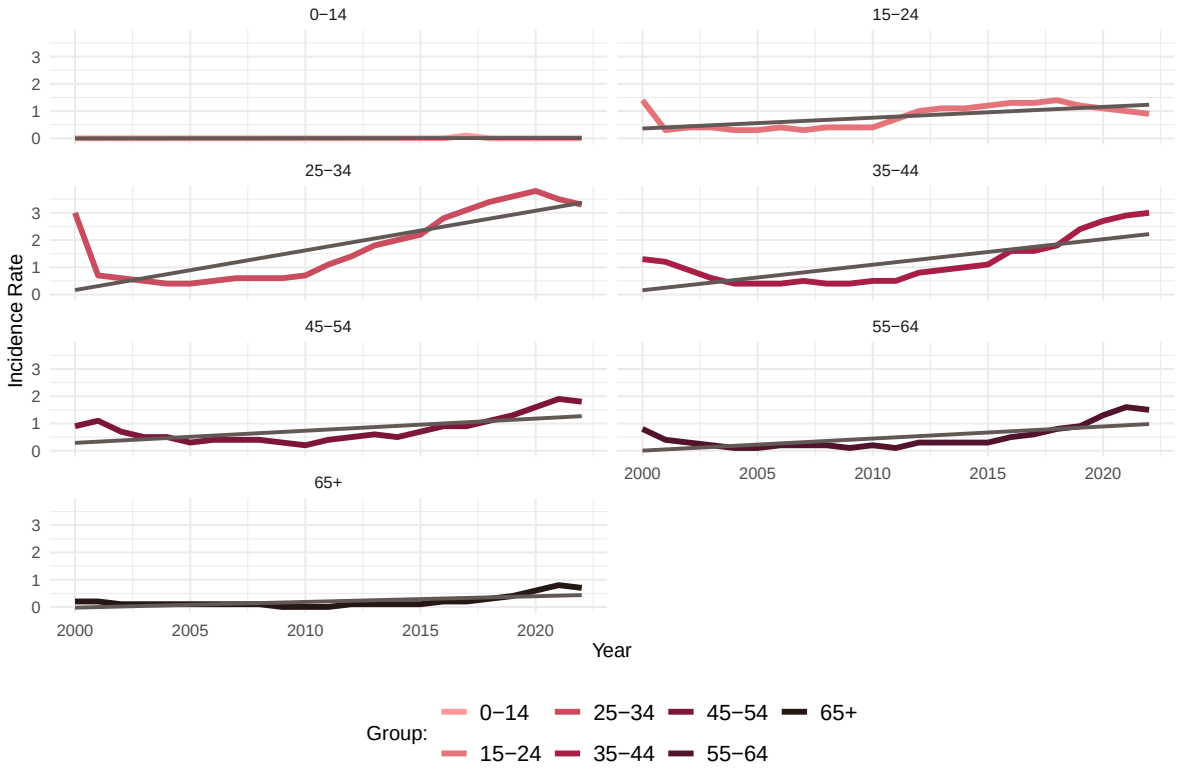


Table 2. Linear Regression Equations for Incidence Rates (per 100,000) by Age (2000-2022)

Age Group	Equation (y = b + m*x)
0-14	y = -1.19 + 0 * x
15-24	y = -78.89 + 0.04 * x
25-34	y = -290.94 + 0.15 * x
35-44	y = -187.2 + 0.09 * x
45-54	y = -88.84 + 0.04 * x
55-64	y = -88.73 + 0.04 * x
65+	y = -41.92 + 0.02 * x

Figure 4a. Lineplot of Yearly Incidence (per 100,000) by Sex (2000-2022)

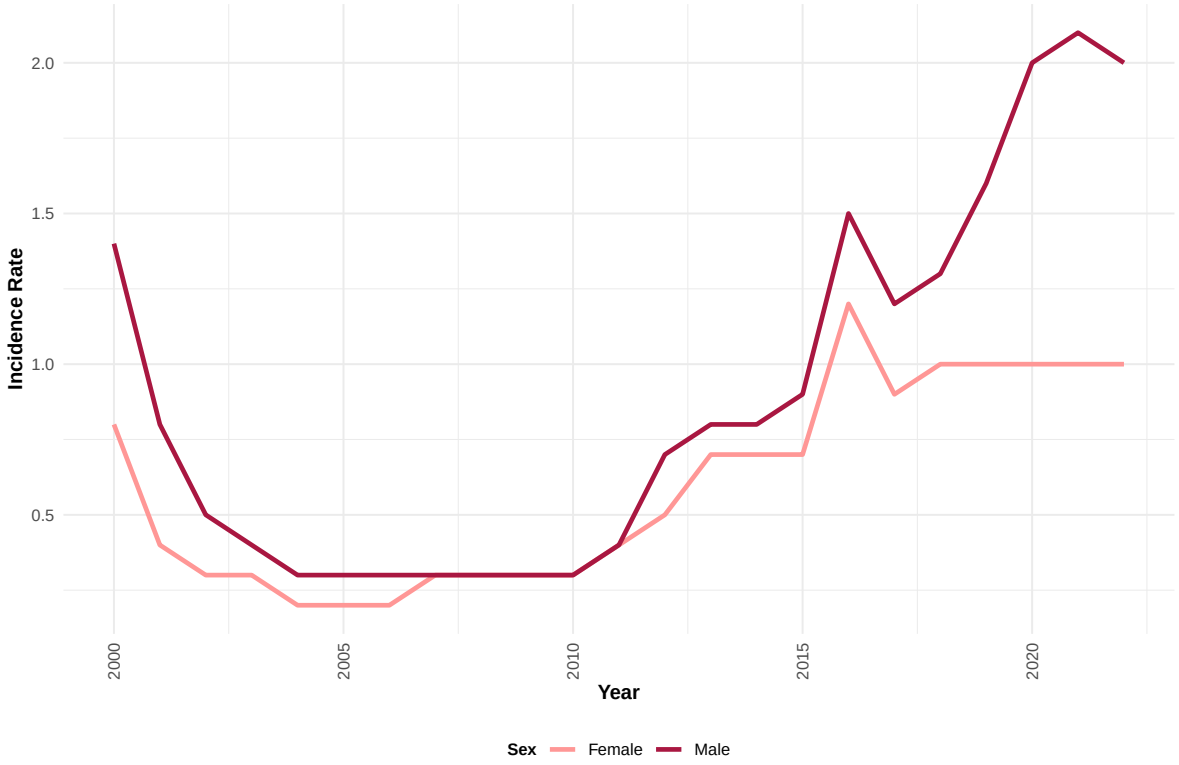


Figure 4b. Faceted Lineplot of Yearly Incidence (per 100,000) by Sex (2000-2022)

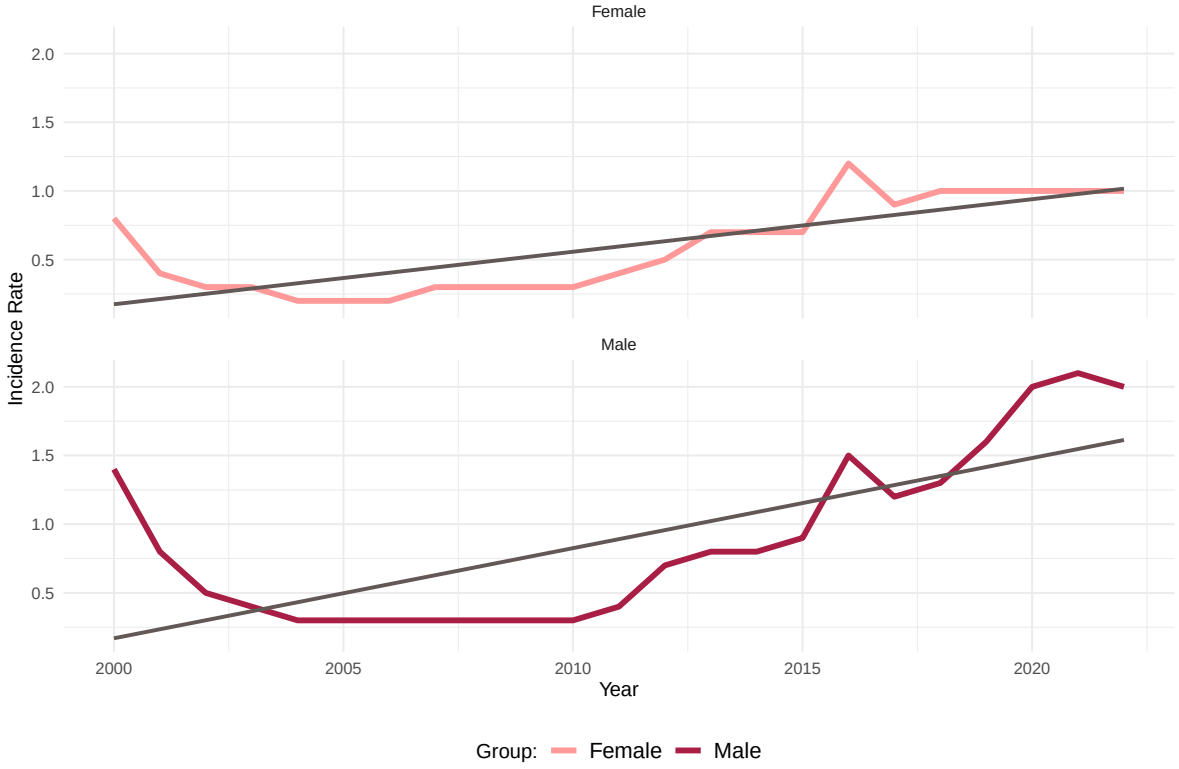


Table 3. Linear Regression Equations for Incidence Rates (per 100,000) by Sex (2000-2022)

Sex	Equation (y = b + m*x)
Female	y = -76.31 + 0.04 * x
Male	y = -131.06 + 0.07 * x

Figure 5a. Lineplot of Yearly Incidence (per 100,000) by Urbanicity (2018-2022)

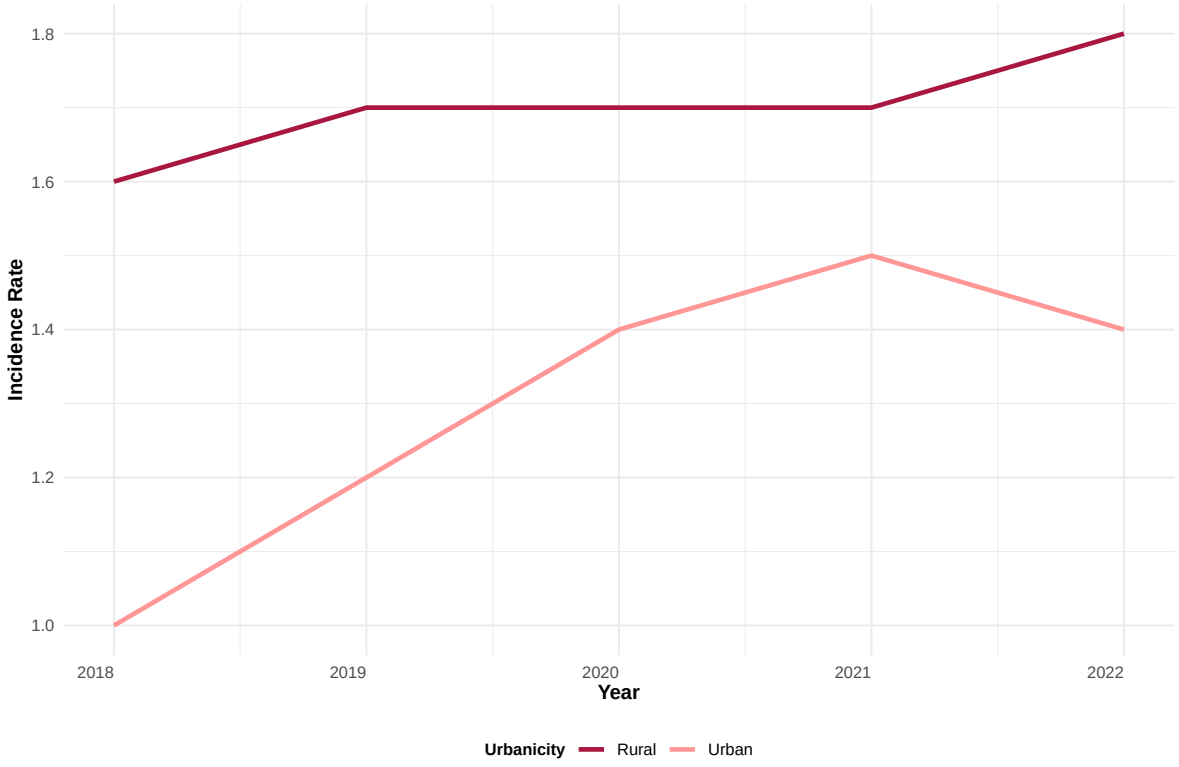


Figure 5b. Faceted Lineplot of Yearly Incidence (per 100,000) by Urbanicity (2018-2022)

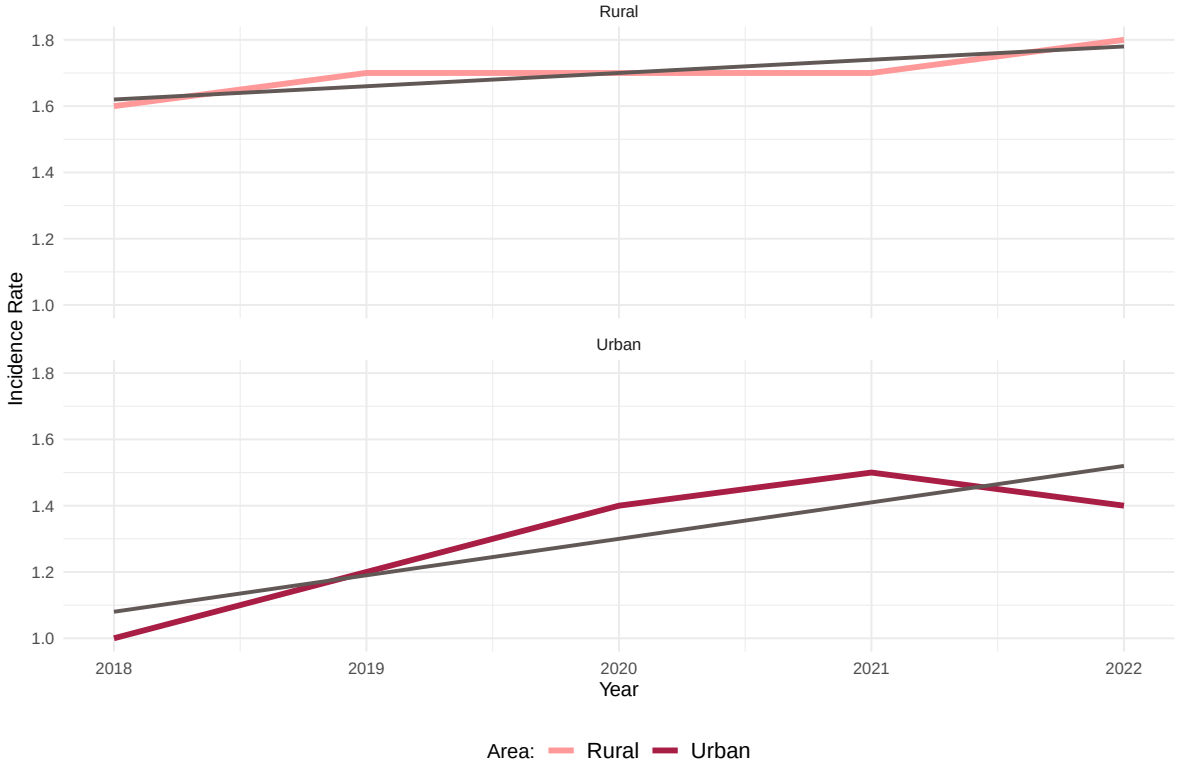


Table 4. Linear Regression Equations for Incidence Rates (per 100,000) by Urbanicity (2018-2022)

Area	Equation (y = b + m*x)
Rural	y = -79.1 + 0.04 * x
Urban	y = -220.9 + 0.11 * x

Conclusion and Summary

Conclusions were made by assessing the slopes of each line ("m" in "y = b + m*x). The group(s) with the greatest overall (positive) slope value (i.e., greatest increase in annual incidence per year) were interpreted as having the greatest rise in incidence for the period of interest.

In addition to consistently experiencing the highest yearly incidence rates for 2000-2022 (Figures 2a and 2b), **the American Indian/Alaska Native (AI/AN) population likewise**

experienced the greatest rise in incidence for the same period, with its linear model having a slope of 0.14 (Table 1).

The White group similarly had the second highest yearly incidence, as well as the second greatest rise with a slope of 0.07. Both the Black/African American and Hispanic/Latino groups experienced a rise of 0.03, although yearly incidence was consistently higher for the former group between 2000-2022. Finally, the Asian/Pacific Islander group experienced the lowest yearly incidence rates as well as the smallest rise over the period of interest (slope = 0.01).

Similarly, **the 25-34 year-old group experienced the greatest rise in incidence for the period of interest (slope = 0.15, Table 1)**, as well as the greatest yearly incidence rates for the same period (Figures 3a and 3b).

The second greatest rise in incidence was observed with the 35-44 year-old group (slope = 0.09) which likewise had the second greatest incidence rates over the period 2000-2022. The 15-24, 45-54, and 55-64 year-old groups all demonstrated similar rises incidence (slope = 0.04 for all). The 65+ year-old group experienced the second smallest rise, with a slope of 0.02. Finally, in addition to having the lowest yearly incidence rates over the period of interest, the 0-14 year-old group has seen almost no change in incidence, resulting in a slope of 0.0.

Finally, **males have seen the greatest rise in incidence of the two sexes included in this analysis, with a slope of 0.07.** Males similarly saw the greatest yearly incidence rates of the two groups. The female group, on the other hand, saw an overall rise of nearly half of what the male group experienced, with a slope of 0.04.

As alluded to in the introduction, injection drug use (IDU) is now identified as the primary risk factor for acute HCV infection. The association between this risk factor and HCV transmission has been particularly notable for the nation's AI/AN population. Although limited, ongoing literature partially attributes this disparity to the impact of generational trauma on engagement with high-risk behaviors, like IDU. Settler colonialism, forced assimilation and relocation policies, and ongoing social marginalization and cultural loss have all contributed to lesser health outcomes, like disproportionately higher HCV incidence rates, for this group. Further, historical trauma has led to decreased engagement with and distrust in healthcare, impacting this group's ability to access appropriate screening and treatment resources.

Literature regarding the increasing incidence in young individuals points to the shifting demographics of the nation's opioid epidemic. Notably, investigators point to the fact that injection drugs (e.g., heroin) have become more accessible than prescription drugs, leading to their higher abuse. Consequently, total incidence of bloodborne infections like HCV and HIV have seen major rises. Another consequence of this particular trend is the potential for future rises in perinatal and pediatric HCV cases as more women of child-bearing age may soon become infected.

Although incidence rates, as well as rises in incidence, have been greater for males than females, recent studies indicate that this trend may not hold within populations of injection drug users.

This may be the case as behavioral and social risk factors have been shown to vary between male and female injectors (e.g., many studies have demonstrated higher rates of needle and equipment sharing among women). However, more research is needed to fully examine and understand this discrepancy.

References

1. CDC. (2024a, May 30). Hepatitis C Basics. Hepatitis C. https://www.cdc.gov/hepatitis-c/about/index.html
2. CDC. (2024b, September 26). Infectious Diseases in Persons Who Inject Drugs. Persons Who Inject Drugs (PWID). https://www.cdc.gov/persons-who-inject-drugs/about/index.html
3. Jeong, D., Wong, S., Karim, M. E., Manges, A. R., Makuza, J. D., Bartlett, S. R., Velásquez García, H. A., Luster, D., Adu, P. A., Binka, M., Yu, A., Krajden, M., & Janjua, N. Z. (2024). Treatment of HCV with direct-acting antivirals on reducing mortality related to extrahepatic manifestations: a large population-based study in British Columbia, Canada. The Lancet Regional Health - Americas, 29, 100658. https://doi.org/10.1016/j.lana.2023.100658
4. Klevens, R. M., Liu, S., Roberts, H., Jiles, R. B., & Holmberg, S. D. (2014). Estimating Acute Viral Hepatitis Infections From Nationally Reported Cases. American Journal of Public Health, 104(3), 482–487. https://doi.org/10.2105/AJPH.2013.301601
5. Rose, M., Allen Myers, J., Ryan, N., Prince, A., Talbot, M., & Espinosa, C. M. (2019). Hepatitis C is now a Millennial Disease in Response to the Opioid Crisis: A Demographic Shift in Hepatitis C Infection. Open Forum Infectious Diseases, 6(Supplement_2), S159–S159. https://doi.org/10.1093/ofid/ofz360.368