**Does health insurance status influence the hospital length of stay for patients with asthma conditions?**

**Naixin Zhang**

**Motivation**

Despite being among the top world economic powers, the US remains the sole industrialized nation in the world without universal health care coverage[[1]](#footnote-1). The issue of providing health care service bothers the United States all the time. Health insurance is one of the important indices to health care servers providing. The data shows that private health insurance coverage continued to be more prevalent than public coverage, covering 67.3 percent of the population and 34.4 percent of the population, respectively in 2018[[2]](#footnote-2). Past research has demonstrated that different health insurance types lead to different medical resource use. For some chronic disease, like asthma condition, if good quality of health care service can't be provided, it will lead to an increasing use of outpatient departments which means reduce medical efficiency.

Furthermore, Length of stay in hospital(LOS) is a typical index to measure the health care quality due to its availability and objective nature[[3]](#footnote-3). The national average length of stay in hospital is 4.5 days, according to the Agency for Healthcare Research and Quality, at an average cost of $10,400 per day. LOS has been widely linked to health outcomes, such as readmission rates. Looking specifically at LOS and its relationship to insurance status in asthma patients, the research has been discordant. This study aims to prove that the different health insurance types has an influence on LOS in asthma conditions patients. Additional research is needed to examine causes and outcomes from these differences to increase efficiency in the health care system, decrease costs and shrink disparities in health outcomes.

**Data Description**

The Data used for this research has been obtained from the website of Health Data (https://healthdata.gov). This data contains over 7 million basic record level details regarding the discharge of inpatients in the State of New York in different health service areas from 2014 to 2016,including personal information such as age, gender, race, facility ID, diagnosis, length of stay, payment typology and so on . We used this dataset to compare differences in LOS with asthma conditions among different insurance groups.

Since there are many features in the dataset, to build good intuition about our dataset, in this section, we will investigate two aspects: data preprocessing and data visualization.

In data preprocessing, four steps were performed: (1) Extract the data we need (2) Check for missing values (3) Check each column's value and process it (4) Export the clean data. The details can be check in the appendix 1. Table 1 and 2 provides a statistic summary of the features' we interested in the dataset after processing.

Asthma patients, aged 8 to 70 years, were included in the analysis. The analysis excludes patients lacking information for birth weight. Patients with LOS in the highest 0.4 percent of the sample (LOS >15 days) were considered outliers and were excluded to prevent their extended LOS from artificially inflating the mean LOS of our study population.

Patients were categorized into one of four insurance groups for analysis based on primary payer status: 1) private insurance 2) Blue Cross/Blue Shield 3) Medicaid 4) Medicare. The main outcome measure for the study was hospital LOS.

**Table1 statistic summary of patients' information for category variable**

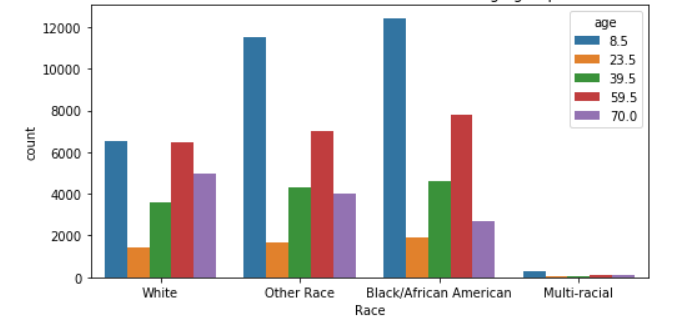
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable name | | | Count of patients | Percentage |
| Demographic | gender | Female | 45859 | 59.5% |
| Male | 31208 | 40.5% |
| Race | Black | 28336 | 36.8% |
| White | 22245 | 28.9% |
| Multi-Race | 537 | 0.7% |
| Others | 25949 | 33.7% |
| Ethnicity | Spanish/Hispanic | 20770 | 26.9% |
| Not Span/Hispanic | 55854 | 72.5% |
| Multi-ethnic | 443 | 0.6% |
| Medical conditions | Admission Type | Urgent | 1993 | 2.6% |
| Emergency | 74454 | 96.6% |
| Elective | 620 | 0.8% |
| Severity code | 1 | 32864 | 42.6% |
| 2 | 33882 | 44.0% |
| 3 | 12602 | 16.4% |
| 4 | 1749 | 2.3% |
| Risk of Mortality | Minor | 55695 | 72.3% |
| Moderate | 13223 | 17.2% |
| Major | 7225 | 9.4% |
| Extreme | 924 | 1.2% |
| treatment method | Medical | 76801 | 99.6% |
| Surgical | 266 | 0.3% |
| Discharge year | year | 2014 | 30798 | 40.0% |
| 2015 | 26821 | 34.8% |
| 2016 | 19448 | 25.2% |
| insurance type | / | Private | 9248 | 12.0% |
| Medicaid | 40917 | 53.1% |
| Medicare | 19326 | 25.1% |
| Blue cross/shield | 7576 | 9.8% |

**Table2 Statistic summary of patients' information for continuous variable**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| variable | mean | std | min | max |
| Length of stay in hospital | 3.12 | 2.57 | 1 | 15 |
| Age | 36.59 | 24.69 | 8 | 70 |
| Total charges | 22771 | 21403 | 1 | 602269 |

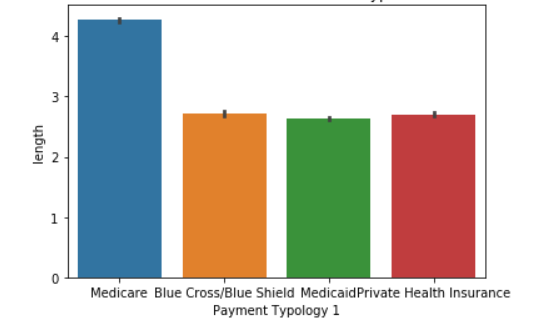
In data visualization part, we mainly focus on checking if there are co-founders might affect LOS.

Figure1 indicates there is a balance count of patients distribution for each race group. Teenager accounts for a large proportion while 70 and older patients account for a small proportion.



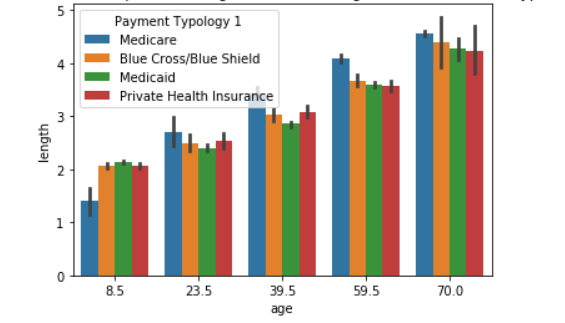
**Figure1 The count distribution of race based on age group**

Overall, patients with Medicare health insurance tends to stay in the hospital more than one day longer than other insurance groups. Patients with blue cross/blue shield, Medicaid or health insurance have similar length of stay in hospital. It shows in Figure 2.



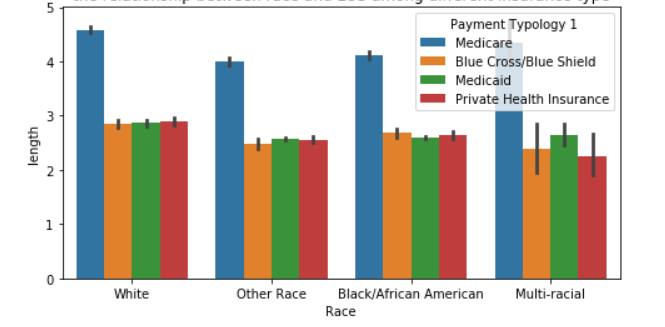
**Figure2 The LOS for each insurance type**

Figure 3 indicates the elderly patients tend to stay longer in hospital and young patients tend to stay shorter in hospital. However, among younger patients' group, Medicare patients tend to stay in the hospital shorter than other health care insurance patients which is different compared with other age groups. It shows that age does influence the length of stay in the hospital.

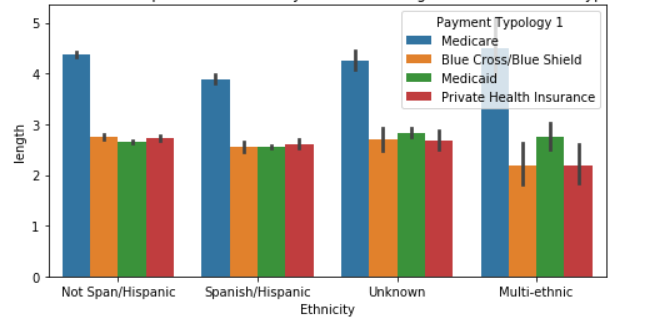


**Figure3 The relationship between age and LOS among different insurance type**

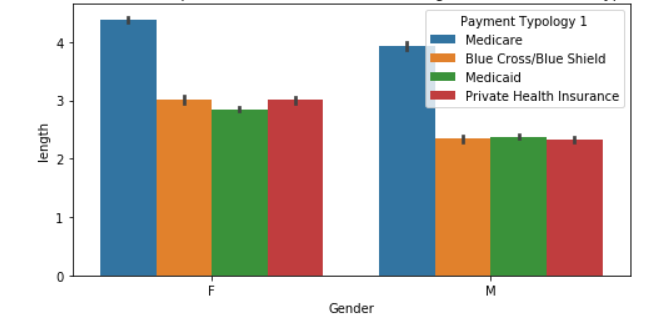
Figure 4 indicates that white patients have longer LOS compared with other race group. Among each group, Medicare patients tends to stay in hospital longer and other three insurance groups.



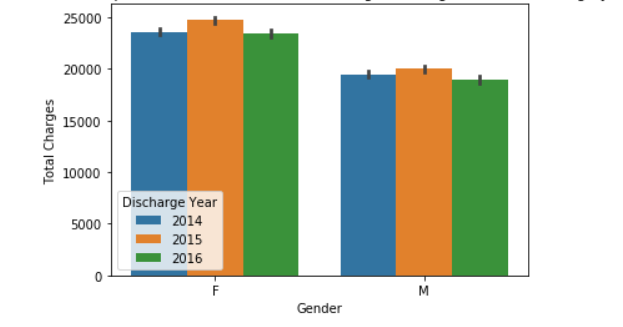
**Figure4 The relationship between race and LOS among different insurance type**



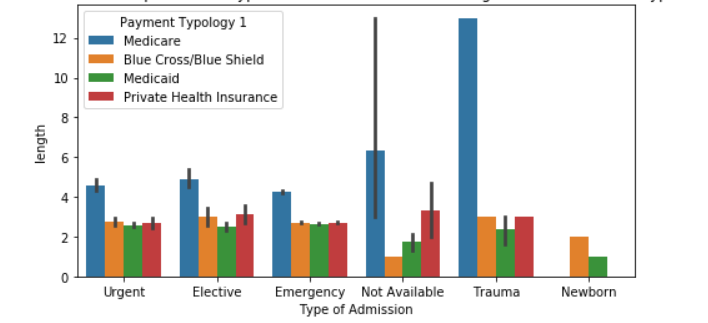
**Figure5 The relationship between Ethnicity and LOS among different insurance type**



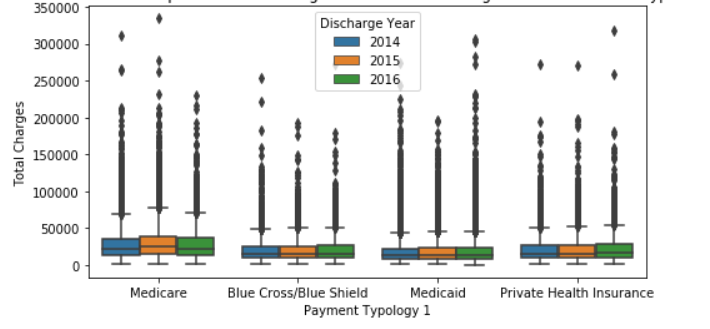
**Figure6 The relationship between Gender and LOS among different insurance type**



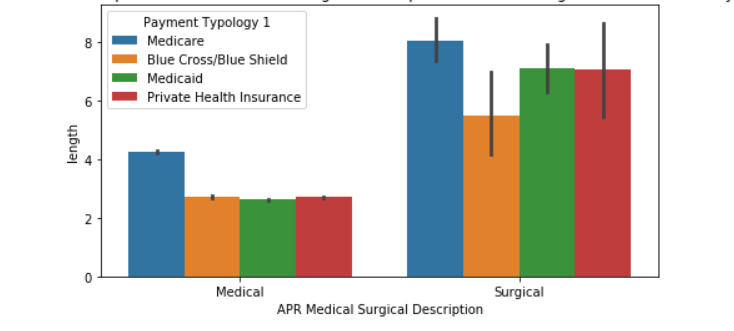
**Figure7 The relationship between Gender and total charges among different discharge year**



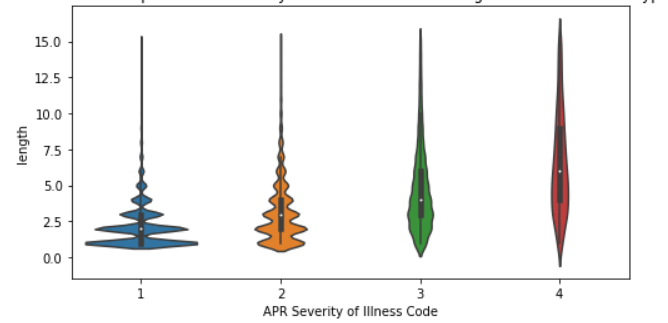
**Figure8 The relationship between Type of Admission and LOS among different insurance type**



**Figure9 The relationship between Discharge Year and LOS among different insurance type**



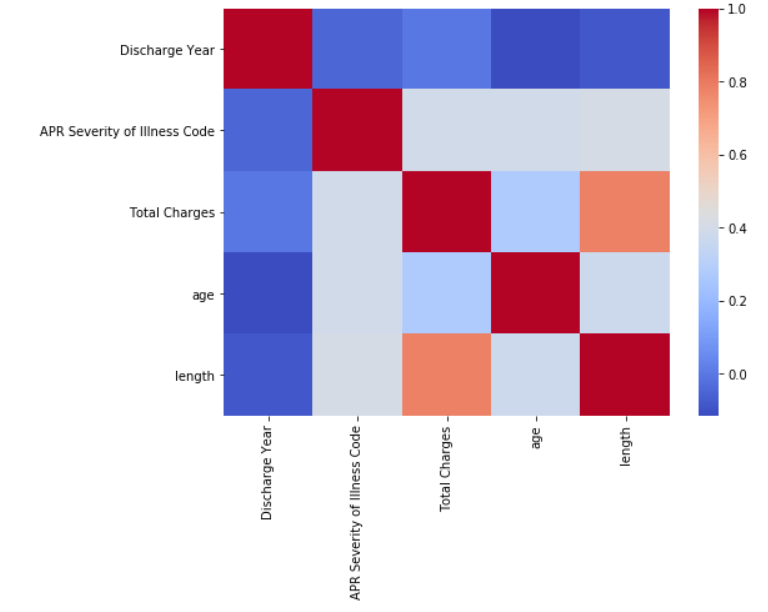
**Figure10 The relationship between APR Medical Surgical Description and LOS among different insurance type**



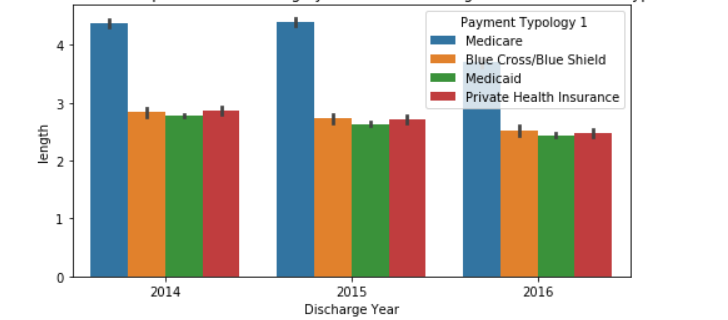
**Figure11 The relationship between Severity of Illness and LOS among different insurance type**

The data visualization part has consistently demonstrated that many nonclinical factors affect LOS including race, age, sex, severity of illness, Medical Surgical and so on.

Conducting a correlation test to eliminate the redundant features.

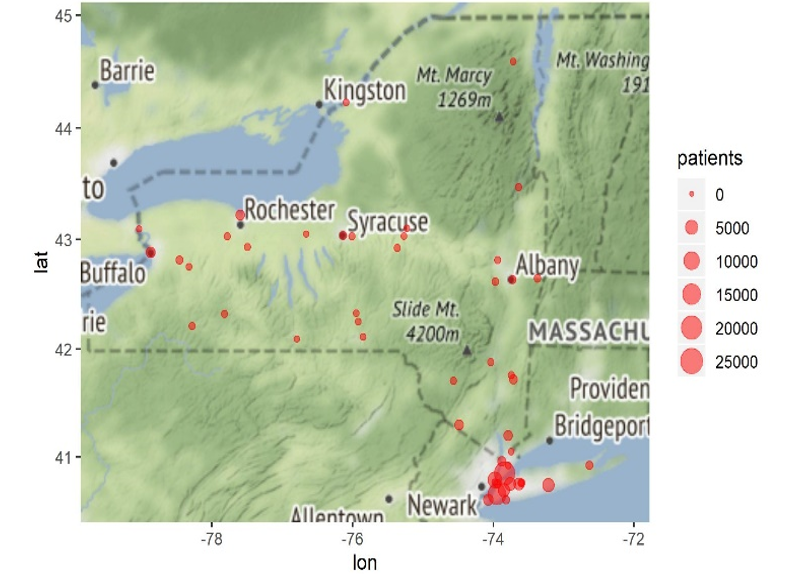


**Figure12 The heating map of each two factors**

****

There is a time trend, with year goes by, the length of stay in hospital tends to decrease. We'd better control the time trend that are constant across patients but vary over time.

We can see the distribution of patients in each hospital from Figure1, most of the patients are in New York City. However, there are also many patients scattered in other counties. We consider controlling the characteristics of each hospital that might affect the length of stay in the hospital. For example, patients in urban hospitals might have higher quality medical facilities and income than in rural areas, which might influence the length of stay in the hospital. We know that regressions relying on across hospital variation are problematic due to potential omitted variable bias. Thus, we add hospital-level fixed effects to focus on within hospital variation.



**Figure1 The distribution of patients in hospitals for New York Area**

**Regression model**

To investigate how the different medical insurance types influence the length of stay in hospital in asthma patients after control the time and hospital level fixed effect, we build the following specification:

*= ++ +*

*+ + +*

i = 1,2…. 81097

t = 2014,2015,2016

h = hospital

For patient i, the coefficient of the 𝛼\_0 represent the length of stay in hospital for patients with black race, multi-ethnicity, extreme risk of mortality and self-selected into hospital and discharge in 2014, paid by private health care insurance.

Demographic variable includes the patient information, such as race, ethnicity, gender and age. Medical condition includes admission type, discharge year, severity of the illness, treatment method, total charges.

**Regression results**

As we can see from Table3, after we control the characteristics of the hospital and time trend, if we keep other factors unchanged, compared with private health insurance type, patients with blue cross and blue shield health insurance status don't have a significant difference in length of stay in hospital. Patients with Medicare or Medicaid insurance have a significantly longer length of stay in hospital compared with private health insurance status. Generally, compared with private health insurance type, there is a 99.9% probability, patients with Medicare insurance tend to stay 0.21 days longer in hospital, and there is a 95% probability, patients with Medicaid insurance tend to stay 0.03 days longer in hospital.

It might result from two reasons. On the one hand, Blue Cross Blue Shield is a typical health insurance provider because it is neither a for-profit privately-owned company nor a program run by the federal government. Blue Cross Blue Shield has partnered with several private health insurance. It has little difference with some private insurance. On the other hand, different private health insurance in the various counties might have different terms or conditions, which leads to a different length of stay in the hospital. After we control these across county variation and focus on the within county variation, we get a different result.

Let's look at other factors that might influence the length of stay in hospital:

For demographic attributes, the race, in model 2, we can see white patients have a significantly longer length of stay in the hospital, patients whose race are white tends to stay in the hospital longer than black patients. However, ethnicity doesn't play a significant role in the length of stay in the hospital. Age and total charges play a significant role in the length of stay in the hospital. The older the patient, the longer they stay in the hospital.

For Medical condition attributes, the emergency type of admission has a significant influence on the length of the hospital. Compared with elective admission type of patients, if the patients have an emergency type of access, he tends to stay in hospital 0.21 days shorter. The regression result also shows the more severe the disease, the longer the length of stay in the hospital. For the risk of mortality, the higher risk of death, the longer length of stay in the hospital. For the treatment method, compared with patients who are treated as a medical method, patients who are treated as surgical tends to have a significantly two days shorter length of stay in the hospital. It might because patients who have conservative treatment tend to stay in the hospital for further observation.

Table3 Regression Result

|  |  |  |
| --- | --- | --- |
|  | Model 1 | Model 2 |
| (Intercept) | 1.11 \*\*\* |  |
|  | (0.11) |  |
| Multi-racial | -0.52 \*\*\* | 0.02 |
|  | (0.07) | (0.06) |
| Other Race | -0.14 \*\*\* | 0.00 |
|  | (0.01) | (0.01) |
| White | 0.13 \*\*\* | 0.06 \*\*\* |
|  | (0.01) | (0.01) |
| Not Span/Hispanic | 0.27 \*\*\* | -0.11 |
|  | (0.07) | (0.07) |
| Spanish/Hispanic | 0.20 \*\* | -0.10 |
|  | (0.07) | (0.07) |
| Emergency admission | -0.70 \*\*\* | -0.21 \*\*\* |
|  | (0.06) | (0.05) |
| Urgent admission | -0.44 \*\*\* | -0.11 |
|  | (0.07) | (0.06) |
| year2015 | -0.17 \*\*\* |  |
|  | (0.01) |  |
| year2016 | -0.37 \*\*\* |  |
|  | (0.01) |  |
| severity\_2 | 0.09 \*\*\* | 0.10 \*\*\* |
|  | (0.01) | (0.01) |
| severity\_3 | 0.40 \*\*\* | 0.28 \*\*\* |
|  | (0.02) | (0.02) |
| severity\_4 | 0.23 \*\*\* | -0.22 \*\*\* |
|  | (0.05) | (0.04) |
| Major mortality | 0.04 | 0.29 \*\*\* |
|  | (0.05) | (0.05) |
| Minor mortality | -0.17 \*\* | 0.19 \*\*\* |
|  | (0.06) | (0.05) |
| Moderate mortality | -0.16 \*\* | 0.15 \*\* |
|  | (0.06) | (0.05) |
| Surgical treatment | -1.21 \*\*\* | -2.01 \*\*\* |
|  | (0.09) | (0.08) |
| Blue cross/ blue shield | 0.07 \*\* | 0.01 |
|  | (0.02) | (0.02) |
| Medicaid | 0.23 \*\*\* | 0.03 \* |
|  | (0.02) | (0.01) |
| Medicare | 0.36 \*\*\* | 0.21 \*\*\* |
|  | (0.02) | (0.02) |
| Charges | 0.00 \*\*\* | 0.00 \*\*\* |
|  | (0.00) | (0.00) |
| Age | 0.01 \*\*\* | 0.01 \*\*\* |
|  | (0.00) | (0.00) |
| nobs | 77076 | 77076 |
| r.squared | 0.68 | 0.76 |
| adj.r.squared | 0.68 | 0.76 |
| sigma | 1.46 | 1.25 |
| statistic | 7057.56 | 1150.68 |
| p.value | 0.00 | 0.00 |
| df | 77052.00 | 76859.00 |
| df.residual | 77052.00 | 76859.00 |
| \*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05. | | |

4.Hypothesis Test

Since we have known how different health insurance types influence the length of stay in hospital compared with private health insurance. We now further investigate the relationship between each of the two-health insurance.

The first hypothesis is that: we assume that patients with Medicaid or Medicare insurance type have no significant influence on the length of stay in hospitals. The null hypothesis and alternative hypothesis are as follows:

The second hypothesis is that: we assume that patients with Private or Medicare insurance type has no significant influence on the length of stay in hospitals. The null hypothesis and alternative hypothesis are as follows:

The third hypothesis is that: we assume that patients with Private or Medicaid insurance type has no significant influence on the length of stay in hospitals. The null hypothesis and alternative hypothesis are as follows:

After doing the Wald test, we got this statistical result summarized in table 4. There is no significant difference in length of stay in the hospital for patients with The Blue Cross/Blue Shield and Medicaid insurance. However, patients with Medicare insurance type shows a significant difference in length of stay in hospital compared with patients with Blue cross/blue shield insurance.

Table 4 Wald test result for hypothesis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Insurance type | chi2 | df1 | qchisq(0.9,df1) | result |
| Blue cross/Blue shield and Medicaid | 1.6875 | 1 | 2.706 | Not rejected |
| Blue cross/Blue shield and Medicare | 1.025e+02 | 1 | 2.706 | Reject |
| Medicaid and Medicare | 1.461e+02 | 1 | 2.706 | Reject |

**Conclusion**

**I am still working on this part.**

1. Health insurance coverage in the United States https://en.wikipedia.org/wiki/Health\_insurance\_coverage\_in\_the\_United\_States [↑](#footnote-ref-1)
2. Berchick, Edward R., Emily Hood, and Jessica C. Barnett. "Health insurance coverage in the United States: 2017." Current population reports. Washington DC: US Government Printing Office (2018). [↑](#footnote-ref-2)
3. Brasel, Karen J., et al. "Length of stay: an appropriate quality measure?." Archives of Surgery 142.5 (2007): 461-466. [↑](#footnote-ref-3)