Statistical analysis on length of stay in hospital

Dung Tien Nguyen1,2,*, Phuc Dang Ho3, Thien Chi Nguyen4, Nguyen Thi Cam Van1,2

ABSTRACT
The rising financial problems of healthcare institutions make studies of resource distribution more and more important and valuable. Among these studies, identification of length of stay of hospital patients (LOS) has attracted many scientists recently since it contributes to better knowledge of hospital costs and helps these institutions control the costs. This paper is devoted to study the length of stay of inpatients in hospital. Although predicting the length of stay is difficult, it is actually useful and beneficial if some key factors that have influence on patient length of stay could be determined. This paper will be the basis for a running example that illustrates alternative models of the length of stay of hospital patients. A total of 1189 episodes, which contains patient records, were analyzed by using some parametric and nonparametric statistical methods. In this study, several factors are first considered and investigated, including date of admission, medical admission unit, diagnosis result, international classification of diseases (icd), age, province, profession, recovery status when discharged, ethnic, and etc. Multiple regression analysis was also carried out for modeling length of stay as a function of several independent variables. Since the number of inpatient hospital stays is concerned, the family of Poisson distributions is used in this study. This approach is also supported by the corresponding histogram. Furthermore, univariate analyses showed that age, province, profession, admission quarter, recovery status when discharged, and diseases significantly influence on LOS. Finally, multivariate analysis of multiple regression model emphasized that type of disease, admission quarter, age group, and profession are the key factors that influence the LOS. These results may have some economic and clinical implications for not only patients but also hospitals.

Key words: influence factors, length of stay, non-parametric tests, Poisson regression

INTRODUCTION
With the quickly increasing in health care costs, governments and humanitarian funding agencies are still looking for mechanisms to control and evaluate the effectiveness of medical care. One of key factors that have high impact on the medical examination and treatment process, especially their cost, is the length of stay (LOS) of inpatients. In fact, a valid approximation of LOS and accurately identifying some factors that influence the LOS can help significantly improve hospital discharge planning. In addition, this information may help patients and their family to get better preparation.

There have been many statistical models studying LOS of inpatient1–4. These studies were conducted using data sets of about 1000 to 2000 patients on various pathologies at some local hospitals. Their results show that many factors influence inpatient hospital stay. Regarding to research results on inpatients in Vietnam, Viet et al.5 studied the number of inpatients changing during 2003–2007. However, the authors simply used descriptive statistics and neither inferential statistics nor modeling was carried out. In this study, we will explore some factors that potentially influence LOS at a local hospital. The aim of this study is to determine whether there are differences in LOS among categories defined by independent variables, and which factors significantly predict the variation in LOS. The rest of this paper is arranged as follows. In Section Method, we introduce the methods used in this research. Section Results and Multivariate analysis discuss the main results of the paper. Finally, some discussion and conclusion are given in Section Discussion and conclusion.

METHODS
Subjects
The hospital data for the period December 2015 to December 2016 were obtained from the database system of a hospital in Central Highlands of Vietnam, resulting in a total of 1209 episodes. Both reference tables of ICD-9 and ICD-10 published by the World Health Organization (WHO) (see Cartwright6 and Weatherspoon7 for more details) were used to classify disease and other health problems recorded in the database.

Cite this article: Nguyen D T, Ho P D, Nguyen T C, Van N T C. Statistical analysis on length of stay in hospital. Sci. Tech. Dev. J. – Engineering and Technology; 3(513):SI97-SI106.
The ICD-10 diagnosis codes were groups based on advice from local clinical doctors. In order to ensure adequate sample size, the groups that consist of less than 0.8% number of all episodes were excluded. This reduced the numbers of episodes to a total of 1189. The following major diagnostic category were concerned: i) certain infectious and parasitic diseases, ii) diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism, iii) diseases of the nervous system, iv) diseases of the ear and mastoid process, v) diseases of the circulatory system, vi) diseases of the respiratory system, vii) diseases of the digestive system, viii) diseases of the skin and subcutaneous tissue, ix) diseases of the musculoskeletal system and connective tissue, x) diseases of the genitourinary system, xi) injury - poisoning - and certain other consequences of external causes.

The dependent variable in the study is the LOS. The factors of a) age, b) the province in which the patient resides, c) residential community, d) profession, e) seasonality of admission and discharge, f) admitted day of the week, g) status of patients when discharged, h) gender, i) ethnicity, and j) diagnosis classification were analyzed to check the influences on LOS.

Statistical Analysis
The Shapiro-Wilk test was used to test the normality of LOS for each group. If the distribution is reasonably normal, then the t-test and the Anova are used to test the significant effect on LOS. Mann-Whitney test was used to compare LOS between two groups, and Kruskal-Wallis test was used to compare LOS between more than two groups. The results were considered to be significant provided that p-value is less than 5%. An appropriate transformation of the LOS was performed before Poisson regression was used. The data were randomly split into 90% of records of training part, and the remaining 10% of records part to test model. The analysis in this study was carried out by using R.

RESULTS
The overall average LOS was 15.7 days with the standard deviation 4.4 days whereas the median LOS was 16 days and the interquartile range (IQR) was 5 days. This shows a negatively skewed distribution of LOS. The average age of patients was 51 years old (range 1 to 84). The plot of LOS versus age shows no pattern. However, if the patients were divided into four groups of age: 1-17, 18-39, 40-60, and 61-84 then the boxplot shows an increasing pattern. The Kruskal-Wallis test indicates that LOS is significantly different among the age groups (p = 2.096 × 10⁻⁷). Moreover, the Dunn test in Table 1 shows the significant increasing tendency of LOS across the age groups, presenting the phenomenon that older people need more time for recovery from illness. For convenience, we will use [*], [**], and [***] to indicate that 0.01 < p < 0.05, 0.005 < p < 0.01, and p < 0.005, respectively.

| Comp. | Z     | Pu       | Pad     | Sig.  |
|-------|-------|----------|---------|-------|
| 1-2   | -2.6982 | 6.9697e-03 | 8.3637e-03 | **    |
| 1-3   | -4.2759 | 1.9036e-05 | 5.7109e-05 | ***   |
| 2-3   | -3.3213 | 8.9581e-04 | 1.3437e-03 | ***   |
| 1-4   | -4.6252 | 3.7407e-06 | 2.2444e-05 | ***   |
| 2-4   | -3.9413 | 8.1032e-05 | 1.6206e-04 | ***   |
| 3-4   | -1.1529 | 2.4894e-01 | 2.4894e-01 |       |

The provinces where the patients reside were collapsed into two groups: the local province and the others. By the virtue of Wilcoxon rank-sum test, we notice that LOS significantly decreases for those who reside locally with p = 0.01451. The patients’ professions were classified into six groups: government officers (1), intellectuals (2), senior citizens (3), students (4), farmers (5), the others (6). The Kruskal-Wallis test confirms that LOS is significant different among these groups with p = 2.535 × 10⁻⁷. Table 2 provides more details on multiple
comparisons between the groups. Thus, the profession groups can be arranged in increasing order of LOS as follows: intellectuals (2) and students (4), government officers (1) and farmers (5), senior citizens (3) and the others (6).

Table 2: Multiple comparison between professions

| Comp. | Z         | P_unadj | P_adj  | Sig  |
|-------|-----------|---------|--------|------|
| 1 - 2 | 2.7767    | 5.4921e-03 | 0.0082 | **   |
| 1 - 3 | -1.7622   | 7.8037e-02 | 0.0975 |      |
| 2 - 3 | -4.0007   | 6.3151e-05 | 0.0003 | ***  |
| 1 - 4 | 3.2078    | 1.3373e-03 | 0.0040 | ***  |
| 2 - 4 | 1.1453    | 2.5208e-01 | 0.2909 |      |
| 3 - 4 | 3.9998    | 6.3394e-05 | 0.0002 | ***  |
| 1 - 5 | 0.4040    | 6.8621e-01 | 0.6862 |      |
| 2 - 5 | -2.8205   | 4.7946e-03 | 0.0080 | **   |
| 3 - 5 | 3.1409    | 1.6843e-03 | 0.0036 | ***  |
| 4 - 5 | -3.1973   | 1.3869e-03 | 0.0035 | ***  |
| 1 - 6 | -2.0629   | 3.9125e-02 | 0.0534 |      |
| 2 - 6 | -4.0825   | 4.4557e-05 | 0.0003 | ***  |
| 3 - 6 | -0.7862   | 4.3174e-01 | 0.4626 |      |
| 4 - 6 | -4.1450   | 3.3974e-05 | 0.0005 | ***  |
| 5 - 6 | -2.8615   | 4.2169e-03 | 0.0079 | **   |

Patients’ dates of admission to the hospital were categorized into quarters: January to March (1), April to June (4), July to September (7), and October to December (10). The Kruskal-Wallis test shows a significant LOS difference among the groups with p < 2.2 × 10⁻¹⁶ (Table 3). In particular, LOS is longer during the second and the third quarters each year and shorter during other quarters. Besides, omitting 12 records of patients admitted to the hospital during weekends, Kruskal-Wallis test shows that the LOS is not significantly different among weekdays (p-value > 0.5).

There are three categories for the status of patients when discharged: fully recovered, partially recovered, and unrecovered. This status significantly affected the LOS (p = 7.765 × 10⁻⁶). The median LOS for those who were not recovered was 5 days while the median LOS for those fully or partially recovered was 16 days (Table 4).

There was no significant ethical effect on LOS, the p-value of the corresponding Kruskal-Wallis test is 0.228. The median LOS for patients of majority ethnic groups was 16 days while that of minority ethnic groups was 17 days. In addition, there were a few more women (53.7%) than men (46.3%) who were admitted to the hospital, and there is not significant evidence to claim that gender influences LOS, due to p = 0.1944 of Wilcoxon rank-sum test.

Concerning the patients’ diseases, the WHO disease classification standard was used to categorize the recorded diseases into: certain infectious and parasitic diseases (1), diseases of the nervous system (6), diseases of the ear and mastoid process (8), diseases of the circulatory system (9), diseases of the respiratory system (10), diseases of the digestive system (11), diseases of the skin and subcutaneous tissue (12), diseases of the musculoskeletal system and connective tissue (13), diseases of the genitourinary system (14), injury, poisoning and certain other consequences of external causes (19). The Kruskal-Wallis test was then applied to conclude that patients’ diseases affect the LOS (p < 2.2 × 10⁻¹⁶). Furthermore, the Dunn test of multiple comparison (Table 5) allows us to conclude that the LOS of patients in the following groups is significantly longer than that in the other groups: diseases of the nervous system (6), diseases of the circulatory system (9), diseases of the musculoskeletal system and connective tissue (13), and injury, poisoning and certain other consequences of external causes (19).
Table 3: Multiple comparison between admitted day groups

| Comp. | Z   | P unadj | P adj | Sig. |
|-------|-----|---------|-------|------|
| 1 - 10| 1.586| 1.12e-01| 1.12e-01|     |
| 1 - 4 | -3.685| 2.28e-04| 3.42e-04| *** |
| 10 - 4| -6.470| 9.76e-11| 2.92e-10| *** |
| 1 - 7 | -5.534| 3.11e-08| 6.23e-08| *** |
| 10 - 7| -9.072| 1.15e-19| 6.95e-19| *** |
| 4 - 7 | -1.997| 4.58e-02| 5.49e-02| **  |

Table 4: Multiple comparison between recovery status

| Comp. | Z   | P unadj | P adj | Sig. |
|-------|-----|---------|-------|------|
| 1 - 2 | -1.021| 3.071e-01| 3.071e-01|     |
| 1 - 3 | 4.602 | 4.174e-06| 6.261e-06| *** |
| 2 - 3 | 4.831 | 1.356e-06| 4.069e-06| *** |
| Comp. | Z     | Unadj  | Padj  | Sig  |
|-------|-------|--------|-------|------|
| 1 - 10| -1.4048| 1.600e-01| 2.400e-01|      |
| 1 - 11| -1.7617| 7.811e-02| 1.351e-01|      |
| 10 - 11| -0.6425| 5.205e-01| 5.855e-01|      |
| 1 - 12| -1.9764| 4.810e-02| 8.658e-02|      |
| 10 - 12| -0.7602| 4.471e-01| 5.438e-01|      |
| 11 - 12| 0.0277| 9.778e-01| 9.778e-01|      |
| 1 - 13| -4.7135| 2.434e-06| 1.825e-05| *** |
| 10 - 13| -4.8963| 9.763e-07| 1.098e-05| *** |
| 11 - 13| -2.7165| 6.597e-03| 1.484e-02|  *  |
| 12 - 13| -3.9222| 8.772e-05| 3.289e-04| *** |
| 1 - 14| -3.6598| 2.523e-04| 8.111e-04|      |
| 10 - 14| -3.0813| 2.060e-03| 5.151e-03|      |
| 11 - 14| -1.6920| 9.063e-02| 1.510e-01|      |
| 12 - 14| -2.2449| 2.476e-02| 4.846e-02|      |
| 13 - 14| 1.59543| 1.106e-01| 1.777e-01|      |
| 1 - 19| -1.4061| 1.596e-01| 2.477e-01|      |
| 10 - 19| 0.0348| 9.722e-01| 9.943e-01|      |
| 11 - 19| 0.6853| 4.931e-01| 5.689e-01|      |
| 12 - 19| 0.8225| 4.107e-01| 5.281e-01|      |
| 13 - 19| 5.2940| 1.196e-07| 2.691e-06|      |
| 14 - 19| 3.2665| 1.088e-03| 3.062e-03|      |
| 1 - 6 | -4.8904| 1.006e-06| 9.056e-06|      |
| 10 - 6 | -5.1346| 2.827e-07| 4.240e-06|      |
| 11 - 6 | -2.9326| 3.360e-03| 7.959e-03|      |
| 12 - 6 | -4.1838| 2.865e-05| 1.432e-04|      |
| 13 - 6 | -1.0059| 3.144e-01| 4.288e-01|      |
| 14 - 6 | -1.9910| 4.647e-02| 8.713e-02|      |
| 19 - 6 | -5.5314| 3.175e-08| 1.428e-06|      |
| 1 - 8 | -1.0356| 3.003e-01| 4.360e-01|      |
| 10 - 8 | 0.2566| 7.974e-01| 8.544e-01|      |
| 11 - 8 | 0.7786| 4.361e-01| 5.452e-01|      |
| 12 - 8 | 0.8769| 3.804e-01| 5.035e-01|      |
| 13 - 8 | 3.8069| 1.406e-04| 4.869e-04|      |
| 14 - 8 | 2.6731| 7.514e-03| 1.610e-02|      |
| 19 - 8 | 0.2344| 8.146e-01| 8.525e-01|      |
| 6 - 8 | 4.0113| 6.036e-05| 2.716e-04|      |
| 1 - 9 | -4.2701| 1.953e-05| 1.255e-04|      |
| 10 - 9 | -3.9772| 6.971e-05| 2.852e-04|      |
| 11 - 9 | -2.3522| 1.865e-02| 3.816e-02|      |
| 12 - 9 | -3.1347| 1.720e-03| 4.553e-03|      |
| 13 - 9 | 0.2743| 7.838e-01| 8.602e-01|      |
| 14 - 9 | -1.0125| 3.112e-01| 4.377e-01|      |
| 19 - 9 | -4.2142| 2.506e-05| 1.409e-04|      |
| 6 - 9 | 0.7447| 4.564e-01| 5.404e-01|      |
| 8 - 9 | -3.3459| 8.199e-04| 2.459e-03|      |
The LOS descriptive parameters on groups of significant factors are summarized in Table 6.

**MULTIVARIATE ANALYSIS**

In this section, Poisson Regression Model, as a popular multivariate statistical method for count data, is applied to investigate the factors that truly affected on LOS, and to predict LOS. In view of Figure 1, it seems that the LOS itself does not follow Poisson distribution. That's why we first apply a transformation $T_{LOS} = 35 - LOS$ to make the distribution look “more Poisson likely” (see Figure 2).

![Figure 2: The transformed LOS distribution](image)

Based on a test statistic proposed by Hoaglin, we get the plot in Figure 3.

![Figure 3: Checking for Poisson distribution](image)

The linear pattern in Figure 3 shows that $T_{LOS}$ may follow Poisson distribution. Then we can use the multiple Poisson regression model to fit the data. We built the model by testing for the inclusion of each new independent variable, one after another. Using the results obtained from univariate tests, we arrange the factors in increasing order of the p-values (see Table 7). If the variable’s addition was significant at the significant level 5%, then we keep it in the model. Notice that the factor “recovery status” will not be included in the model since using this status to predict the LOS makes no sense.

The result on Poisson regression model is summarized in Table 9. In this model, the factor “province” was excluded since the corresponding p-value counted 0.634905. Notice that only 90% of the data set were used to train the model. In this table, the reference baseline includes: certain infectious and parasitic diseases (for factor “disease”), quarter I (for factor “quarter admission”), age group less than 18 years old (for factor “age”), and Government Officers (for factor “profession”).

The AIC value of this model fitting as 5950 and the residual deviance was 793.95 on 1048 degrees of freedom. We also obtained a summary of deviance residuals given in Table 9. Using the 10% remaining data, we tested the model and compute the errors which the difference between the predicted values and the observed values (Table 8). The IQR = 3.134 and SD = 3.819 shows that the prediction was quite reasonable.

**DISCUSSION AND CONCLUSION**

It is no doubt that understanding of factors that influence LOS is helpful for patients, hospitals, and health service system. This study has provided not only information about potential effect of many factors on LOS but also an appropriate model for predicting LOS. This enables better planning of resources and making decision. Similar research has not been studied in Vietnam. Then this study complements and enhances current research related to LOS.

Univariate analyses pointed out that the following factors have significant influence on LOS: age, province, profession, admission quarter, recovery status when discharged, and diseases. It is not a surprise that patients’ profession affects LOS since the professions influence their life style which, in turn, may have some impact on LOS. Moreover, intellectuals and students usually have good knowledge on their health and thus keep their lives healthy. The senior citizen group is the one staying longest at the hospital since senior people are not as strong as the younger and often need more time to recover from diseases.

Concerning the admission quarter, people get sick more likely from April to September, especially during summers. In addition, people normally want to stay at home with their family during Tet holiday and
Table 6: LOS analysis significant factor results

| Factors       | Group          | N   | Prop.  | Mean  | SD    | SE   | Md   | IQR |
|---------------|----------------|-----|--------|-------|-------|------|------|-----|
| Age           | <18            | 33  | 0.028  | 11.5  | 5.95  | 1.040| 13   | 9   |
|               | 18-39          | 222 | 0.187  | 14.6  | 4.87  | 0.327| 15   | 7   |
|               | 40-60          | 592 | 0.498  | 15.9  | 4.06  | 0.167| 16   | 4   |
|               | >60            | 342 | 0.288  | 16.3  | 4.02  | 0.217| 17   | 4   |
| Province      | Local          | 1108| 0.932  | 15.6  | 4.36  | 0.131| 16   | 5   |
|               | Nonlocal       | 81  | 0.068  | 16.5  | 4.33  | 0.481| 17   | 4   |
| Profession    | Government Officers | 132 | 0.111  | 15.7  | 3.99  | 0.348| 16   | 5   |
|               | Intellectuals  | 36  | 0.030  | 12.8  | 5.53  | 0.922| 15   | 10  |
|               | Senior citizens| 286 | 0.241  | 16.5  | 4.05  | 0.239| 17   | 4   |
|               | Students       | 15  | 0.013  | 10.7  | 5.42  | 1.400| 10   | 8   |
|               | Farmers        | 627 | 0.527  | 15.4  | 4.39  | 0.175| 16   | 6   |
|               | The others     | 93  | 0.078  | 16.7  | 3.83  | 0.398| 18   | 4   |
| Quarters      | I              | 166 | 0.140  | 14.7  | 4.65  | 0.361| 16   | 6   |
|               | II             | 282 | 0.237  | 16.3  | 4.24  | 0.252| 17   | 4   |
|               | III            | 359 | 0.302  | 16.9  | 4.27  | 0.225| 18   | 4   |
|               | IV             | 382 | 0.321  | 14.4  | 4.02  | 0.206| 15   | 5   |
| Recovery      | Fully recovered| 619 | 0.521  | 15.7  | 4.46  | 0.179| 16   | 5   |
|               | Partially recovered | 554 | 0.466  | 15.9  | 3.98  | 0.169| 16   | 5   |
|               | Unrecovered    | 16  | 0.013  | 7.81  | 6.09  | 1.52 | 5    | 8.5 |
| Diseases      | Certain infectious and parasitic diseases (1) | 10  | 0.008  | 7.2   | 3.88  | 1.23 | 6.5  | 2.75 |
|               | Diseases of the nervous system (6) | 308 | 0.259  | 16.6  | 3.74  | 0.213| 16   | 4   |
|               | Diseases of the ear and mastoid process (8) | 13  | 0.011  | 10.6  | 5.38  | 1.49 | 9    | 8   |
|               | Diseases of the circulatory system (9) | 57  | 0.048  | 16.5  | 3.42  | 0.453| 15   | 4   |
|               | Diseases of the respiratory system (10) | 26  | 0.022  | 10.8  | 5.53  | 1.080| 10   | 8   |
|               | Diseases of the digestive system (11) | 13  | 0.011  | 11.8  | 6.39  | 1.770| 12   | 8   |
|               | Diseases of the skin and subcutaneous tissue (12) | 27  | 0.023  | 14.1  | 1.97  | 0.380| 15   | 1   |
|               | Diseases of the musculoskeletal system and connective tissue (13) | 655 | 0.551  | 15.9  | 4.20  | 0.164| 17   | 5   |
|               | Diseases of the genitourinary system (14) | 50  | 0.042  | 15.6  | 4.03  | 0.569| 15   | 3.75 |
|               | injury, poisoning and certain other consequences of external causes (19) | 30  | 0.025  | 11.3  | 4.96  | 0.905| 12   | 6.5 |
Table 7: Summary of univariately significant independent variables on LOS

| Variable          | Test          | p-value      | Order |
|-------------------|---------------|--------------|-------|
| Disease           | Kruskal-Wallis| < 2.2e-16    | 1     |
| Admission quarter | Kruskal-Wallis| < 2.2e-16    | 2     |
| Age               | Kruskal-Wallis| 2.096e-7     | 3     |
| Profession        | Kruskal-Wallis| 2.535e-7     | 4     |
| Province          | Wilcoxon rank-sum | 0.01451   | 5     |

Table 8: Summary for residuals and errors

|        | Min   | 1st Q | Median | 3rd Q  | Max   |
|--------|-------|-------|--------|--------|-------|
| Residuals | -2.5807 | -0.5764 | -0.1508 | 0.3748 | 3.1731 |
| Errors   | 0.0791 | 0.9317 | 2.1406 | 4.0659 | 10.7255 |

dependence, leave the hospital as soon as possible in January and February. Many research papers have shown that admission day of week affects the LOS, which did not happen in our study. This fact might be due to small number of observations for admission in weekends (only 12 records).

If treatments for a person do not show expected progress after few days, he/she is often transferred to another hospital with better doctors. That is, a patient who could not get an appropriate treatment at the hospital is likely to leave the hospital sooner than the others.

The factor "ethnic" does not really affect the LOS. This emphasizes the good policies of the government on the equity of providing public healthcare services to all the citizens.

The results obtained here showed that patients in the following disease groups stay longer at the hospital: diseases of the nervous system (6), diseases of the circulatory system (9), diseases of the musculoskeletal system and connective tissue (13), and injury, poisoning and certain other consequences of external causes (19). A possible reason is that these diseases normally require longer inpatient treatments.

In the studied multiple regression model, type of disease, admission quarter, age group, and profession are the key factors that influence the LOS. This model can be applied to patients admitted to any hospital for individual patient expectation of their LOS and resource preparation. Further researches with more sufficient information collected are needed to reduce the error of the regression model. Further studies is also needed to investigate whether factors such as specialty of doctor, admission year, and distance between patients’ residential location and the hospital could be involved in the considered models.

ACKNOWLEDGMENT

The first author’s research is funded by Ho Chi Minh City University of Technology – VNU-HCM under grand number T-KHUD-2018-18.

LIST OF ABBREVIATION

LOS: Length of stay
TLOS: Transformed length of stay

CONFLICT OF INTEREST

We hereby declare that we have no conflict of whatsoever involved in publishing this research.

AUTHORS’ CONTRIBUTION

Dung Tien Nguyen: Conceptualization, Analysis, Writing, Validation, Data curation
Phuc Dang Ho: Editing, Validation
Thien Chi Nguyen: Writing, Analysis
Nguyen Thi Cam Van: Data curation

REFERENCES

1. Carter EM, Potts HW. Predicting Length of Stay from an Electronic Patient Record System: A Primary Total Knee Replacement Example. BMC Medical Informatics and Decision Making. 2014;14. PMID: 24708853. Available from: https://doi.org/10.1186/1472-6947-14-26.
2. Huang JQ, Hooper PM, Marrie TJ. Factors Associated with Length of Stay in Hospital for Suspected Community-Acquired Pneumonia. Canadian Respiratory Journal. 2006;13:317–324. PMID: 16983487. Available from: https://doi.org/10.1155/2006/325087.
3. Liu Y, Phillips M, Codde J. Factors influencing patients’ length of stay. Aust Health Rev. 2001;24:63–70. PMID: 11496473. Available from: https://doi.org/10.1071/AH010663.
4. Stecker M, Stecker M, Falotico J. Predictive Model of Length of Stay and Discharge Destination in Neuroscience Admissions. Surgical Neurology International. 2017;8:17. PMID: 28217396. Available from: https://doi.org/10.4103/2152-7806.199558.
| Coefficients     | Estimate | Std. Error | z-value | Pr(>|z|)     | Sig. | Explanation                                                                 |
|------------------|----------|------------|---------|--------------|------|-----------------------------------------------------------------------------|
| (Intercept)      | 3.485005 | 0.094941   | 36.707  | < 2e-16      | ***  |                                                                              |
| disease (6)      | -0.376245| 0.074801   | -5.030  | 4.91e-07     | ***  | diseases of the nervous system                                              |
| disease (8)      | -0.114318| 0.096562   | -1.184  | 0.236458     |      | diseases of the ear and mastoid process                                      |
| disease (9)      | -0.365574| 0.080327   | -4.551  | 5.34e-06     | ***  | diseases of the circulatory system                                           |
| disease (10)     | -0.120913| 0.084921   | -1.424  | 0.154496     |      | diseases of the respiratory system                                           |
| disease (11)     | -0.174236| 0.097151   | -1.793  | 0.072900     | .    | diseases of the digestive system                                             |
| disease (12)     | -0.238071| 0.085899   | -2.772  | 0.005579     | **   | diseases of the skin and subcutaneous tissue                                 |
| disease (13)     | -0.341208| 0.074075   | -4.606  | 4.10e-06     | ***  | diseases of the musculoskeletal system and connective tissue                 |
| disease (14)     | -0.296136| 0.080858   | -3.662  | 0.000250     | ***  | diseases of the genitourinary system                                         |
| disease (19)     | -0.163074| 0.083051   | -1.964  | 0.049583     | *    | injury, poisoning and certain other consequences of external causes         |
| quarter II       | -0.084461| 0.023536   | -3.589  | 0.000333     | ***  | quarter II                                                                  |
| quarter III      | -0.130390| 0.022693   | -5.746  | 9.14e-09     | ***  | quarter III                                                                  |
| quarter VI       | -0.004991| 0.022152   | -0.225  | 0.821757     |      | quarter IV                                                                   |
| age (2)          | -0.117596| 0.051815   | -2.170  | 0.029986     | *    | age 18-39                                                                    |
| age (3)          | -0.143833| 0.052744   | -2.727  | 0.006391     | **   | age 40-60                                                                    |
| age (4)          | -0.145292| 0.054687   | -2.657  | 0.007889     | **   | age > 61                                                                     |
| profession (2)   | 0.086687 | 0.043057   | 2.013   | 0.044082     | *    | Intellectual                                                                 |
| profession (3)   | -0.028681| 0.026595   | -1.078  | 0.280846     |      | Senior citizens                                                              |
| profession (4)   | 0.013693 | 0.078256   | 0.175   | 0.861099     |      | Students                                                                     |
| profession (5)   | 0.009698 | 0.023249   | 0.417   | 0.676579     |      | Farmers                                                                      |
| profession (6)   | -0.021622| 0.032993   | -0.655  | 0.512237     |      | The others                                                                   |

[*] 0.01 < p < 0.05, [**] 0.005 < p < 0.01, and [***] p < 0.005.
5. Viet NL, Tuan PV, Hung PM, Hanh VD, Quang NN. Study on Disease Patterns in Inpatients at Vietnam National Heart Institute During 2003-2007. Journal of Vietnamese Cardiology. 2010.

6. Cartwright DJ. ICD-9-CM to ICD-10-CM Codes: What? Why? How? Adv Wound Care (New Rochelle). 2013;2:588–592. PMID: 24761333. Available from: https://doi.org/10.1089/wound.2013.0478.

7. Weatherspoon D, Chattopadhyay A. International Classification of Diseases Codes and Their Use in Dentistry. J Dent Oral Craniofac Epidemiol. 2013;1:20–26.

8. Hoaglin DC. A Poissonness Plot. The American Statistician. 1980;34:146–149. Available from: https://doi.org/10.1080/00031305.1980.10483020.

9. Hoaglin DC, Tukey JW. Exploring Data Tables, Trends, and Shapes. Ed. by Hoaglin, D. C., Mosteller, F., and Tukey, J. W. Wiley, 2006. Chap. Checking the Shape of Discrete Distributions; Available from: https://doi.org/10.1002/9781118150702.