Data Analysis

Yves Greatti

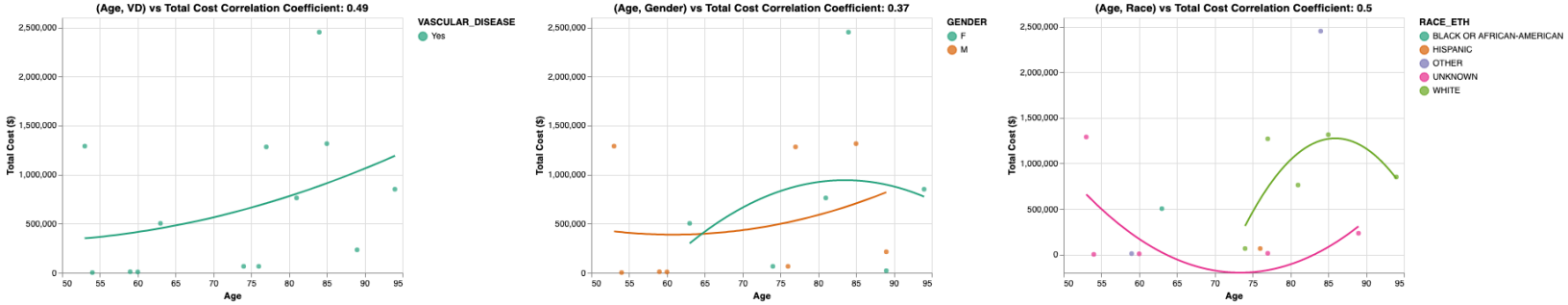
Cohort Age Statistics

	Count	Mean $\pm$ SD	Min	25%	50%	75%	Max
All	80	59 ± 19	13	52	60	74	94
Female	51	58 ± 20	13	49	62	73	94
Male	29	60 ± 18	15	53	59	76	89
Hispanic	14	47 ± 24	13	23	51	60	87
Black African-American	7	51 ± 22	18	38	60	62	80
Asian	1	15	15	15	15	15	15
White	30	65 ± 16	35	53	69	78	94
Unknown	26	61 ± 14	24	54	60	71	89
Other	2	71 ± 17	59	65	71	77	84

Total Cost: increased with age and disease

Left: Disease and age trend. Middle: diseased people and gender trend. Right: diseased people and race trend

Left: diabetic people and age trend. Middle: diabetic people and gender trend. Right: diabetic people and race trend



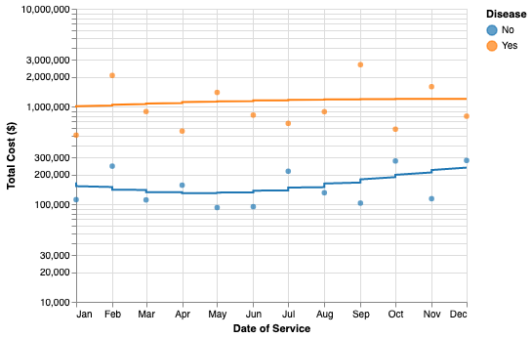
Left: people w. vascular disease (VD) and age trend. Middle: VD people and gender trend. Right: VD people and race trend

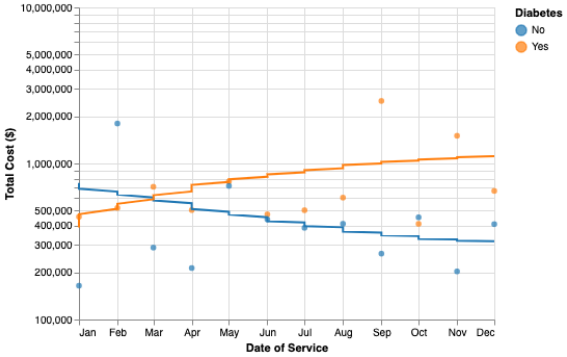
Investigation into the correlation between patient demographics (age, gender, and race/ethnicity) and disease status with total healthcare costs

- Patients without diseases exhibit an upward trend in total costs until age 50, after which their costs plateau (previous page top row charts).
- The total cost for patients with diseases tends to rise until later in life (previous page top row charts).
- As individuals age, the total costs for both men and women tend to increase. However, there is a slightly steeper increase in costs for women as they age, which may be attributed to their longer life expectancy (previous page top row charts).
- The presence of a significant number of patients with 'unknown' or 'other' ethnicities hinders the ability to conduct meaningful trend analysis.
- The trends in total cost are based on the available data points, which indicate that there are sufficient numbers of White and Hispanic patients for a meaningful analysis, whereas the sample sizes for other ethnic groups are relatively smaller.
- After reaching middle age, the total costs for White patients begin to decline before increasing exponentially later in life (previous page top row charts).
- In contrast, the total costs for individuals in the Hispanic population tend to increase gradually as they age (previous page top row charts).
- Due to their larger representation in the dataset, individuals with diabetes exhibit trends similar to those observed for the overall disease population (previous page middle row charts).
- When examining vascular disease patients by age and gender, distinct trends emerge. For women, costs tend to increase in an inverse parabolic pattern before decreasing again in later years. In contrast, men exhibit a steady increase in costs over time. Meanwhile, White and Unknown patients demonstrate opposite parabolic trends, with costs for White individuals increasing in their late 70s before decreasing, while costs for Unknown patients steadily decrease until their 70s and then begin to rise slowly. It's important to note that these results are limited by the relatively small number of data points available, which may affect the interpretation of these trends (previous page bottom row charts).

EXTRA

Total Cost Trend Over 2020





The total costs for people with diseases are generally higher compared to those without diseases. This trend is also observed in people with diabetes, who have consistently higher costs throughout the year compared to non-diabetic people. However, at the beginning of the year, non-diabetic people had higher total costs compared to diabetic people. Over the course of the year, costs for diabetic people slowly increased to reach a total of $1,000,000 at the end of the year. On the other hand, non-diabetic people showed the opposite trend, with initial costs around $700,000 that gradually decreased to $300,000 by year-end.

	AGE	GENDER	RACE_ETH	DISEASE	MONTH	ROLLING_COST
0	87	F	HISPANIC	Yes	1	27200.0
1	87	F	HISPANIC	Yes	1	26200.0
2	87	F	HISPANIC	Yes	1	3500.0
3	63	F	BLACK OR AFRICAN-AMERICAN	Yes	1	3000.0
4	77	M	WHITE	Yes	1	3000.0
5	87	F	HISPANIC	Yes	1	2900.0
6	63	F	BLACK OR AFRICAN-AMERICAN	Yes	1	2900.0
7	71	F	UNKNOWN	Yes	1	2000.0
8	71	F	UNKNOWN	Yes	1	1900.0
9	53	M	WHITE	Yes	1	1600.0
10	77	M	WHITE	Yes	1	1400.0
11	77	M	WHITE	Yes	1	1300.0
12	86	M	WHITE	Yes	1	1200.0
13	53	M	WHITE	Yes	1	1100.0
14	63	F	BLACK OR AFRICAN-AMERICAN	Yes	1	1000.0
15	45	M	UNKNOWN	No	1	900.0
16	52	F	BLACK OR AFRICAN-AMERICAN	No	1	900.0
17	89	M	UNKNOWN	Yes	1	800.0
18	49	F	UNKNOWN	No	1	800.0
19	53	M	WHITE	Yes	1	700.0

As people with disease age, their monthly rolling costs increase, with the highest costs observed among older Hispanic and Black/African American women, followed by White and Black/African American men. The higher costs for women are partially explained by their longer life expectancy compared to men.

SQL

```
SELECT AGE, GENDER, RACE_ETH, DISEASE,
MONTH(DATE_OF_SERVICE) AS MONTH, SUM(TCC_PAID)
OVER ( PARTITION BY AGE, GENDER, RACE_ETH, DISEASE
        ORDER BY DATE_OF_SERVICE ROWS
       BETWEEN 1 PRECEDING AND CURRENT ROW )
'      AS ROLLING_COST FROM CONS_TABLE
GROUP BY AGE, GENDER, RACE_ETH, DISEASE,
MONTH(DATE_OF_SERVICE), DATE_OF_SERVICE, TCC_PAID
 ORDER BY MONTH, ROLLING_COST DESC
```