![Brown School of Public Health logo]

BROWN
School of Public Health

Effectiveness of Two Insurance Companies in Minimizing Length of Hospitalizations for Pneumonia Cases

David Moon
PHP1510

Introduction

In healthcare environments, there is a constant battle to reduce hospitalization costs while ensuring high quality services. One of the key factors contributing to high costs is the length of stay (LOS). It has been suggested that some insurance companies have been more successful than others at minimizing hospital length of stay.

In particular, this study examines differences between two of the largest insurance companies (termed companies A and B in this paper) in the U.S using a provided dataset that includes information from pneumonia patients' medical records about the patients themselves, as well as characteristics from hospitals in the U.S.

Pneumonia is an infection that inflames the air sacs in one or both lungs. It is considered most serious for infants and young children, people older than age 65, and people with weakened immune systems.

The following investigation seeks to test whether one of the two companies (A or B) has been more successful at minimizing the length of hospitalization (and hence the respective healthcare cost).

Methods

Dataset

The total sample size is 787, of which 392 pneumonia cases were insured with company A and 395 pneumonia cases were insured through company B. The patient data comes from 29 randomly selected metropolitan hospitals in the U.S. Refer to table A for a description of variables included in this dataset.

Statistical Procedure

First, statistical analyses were used to examine differences in patient characteristics for both companies. To examine the differences in complications for companies A and B, a two-tailed proportion test (α = 0.05) was performed under the null hypothesis that there is no difference in the proportion of complications per company. A chi-squared test (α = 0.05) of company type and

race was also performed under the null hypothesis that race and company type are independent.

Analyses used to examine the differences in hospital characteristics for companies A and B involved a two-tailed two proportion test ($\alpha = 0.05$) for hospital types for a given company type (under the null hypothesis that the proportions of the patients being treated at private hospitals are the same for both companies). A two-tailed t-test ($\alpha = 0.05$) to determine whether there was a significant difference in the average number of hospital beds available to patients insured by each company.

The primary research question was answered using a two-tailed t-test ($\alpha = 0.05$) in the difference of average LOS for companies A and B.

Results

Upon preliminary exploration of this dataset, there were several key findings from the patient characteristics. In addition to the fact that there were ~1.8x more boys than girls (503 compared to 284), the age range in this dataset was ages 2-17. The distribution of these ages were slightly right-skewed (figure A) for both companies A and B. When comparing the representation of various races in the dataset, analysis revealed that there were far fewer Asian patients, as well as patients categorized as "other" for both companies (figure B). A chi-squared test (df = 4) examining race and company type yielded a p-value of 9.817e-13, leading to a rejection of the null (the race and company variables are not independent).

A two-tailed two proportion test (df = 1) revealed that there is not enough evidence to reject the null that the proportion of complications for each company is different (p-value = 1). The sample proportion of complications out of all patients

insured under A was 0.018, and the same statistic was 0.020 under B.

In the hospital-related characteristics, data analysis revealed that the number of hospital beds for a given hospital ID varied (hospital 17 had the highest average number of beds with 532, while hospital 10 had the lowest with 43.1 beds). With this in mind, a two-tailed t-test (df = 777.21) yielded that there is no significant difference in the average number of beds available to patients insured by companies A and B (p-value = 0.737).

For private hospitals, the distributions of LOS for both insurers was right-skewed. For public hospitals, insurer A's LOS values had a heavily left-skewed distribution, while insurer B had no clear skewness (figure D). In general, there were more patients treated at private hospitals in this dataset, and a two-tailed proportion test (df = 1) yielded a p-value of 1.591e-14. This led to the conclusion that there is a significant difference in the proportion of patients treated at private hospitals (compared to public) between the two insurers. We estimate with a 95% confidence level that the true difference is between (0.155, 0.260), and it appears that the proportion of patients in a private hospital is higher for company A (estimate of 0.936) than B (estimate of 0.729).

The distribution of LOS for both companies was right-skewed (figure C). Company A had a median LOS of 2 and standard deviation of 1.16; company B had a median LOS of 3 and standard deviation of 1.53. Under a two-tailed t-test (df=736.3), we reject the null hypothesis that there is no difference in the average length of stay between company A and B (p-value = 2.03e-9). Furthermore, we are 95% confident that the true difference in means lies within the range (-0.778, -0.398) and, as a result, it appears the LOS for company B is higher than that of company A.

Discussion

The results demonstrate that company A is more successful than company B at minimizing its LOS (and thus its healthcare cost) for pneumonia patients; these findings come from the two-tailed t-test comparing the mean LOS for both companies.

While the overall goal set out to be answered in this paper was whether one of the two companies has been more successful in minimizing LOS, it is clear that there are weaknesses in the dataset, and thus limitations to the analysis. To begin, given that only pneumonia and boys/girls from ages 2-17 are studied, an overall conclusive statement cannot be made about one company's success over another. For instance, a disease with a much longer average LOS can be more prevalent in one company.

For the two-tailed t-test used to answer the primary research question, the sample size is large enough (despite skewness) to state the means of LOS for both companies is normally distributed. However, while it is known that hospitals are randomly selected, we cannot conclude from this dataset that patients were randomly sampled, resulting in a limitation of this statistical analysis.

Future directions for this investigation involves gathering more diverse data points for patients from each insurer (and possibly posing an altered question to reflect this adjusted dataset). Collecting more information on other proven determinants of healthcare costs for insurers could warrant novel statistical analyses involving more paired t-tests.

Appendix

Var Name	Description
Patient Characteristics	
"id"	Patient identification no.
"age	Age (in years)
"sex"	Sex (boy or girl)
"race"	White, Hispanic, African-American, Asian, or other
"los"	Length of hospital stay (in days)
"complic"	If there were any treatment complications
Hospital Characteristics	
"hosp.id"	Hospital identification no.
"beds"	Total number of beds
"type"	Hospital owner (public or private)
"company"	Insurance company (A or B)

Table A. *Information from the study included in this dataset*
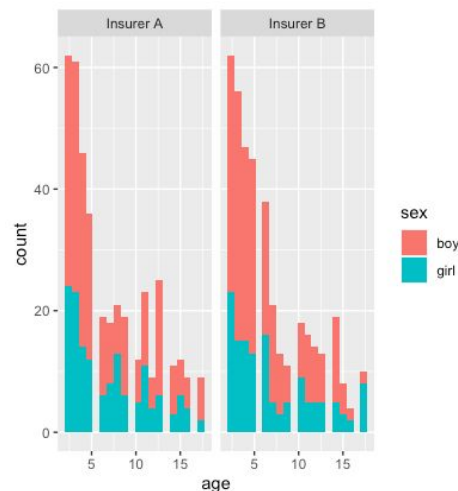

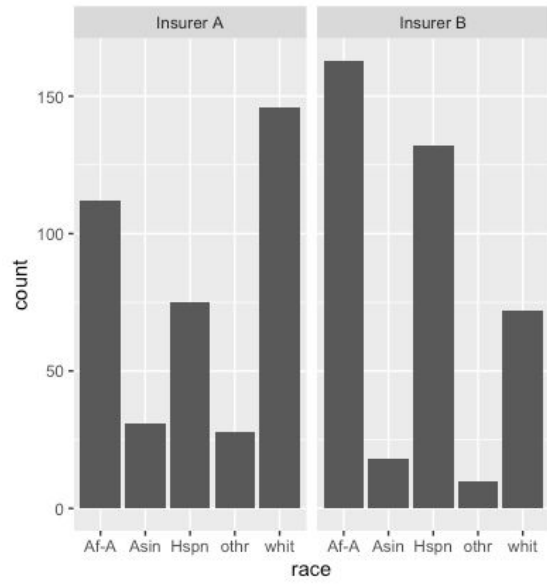Fig A. *Distributions of age filled by sex and faceted by company*

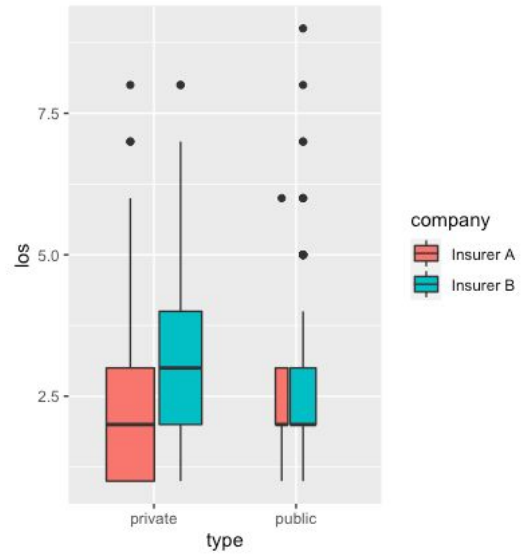Fig B. *Bar chart of various races of patients faceted by company*
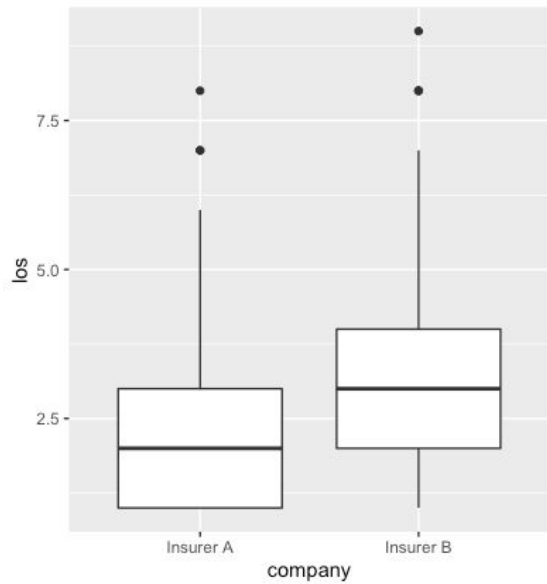

Fig D. *Boxplots of LOS based on hospital type and company type*

Fig C. *Boxplot of LOS for both companies*