

A Study on the Injury Prediction of LOS, Discharge Results, ICU

Young-Hee Nam

Namseoul University, Chungnam, Korea
yhnam14@nsu.ac.kr

Abstract

The purpose of this study is to examine the correlation between the length of stay, use of ICU and discharge results of injured inpatients, and to examine the statistical method that predicts injury characteristics when applying three characteristics to injured inpatients. The data used in the data analysis were 3,773 cases from January 1, 2015 to December 31, 2015. Statistical analysis was Weka ver 3.6 open source software widely used in data mining to perform LOS prediction work was used. The results of the study showed that LOS was best predicted by LR method in the total dataset of injured inpatients, and that the results of discharge and ICU were highly predicted by DT method. In addition, the results of LOS prediction by injury foreigners showed that DT had excellent predictive power in TA and Burn, and DT and BN method had high predictive power in fall. This shows that DT is the best predictor of the independent variables related to injury when LOS, Results, and ICU are the dependent variables of the injured inpatients. In conclusion, applying the quality improvement program for trauma patients can lead to considerable cost reduction, so the treatment process to reduce LOS and complications is very important.

Keywords: Injury, Length of stay, Discharge results, Intensive care unit

1. Introduction

The average life expectancy has increased due to the development of medical technology, but the aspect and cause of damage are diversified due to accelerated industrialization and urbanization. Due to damage, 5 million people die worldwide every year, accounting for 9% of total deaths [8]. In Korea, the death toll from damage was 64.7 per 100,000 populations, the highest among OECD member countries [4], and damage is an important factor that threatens the safety and health of the whole nation. Damage is affected by the patient's hospitalization period according to the patient's condition at the time of hospitalization [6], and the cost of hospitalization is affected by the patient's disease characteristics, hospitalization period, clinical characteristics, and patient's personal characteristics. Especially, the hospitalization period according to the disease is most affected [1][2][3][6]. The average length of stay is higher when other diagnosis is accompanied than when only the main diagnosis [5]. Medical institutions are very interested in managing the period of hospitalization, and if there are many long-term patients, the national medical resources are efficiently used. In addition, each hospital leads to a decrease in profitability [7]. Therefore, this study aims to examine the correlation between the lengths of hospitalization of injured inpatients, whether they use intensive care unit, and the

Article history:

Received (May 12, 2019), Review Result (June 8, 2019), Accepted (July 3, 2019)

outcome of discharge, and what statistical methods predict damage characteristics when applying three characteristics to injured inpatients.

2. Method

2.1. Date

The damage data used in this study were collected from three general hospitals in Seoul and Gyeonggi-do. The total data was 4,047 cases, and the data were collected from January 1, 2015 to December 31, 2015. After removing the number of missing data, the final 3,773 cases were used for analysis.

2.2. Research architecture

Before building the model, each prediction model was created to create a data set, and the data set was preprocessed to select the variables to be used in each prediction model. Then, a prediction model was constructed using three statistical analysis techniques: Decision Tree (DT), Logistic Regression Analysis (LR), and Bayesian Network (BN). In addition, based on the results obtained by five-fold cross-verification, three statistical techniques were compared and analyzed to find the most appropriate statistical analysis technique to predict each injury (traffic accident, burn, fall) LOS (Length of stay, number of days of funding). Table 1 is a 20 basic characteristics used in each prediction, and it is a variable list applicable to LOS, Result, ICU. The data set was determined by basic characteristics with 18 LOS, 19 Results and 17 ICUs.

Table 1. Original features used for each prediction task

Features	Description	LOS	Result	ICU
Admission_Channel	Patient's admission channel	√	√	√
Admission_Month	Month a patient entered a hospital	√	√	√
Age	Patient's age in number	√	√	√
BMI	Level of body mass index	√	√	√
Coming Path	Path of coming hospital	√	√	√
Coming Method	Method of coming hospital	√	√	√
Diagnosis	Code for diagnosis (indicating the name of disease)	√	√	√
Ext_Cause_Code	Code for external cause	√	√	√
Gender	Male, female	√	√	√
ICISS_Score	Score representing the seriousness of injury	√	√	√
ICU	Whether a patient used a ICU or not	√	√	√
Injury_Exp	Whether a patient experienced injury before or not	√	√	√

Injury_Mechanism	Type of injury	√	√	√
Insurance Type	Type of insurance	√	√	√
LOC	Whether a patient is conscious or not	√	√	√
LOS	a (1 ~ 7 days), b (8 ~ 14 days), c (more than 15 days)	√	√	-
Main Treat_Dep	Main treatment department	√	√	√
Min(C)_Min(O)	The difference between coming time and injury occurrence time in minute	√	√	√
Operation	Whether a patient had a operation or not	√	√	-
Result	The outcome of discharge	-	√	-

2.3. Experiments' and model for prediction of LOS

The period of the funding was divided into three groups, a (~7 days), b (8~14 days), and c (more than 15 days). In addition, the data set was adjusted to predict LOS so that the number of cases is the same for each group. Therefore, a total of 2,436 cases (i.e., 812 cases in each stage) were used to construct LOS prediction models. To predict LOS, we discussed with experts and reviewed the literature on LOS prediction [1][2][3][6]. After evaluating the characteristics of the data set for LOS prediction using each statistical analysis technique, the final data set that has a greater impact was determined by predicting the target characteristics by stepwise backward removal using the wrapper approach.

Since it is possible to construct a prediction model with a similar patient group, the total data set for LOS prediction was divided into three groups, each of which included traffic accidents, burns and falls. The data set consisted of three damages, the top three with high incidence; the three damage groups were 993, 525 and 522, respectively. Unlike the characteristics of the entire data set, the new three data sets eliminate the characteristic of damage mechanism. Because in each of the three data sets, the characteristic value has only one value that represents the type of damage. In addition, a new function called accident type was added to the traffic accident data set to provide more specific information when constructing the prediction model. Weka ver. 3.6 open source software widely used in data mining, was used to perform LOS prediction. Based on the ranking search method, 18 characteristics were evaluated using Gain Ratio attribute evaluator. [Table 2] shows the importance rank of the input function for LOS prediction of each data set.

Through a sequential list of input variables, we adopted a wrapper approach using a stepwise backward method to select appropriate subsets for three different statistical analysis techniques, DT, LR, and BN. To construct a decision tree model, the C4.5 algorithm, which showed good results in previous studies, was used, and the parameters of the three statistical analysis techniques were set as basic values in Weka. The prediction model was constructed using three statistical analysis techniques, and the prediction results obtained from 5-fold cross-validation were evaluated and compared. For five-fold cross-validation, the data set is divided into five sub-data sets, one is used as a set of test data, and the rest is used to create a prediction model. Other sub data sets are then used as test data sets and the rest are used to build different prediction models. This is repeated five times, and the final prediction accuracy is obtained by the average of the five-time prediction model. The above approach to the selected

characteristics and constructed prediction model is similar to the discharge result and ICU prediction model.

Table 2. Ranking of features for prediction of LOS in each dataset

Ranking	Total	Traffic accident	Burn	Fall
1	ICU	Operation	Operation	ICU
2	Operation	ICISS_Score	Age	Diagnosis
3	ICISS_Score	Insurance_Type	Main_Treat_Dep	Operation
4	Ext_Cause_Code	Main_Treat_Dep	Insurance_Type	ICISS_Score
5	Diagnosis	Diagnosis	ICU	LOC
6	Main_Treat_Dep	Ext_Cause_Code	ICISS_Score	Ext_Cause_Code
7	Injury_Mechanism	Age	Ext_Cause_Code	Insurance_Type
8	Coming Method	Accident Type	Diagnosis	Coming Path
9	Insurance Type	Coming Method	LOC	Coming Method
10	LOC	ICU	Injury_Exp	Main_Treat_Dep
11	Coming Path	Admission Channel	BMI	Admission Month
12	Age	LOC	Admission Month	BMI
13	Admission Channel	Coming Path	Coming Method	Admission Channel
14	Admission Month	Admission Month	Admission Channel	Injury_Exp
15	Injury_Exp	BMI	Coming Path	Gender
16	BMI	Injury_Exp	Gender	Min(C)_Min(O)
17	Gender	Gender	Min(C)_Min(O)	Age
18	Min(C)_Min(O)	Min(C)_Min(O)	-	-

3. Results and Conclusion

3.1. Results of LOS prediction model

Three statistical techniques were used to analyze the data sets of traffic accidents, burns and falls with 18, 17 and 17 variables, respectively, with the accuracy of prediction for the entire data set and the accuracy of prediction for each injury. [Figure 1] showed the prediction accuracy obtained from the stepwise backward removal in each data set. As input variables decrease, the prediction accuracy is the highest when the most influential 11, 3, 6 and 3 functions are used in traffic accidents, burns and falls.

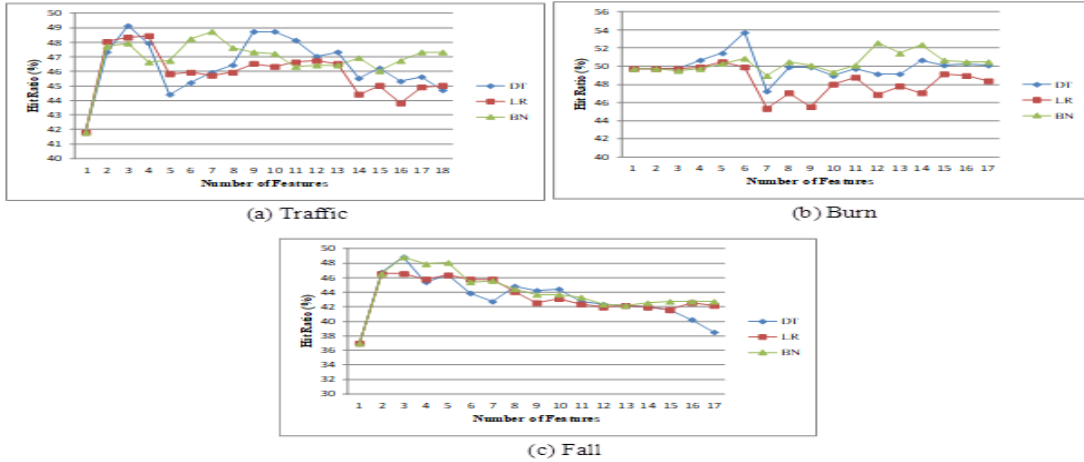


Figure 1. Prediction accuracy of LOS

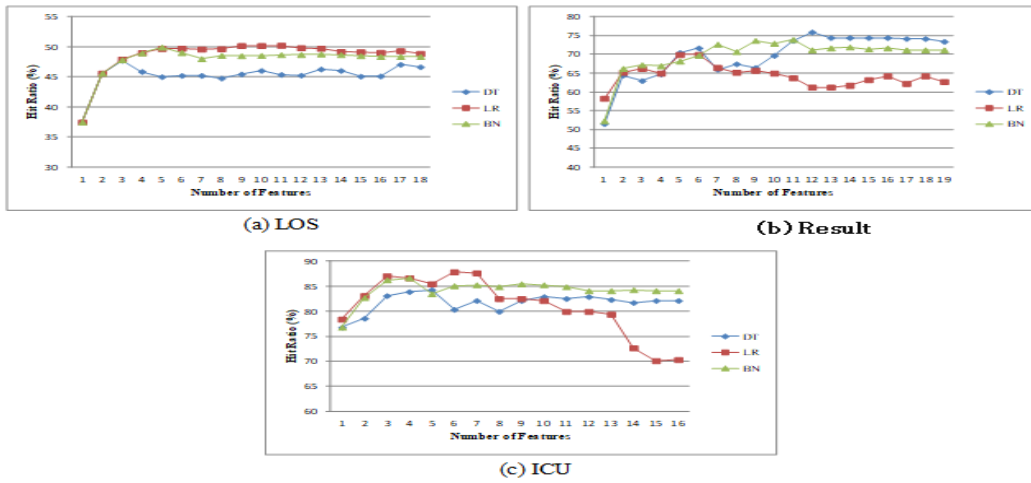


Figure 2. Prediction accuracy of total dataset

In the analysis of the total data set of the injured inpatients, logistic regression analysis (LR) showed optimal prediction accuracy [Table 3], [Figure 2]. In traffic accidents and burns, decision trees showed optimal prediction accuracy, and Fall showed optimal prediction accuracy similar to decision trees and beige eye net. Result and ICU showed high predictive power by DT method.

Table 3. Confusion matrix (LOS)

Total (LR)			
Predicted Actual	a	b	c
a	499	173	140
b	227	248	337
c	111	225	476
Traffic (DT)			
Predicted Actual	a	b	c

a	150	157	24
b	104	176	51
c	55	114	162
Burn (DT)			
Predicted Actual	a	b	c
a	74	91	10
b	40	108	27
c	22	53	100
Fall (DT)			
Predicted Actual	a	b	c
a	92	48	34
b	46	77	51
c	36	52	86
Fall (BN)			
Predicted Actual	a	b	c
a	91	52	31
b	51	66	57
c	34	42	98

3.2. Conclusion

LOS was best predicted by LR method in the total data set of injured inpatients, and Result and ICU were highly predicted by DT method. In addition, the results of LOS prediction by damage foreigners showed that DT had excellent predictive power in TA and Burn, and DT and BN method had high predictive power in Fall. This shows that DT is the best predictor of the independent variables related to injury when LOS, Results, and ICU are the dependent variables of the injured inpatients.

References

- [1] J. G. Moon, "Factor analysis on influence of length of stay," *Journal of the Korean Hospital Association*, vol.25 no.6, pp.20-29, (1996)
- [2] Kathleen Carey, "Hospital length of stay and cost: A multilevel modeling analysis," *Health Services and Outcomes Research Methodology*, vol.3 no.1, pp. 41-56, (2002) [DOI: 10.1023/a:1021530924455]
- [3] K. J. Lee, J. Y. Kim, K. H. Lee, G. J. Suh, and Y. K. Youn, "General scheme for the level I trauma center in South Korea," *Journal Korean Trauma*, vol.18, no.1, pp.1-16, (2005)
- [4] OECD, *OECD Health Date 2010: Statistics and Indicators for 30 Countries*. Paris, Organization for Economic Co-operation and Development, (2010)
- [5] S. K. Cha and S. S. Kim, "The determinant of the length of stay in hospital for schizophrenic patients: using data from the in-depth injury patient surveillance system," *The Journal of digital policy & management*, vol.11, no.4, pp.351-359, (2013)
- [6] S. Shafi, S. Barnes, D. Nicewander, D. Ballard, A. B. Nathens, A. M. Ingraham, M. Hemmila, S. Goble, M. Neal, M. Pasquale, J. J. Fildes, and L. M. Gentilello, "Care reform at trauma centers-mortality, complications, and length of stay," *The Journal of Trauma: Injury, Infection, and Critical Care*, vol.69, no.6, pp.1367-1371, (2010) DOI: 10.1097/ta.0b013e3181fb785d
- [7] Y. H. Cho, "Demand for long-term hospitalized patients in nursing homes and Pay-Pre-Care centers, graduate school," *Inje University in Korea*, (1998)

[8] WHO, The Injury chart book, (2011)

Authors



Young-Hee Nam

Associate Professor, Namseoul University, Chungnam, Korea.

This page is empty by intention.