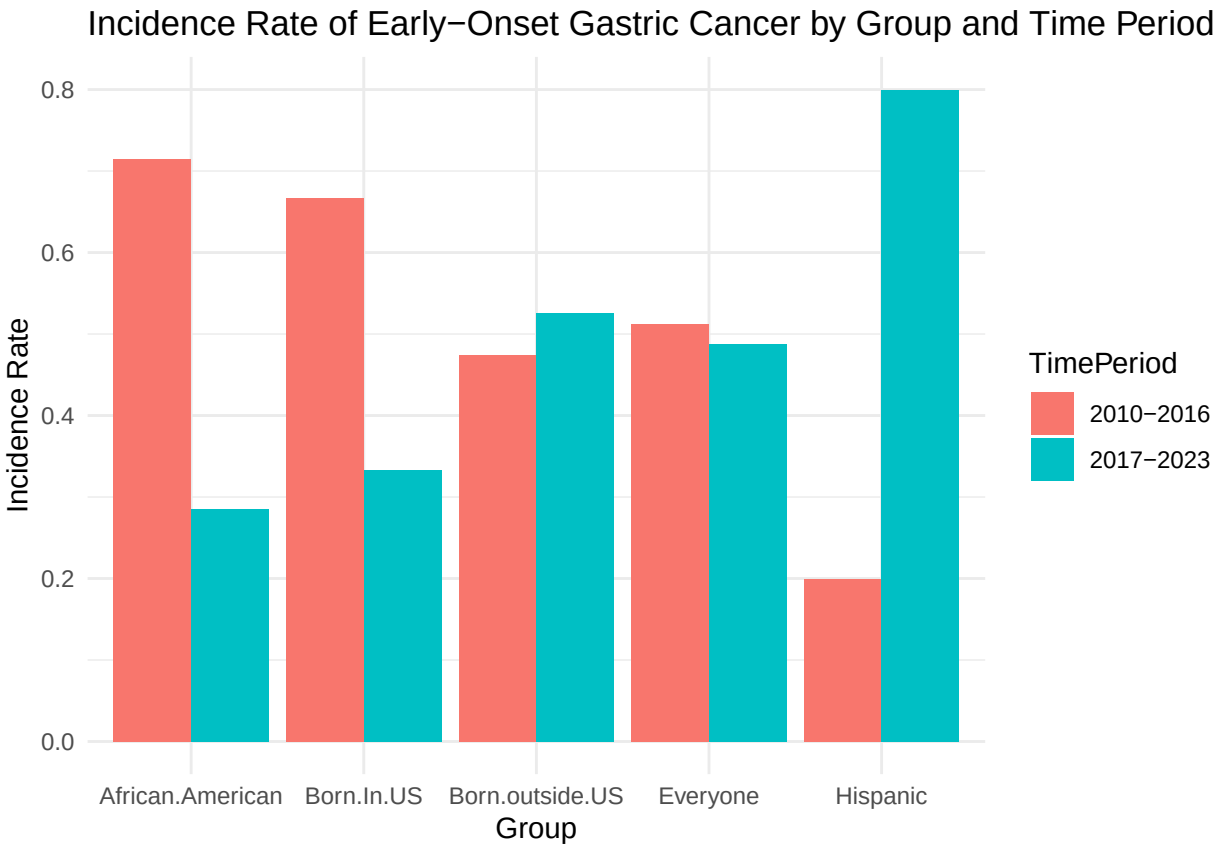
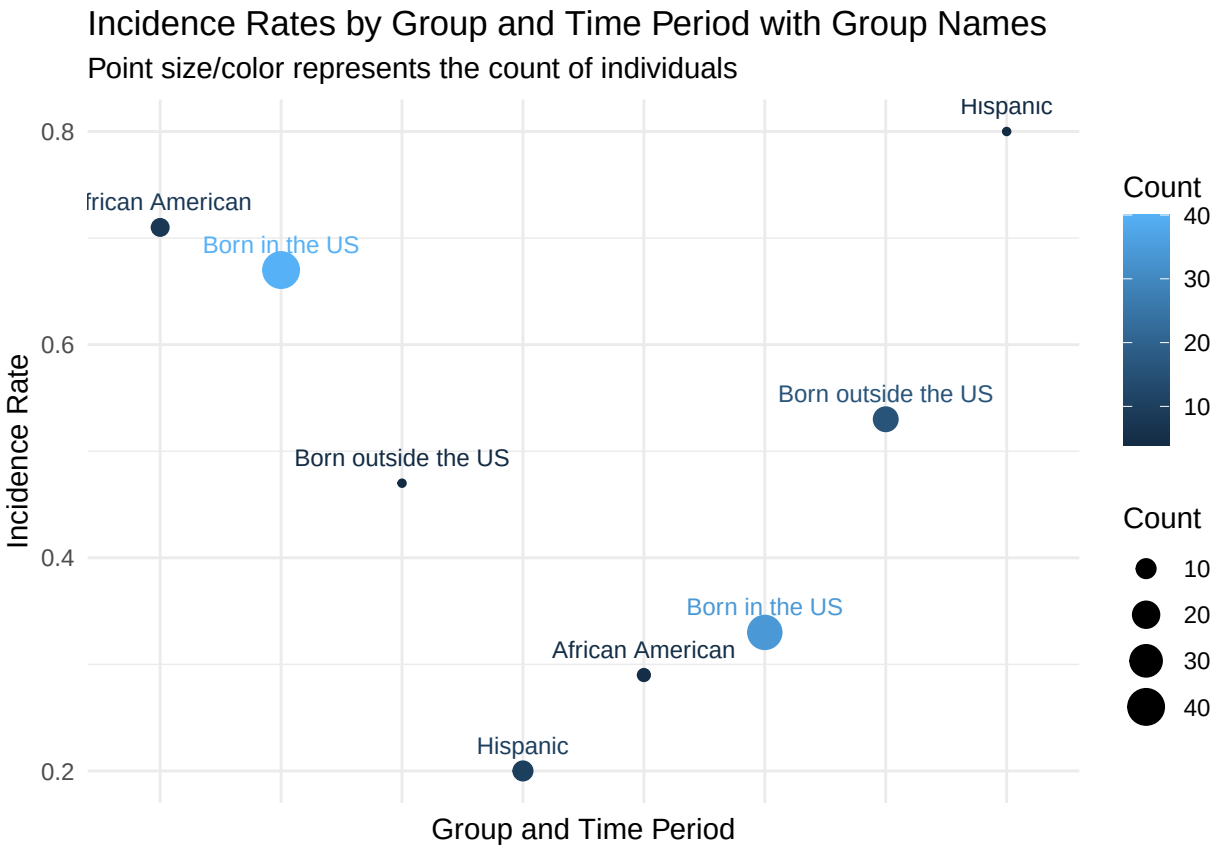
Preliminary Early Onset Gastric Cancer Analysis

Madeline Dabney

2024-12-04



The bar graph displays the distribution of incidence rates for early-onset gastric cancer across different groups (e.g., Hispanic, African American, Born in the US, etc.) and time periods (2010-2016 and 2017-2023). The x-axis represents the incidence rates, while the y-axis shows the frequency (the number of cases within each incidence rate range). The color coding differentiates between the groups, and the facet grids separate the data by time period. From this graph, you can compare how the incidence rate of early-onset gastric cancer varies across different groups and time periods, looking for any notable differences or trends.



Each point represents a group-time combination (e.g., "Hispanic 2010-2016"). The size and color of each point reflect the number of individuals in that group-time period, so larger and more intense points show higher counts. The y-axis shows the incidence rate, helping to visualize how it varies across different groups and time periods.

Descriptive Statistics

```
##
##                       2010-2016 2017-2023
##   African American           10         4
##   Born in the US              8         5
##   Born outside the US        40        34
##   Hispanic                    4        16
```

```
##               Group TimePeriod      Mean    Median SD       Min       Max IQR
## 1          Hispanic  2010-2016 0.2000000 0.2000000 NA 0.2000000 0.2000000   0
## 2          Hispanic  2017-2023 0.8000000 0.8000000 NA 0.8000000 0.8000000   0
## 3  African.American  2010-2016 0.7142857 0.7142857 NA 0.7142857 0.7142857   0
## 4  African.American  2017-2023 0.2857143 0.2857143 NA 0.2857143 0.2857143   0
## 5        Born.In.US  2010-2016 0.6666667 0.6666667 NA 0.6666667 0.6666667   0
## 6        Born.In.US  2017-2023 0.3333333 0.3333333 NA 0.3333333 0.3333333   0
## 7   Born.outside.US  2010-2016 0.4742268 0.4742268 NA 0.4742268 0.4742268   0
## 8   Born.outside.US  2017-2023 0.5257732 0.5257732 NA 0.5257732 0.5257732   0
## 9          Everyone  2010-2016 0.5123967 0.5123967 NA 0.5123967 0.5123967   0
## 10         Everyone  2017-2023 0.4876033 0.4876033 NA 0.4876033 0.4876033   0
```

Hispanic: In the time period from 2010-2016, the incidence rate is 0.20, and in 2017-2023, it increases to 0.80. This suggests a significant rise in the incidence rate among the Hispanic group over the two time

periods. African American: The incidence rate for 2010-2016 is 0.71, which decreases to 0.29 in 2017-2023. This indicates a decline in the incidence rate for African Americans in the second time period. Born in the US: For individuals Born in the US, the incidence rate is 0.67 in 2010-2016 and drops to 0.33 in 2017-2023, showing a decrease in incidence over time. Born outside the US: For individuals Born outside the US, the incidence rate is 0.47 in 2010-2016, and slightly increases to 0.53 in 2017-2023, showing a slight upward trend in the incidence rate.

Mean and Median: The mean and median values are identical for all groups and time periods because oof the small sample size. Standard Deviation: The standard deviation is not available (NA) due to the lack of variation in each group. There is only one data point in each group, so the variance (and standard deviation) cannot be computed. Min and Max: The minimum and maximum values are the same for each group and time period since there's only one data point, and thus no range of values. Interquartile Range (IQR): The IQR is 0 for all groups, as the data does not have enough variability to produce an interquartile range.

The table shows that the Hispanic group experienced a substantial increase in incidence from 2010-2016 to 2017-2023, while the African American and Born in the US groups saw a decline. The Born outside the US group shows a slight increase in incidence, but the change is minimal. The small sample size (only one data point per group and time period) prevents more detailed analysis, such as calculating variability (SD) or conducting more sophisticated statistical tests

Thus, while we can describe trends between groups and time periods, we cannot assess the variability or statistical significance due to the size limitations of the dataset. Further data collection would be necessary to gain valuable deeper insights into these trends.