The Impact of Vaccination Programs on Infectious Disease Rates in the United States

Katarzyna Kopania, Malwina Schonhofer, Juliusz Szymajda

November 5, 2024

1 Introduction

Vaccines have saved millions of lives, dramatically reducing the incidence of deadly infectious diseases such as smallpox and polio. However, despite clear scientific evidence supporting vaccination, public perception has become divided. This report aims to clarify the impact of vaccination programs through an analysis of data from the Tycho Project, focusing on disease rates before and after vaccine introductions.

2 Methodology

2.1 Data Preparation

The dataset, `us_contagious_disease.csv`, contains weekly disease counts from 1928 to 2011 across all 50 states. We focus on Measles, adjusting for population using rates per 100,000 people. Alaska and Hawaii are excluded due to their later statehood, which may impact the comparability of early data.

2.2 Analysis Tasks

1. **Time-Series Analysis:** Calculate and plot the yearly rate of Measles in California, marking the introduction year of the Measles vaccine.
2. **Variability Analysis:** Compare disease rate variability across states in 1950, 1960, and 1970 with and without square root transformation.
3. **Cross-State Trends:** Generate time-series plots for all states, using boxplots and line graphs to analyze trends. Overlay the national average for comparison.
4. **Hypothesis Testing:** Evaluate the effectiveness of vaccines across three diseases using statistical tests.

3 Results

3.1 Measles Rate in California

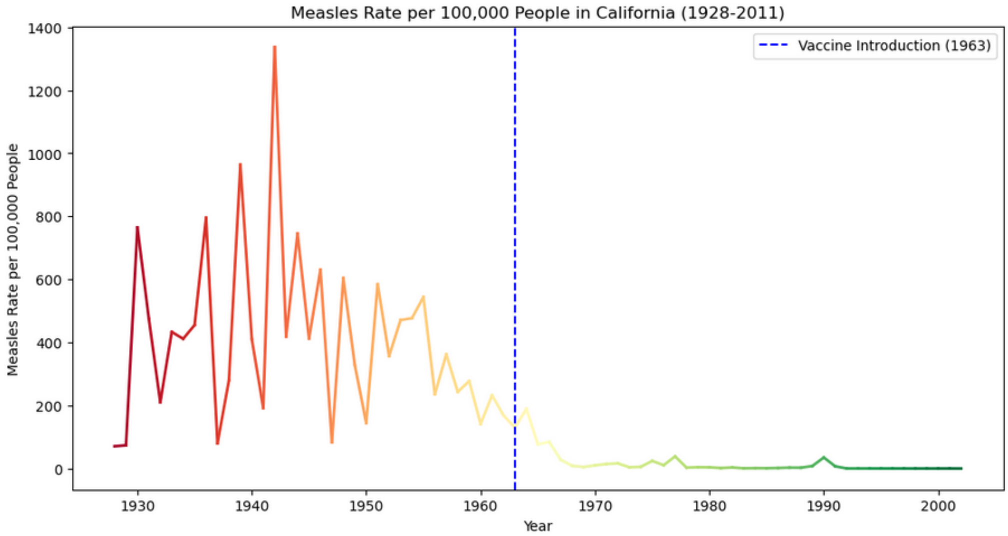


Figure 1: Yearly Measles rate per 100,000 in California with vaccine introduction marked.

Measles Rate in California: The plot showcases rate of measles cases per 100,000 in California (1928-2011) it highlights a significant decline in reported cases of measles after introduction of the measles vaccine in 1963. The peak of cases was reported in **1944**, where it reached value above **1200** per 100,000 people.

3.2 Measles Rate in California with square root transformation

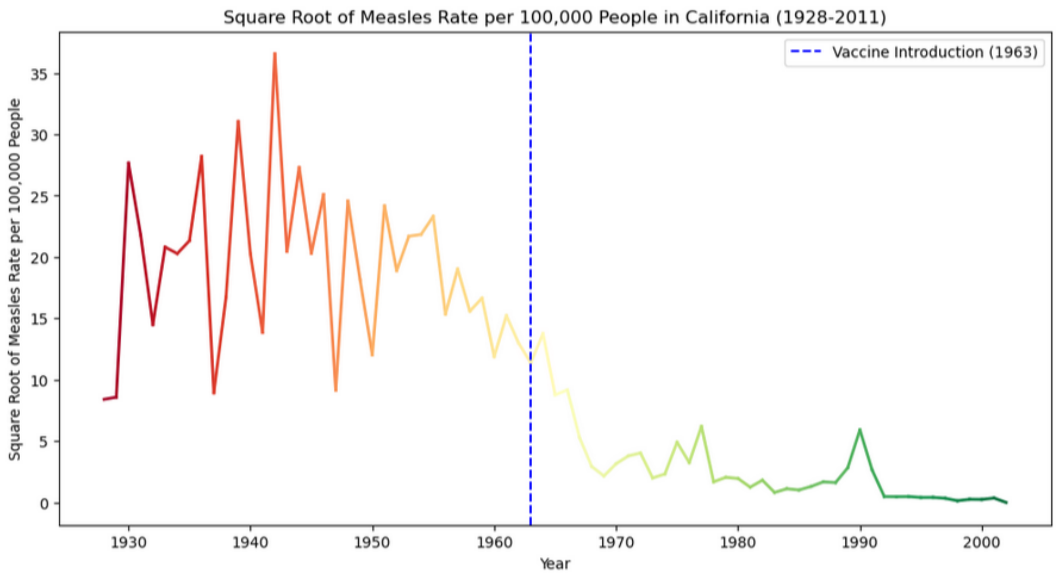
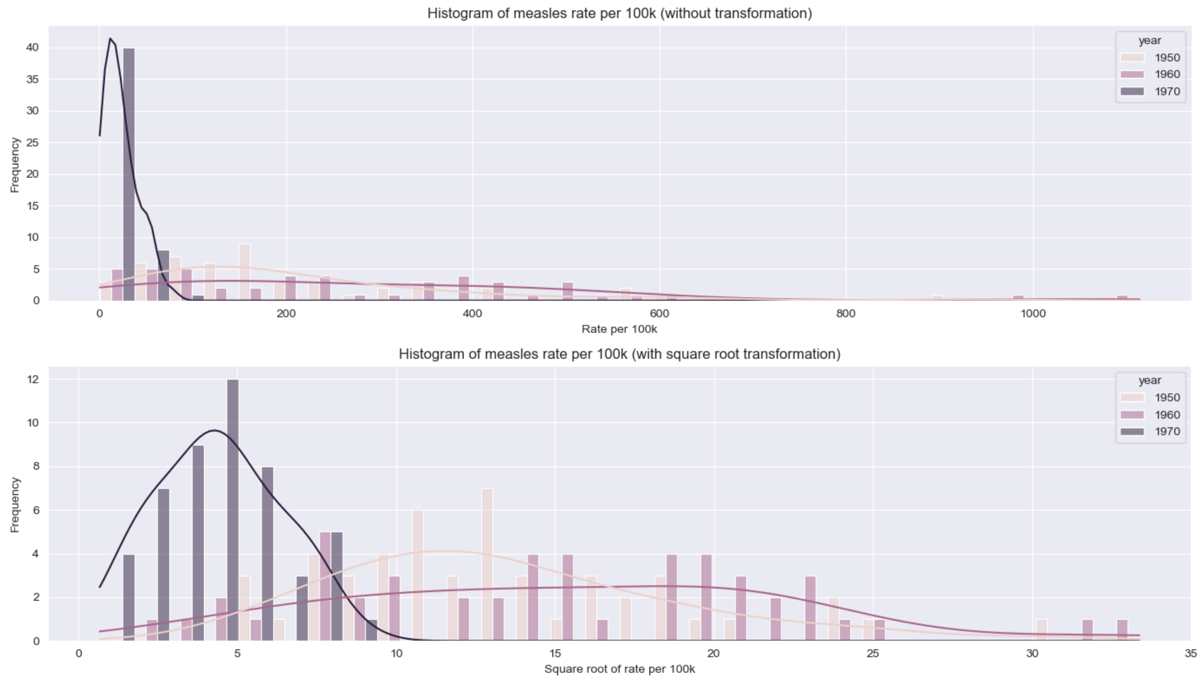


Figure 2: Yearly Measles rate per 100,000 in California with vaccine introduction marked and square root transformation.

Measles Rate in California (Square root transformation): The square root transformation was applied to enhance the visibility of trends or patterns and helps stabilize data. From the **1928-1963** the peaks of measles rates (sqrt) were reaching **35-37**.

3.3 Variability Across States



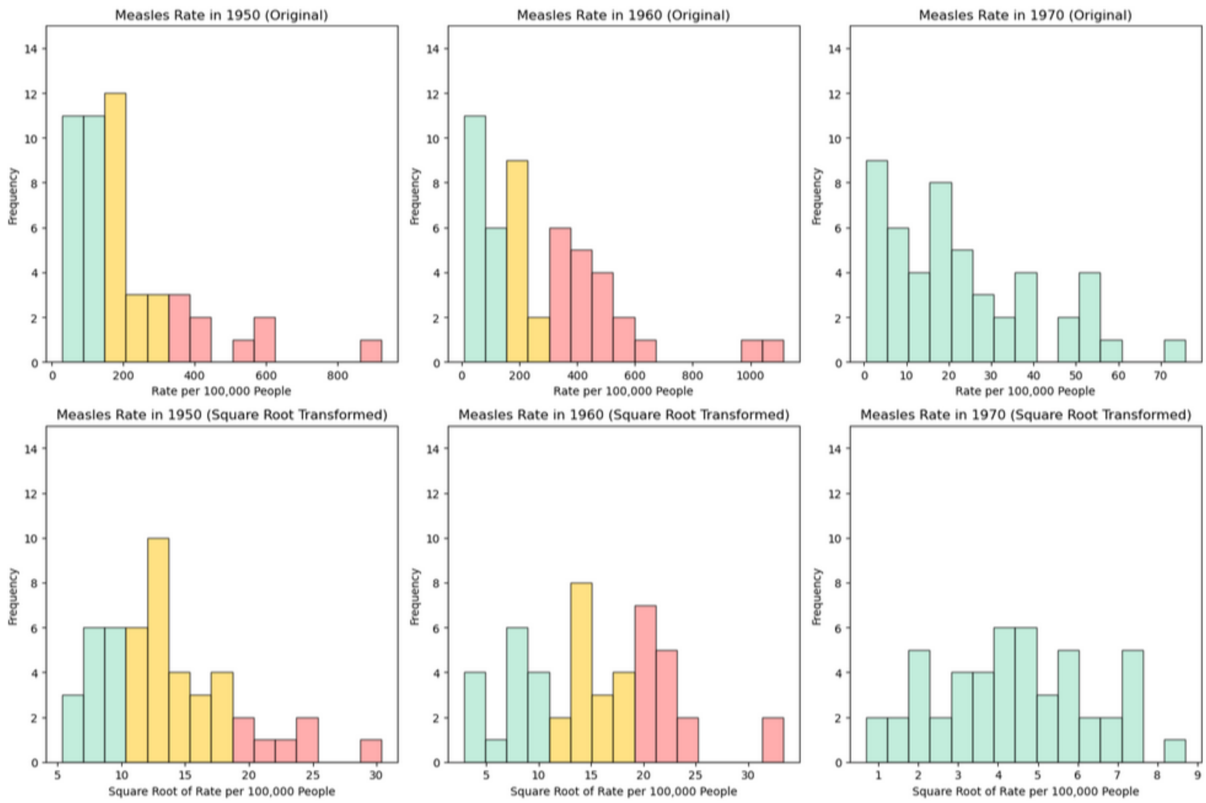


Figure 3: Enter Caption

Figure 4: Histograms of Measles rates across states in 1950, 1960 and 1970.

Top Row (Original Data):

1. **1950 -** rates per 100,000 people, with frequencies peaking around **200**
2. **1960 -** slightly different distribution with frequencies peaking around **200 - 400**
3. **1970 -** decline in rates, with scale reduced to **0 - 70**

Bottom Row (Square Root Transformed):

1. **1950 -** rates per 100,000 people with square root, ranging from **5 - 30**
2. **1960 -** ranging from **5 - 30**
3. **1970 -** after introduction of vaccine, the rates are ranging **1 - 9**

Key Observation:

- **The data shows a general decrease in measles rates over the 20-year period after implementation of vaccines**

3.4 Box Plot Analysis of Measles Rates Across Years

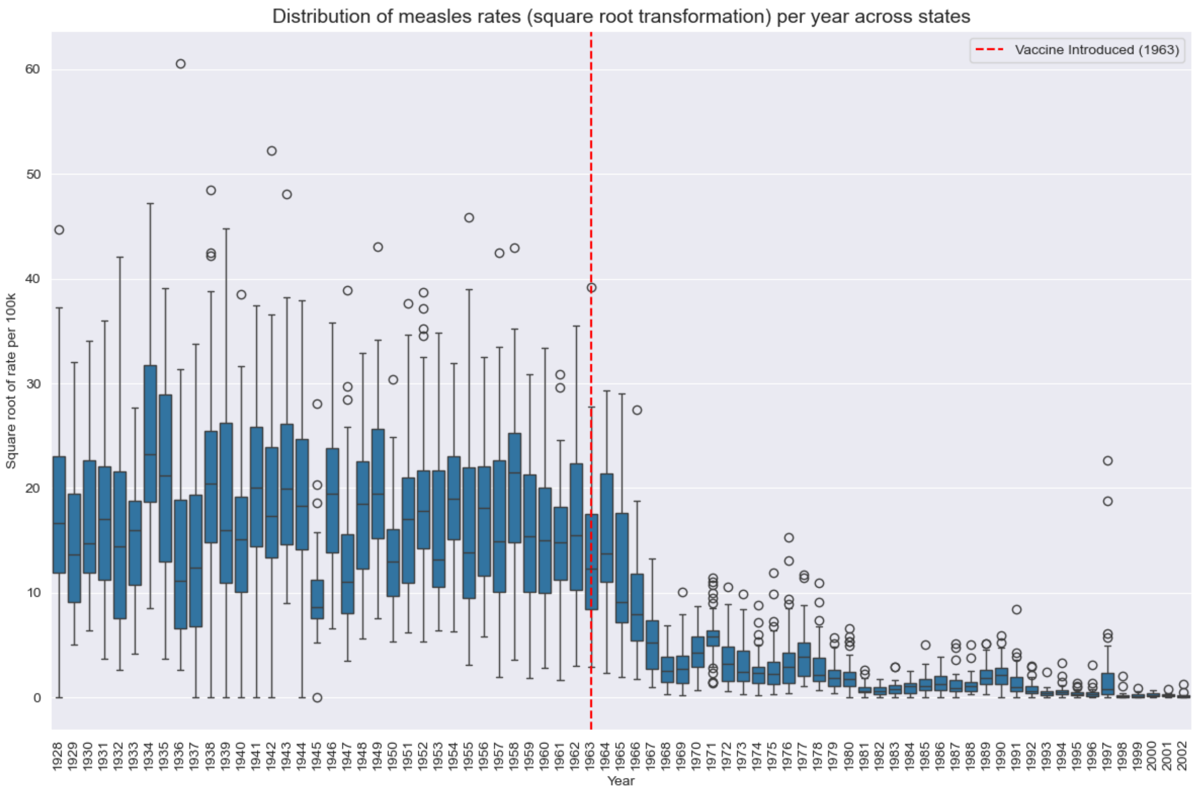


Figure 5: Distribution of measles rates (square root transformation) per year across states.

Measles Rates Distribution Across Years (All State Combined): The boxplot shows the distribution of measles rates (square root transformed) per 100,000 people across U.S. states from 1928 to 2002. Each boxplot represents the spread of measles rates for a specific year, with the median rate shown as a line inside each box, the interquartile range (IQR) as the box itself and whiskers extending to 1.5 times the IQR. The individual dots outside the whiskers represent outliers, which are unusually high or low rates for that year. The red dashed vertical line marks the year 1963 – the introduction of measles vaccine. **Before 1963, rates are high with significant variability and many outliers, indicating frequent outbreaks. After 1963, there is a marked decline in rates,** with smaller, more compact boxplots showing lower and more consistent rates across states, which proves that the vaccine was effective in controlling measles.

3.5 Cross-State Analysis

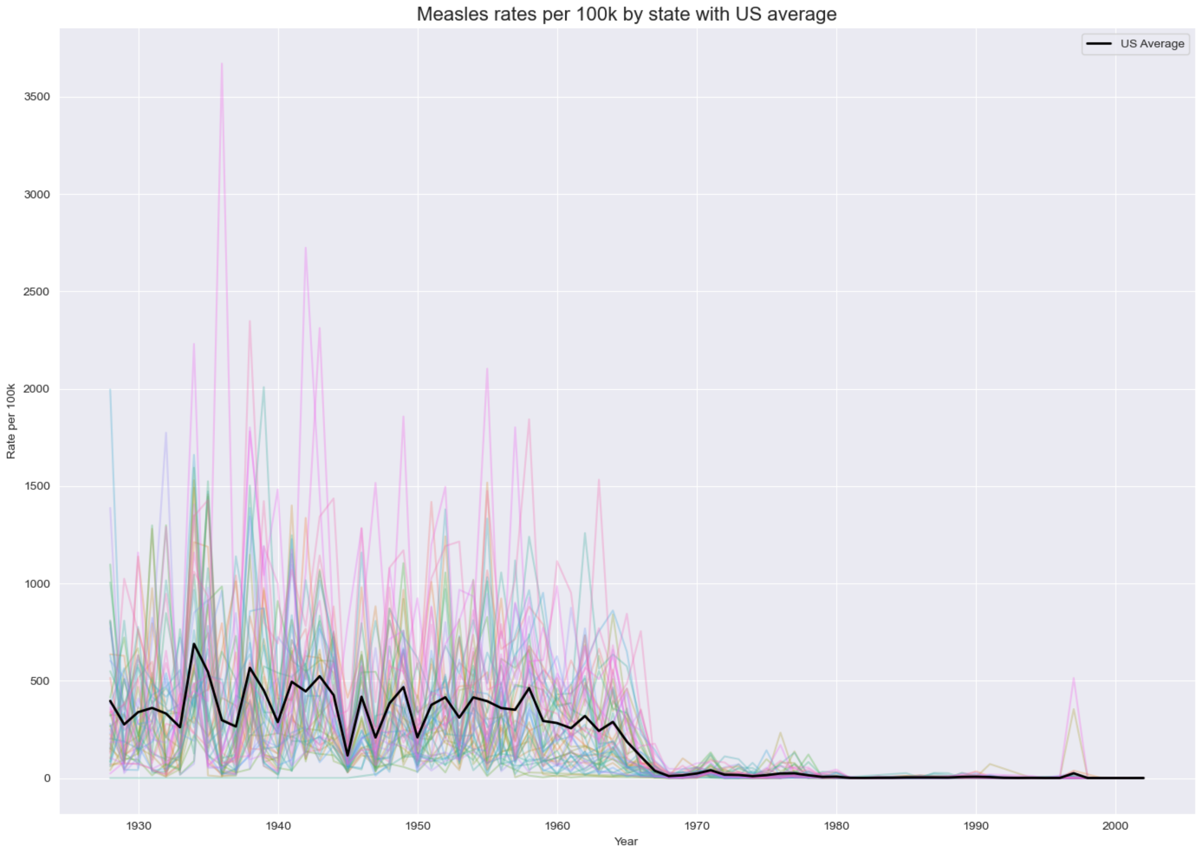


Figure 6: Trends of Measles rates over time in each state with national average overlay.

Cross-State Trend Analysis: This line plot displays the trends in measles rates per 100,000 people across individual U.S. states over time, with a black line representing the **national average**. Following the introduction of the measles vaccine in 1963, the rates decline sharply, with both state-level and national averages approaching zero, reflecting the vaccine's effectiveness in reducing measles incidence.

3.6 Measles Rates Across All States Analysis

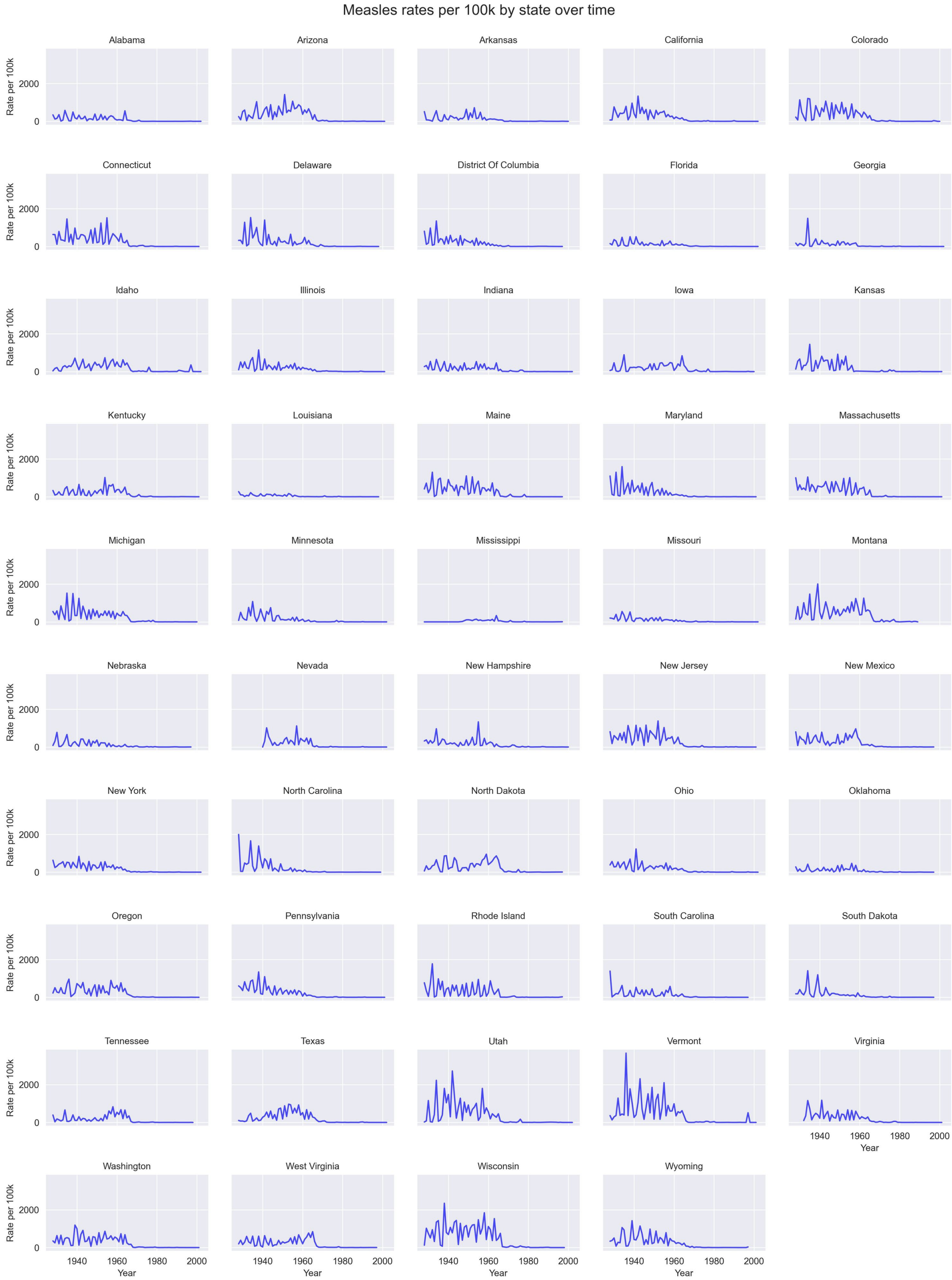


Figure 7: Trends of Measles rates over time in each state.

Measles Rates Across All States: In comparison to the plot of measles rates in California, the plots for other US states also reveal a tendency for **high peaks of reporting occurring roughly between the 1940s-1960s, followed by a rapid decline in rates after the introduction of the measles vaccine in 1963.**

3.7 Hypothesis Testing on Vaccine Effectiveness

```
Measles Vaccine Effectiveness Test:
Pre-vaccine mean rate per 100k: 353.68
Post-vaccine mean rate per 100k: 23.81
T-test statistic: 14.47, p-value: 0.00000
Significant reduction in disease rate post-vaccine.
```

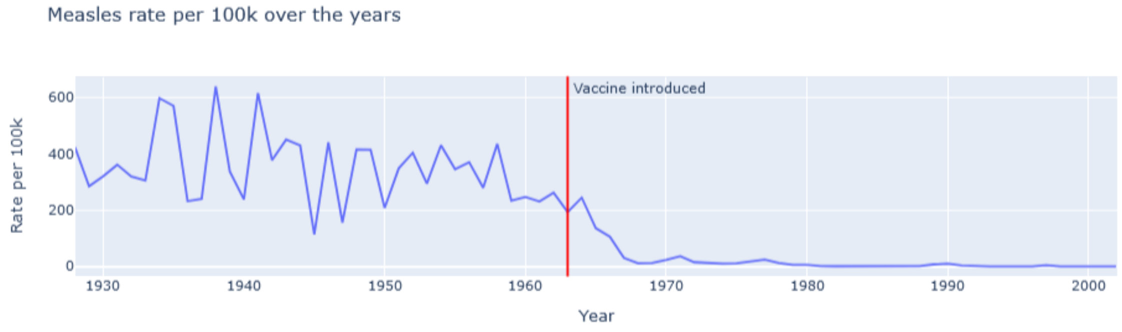


Figure 8: Impact of Measles vaccine on disease rates over time.

```
Polio Vaccine Effectiveness Test:
Pre-vaccine mean rate per 100k: 12.56
Post-vaccine mean rate per 100k: 3.40
T-test statistic: 3.64, p-value: 0.00080
Significant reduction in disease rate post-vaccine.
```

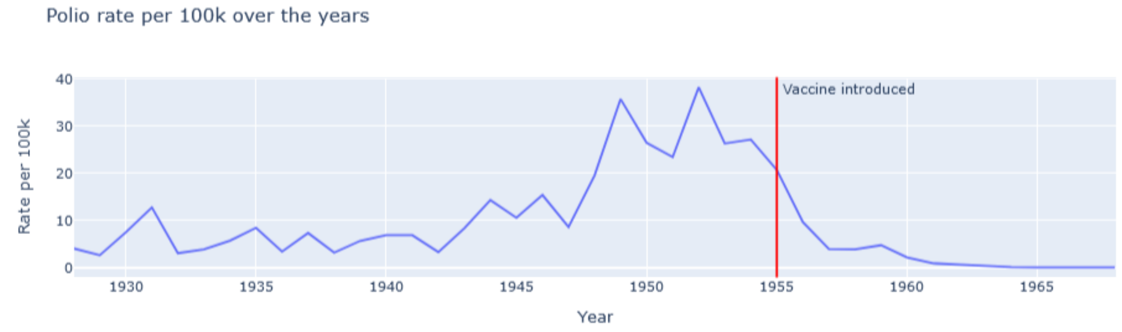


Figure 9: Impact of Polio vaccine on disease rates over time.

```
Hepatitis A Vaccine Effectiveness Test:
Pre-vaccine mean rate per 100k: 15.40
Post-vaccine mean rate per 100k: 4.20
T-test statistic: 7.32, p-value: 0.00000
Significant reduction in disease rate post-vaccine.
```

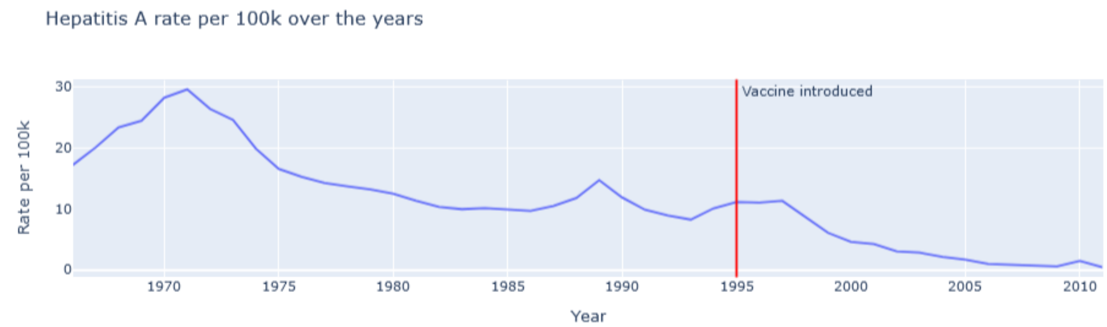


Figure 10: Impact of Hepatitis vaccine on disease rates over time.

Hypothesis Testing on Vaccine Effectiveness: The analysis evaluates the effectiveness of vaccines against **Hepatitis A, Polio and Measles** by comparing disease incidence rates before and after vaccine introduction. The pre-vaccine mean is the average rate of disease per 100,000 population before the vaccine was introduced, while the post-vaccine mean reflects the average rate after vaccination. To assess whether the differences in these rates are meaningful, the t-test statistic is used in the analysis. Moreover, the p-value indicates whether the observed differences are statistically significant. It helps to determine if the results are likely due to the vaccine rather than random chance. **A p-value less than 0.05 suggests a significant reduction in disease rates post-vaccine, implying effective vaccination, whereas a p-value greater than 0.05 indicates no significant change**. Overall, these findings highlight the important role of vaccination in controlling these diseases and demonstrate its effectiveness in overcoming them.

4 Conclusion

Our findings reinforce the public health benefits of vaccination programs. Visual and statistical analyses clearly indicate a decrease in disease rates following vaccine introductions. Despite historical controversies and misinformation, the scientific evidence overwhelmingly supports the positive impact of vaccines. This report highlights the importance of vaccination programs in controlling infectious diseases. By demonstrating disease rate reductions with visual and statistical methods, we provide a data-driven basis for vaccine advocacy, crucial in addressing vaccine hesitancy.