

Short Communication

Predicting Length of Stay in Cardiovascular Patients Using Count Regression and Machine Learning Approaches

Samaneh Sabouri^{1,2*} , Maryam Salari^{1,2} 

1. Department of Biostatistics, School of Health, Mashhad University of Medical Sciences, Mashhad, Iran.

2. Social Determinants of Health Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.

**Citation** Sabouri S, Salari M. Predicting Length of Stay in Cardiovascular Patients Using Count Regression and Machine Learning Approaches. *Journal of Research & Health*. 2026; 16(2):195-202. <http://dx.doi.org/10.32598/JRH.16.2.2598.1> <http://dx.doi.org/10.32598/JRH.16.2.2598.1>**ABSTRACT**

Cardiovascular diseases often involve hospital readmissions and prolonged length of stay (LOS), that is a key measure for evaluating efficiency and quality of hospital care. This cross-sectional study aimed to predict and identify factors associated with LOS using various count regression and machine learning (ML) approaches. A total of 12,752 patients with cardiovascular disease were included. They were admitted to 43 hospitals affiliated with Mashhad University of Medical Sciences between March 2020 and February 2021. The mean patient age was 62.7 ± 13.6 years, with an average LOS of 4.4 ± 5.7 days and a median of 3 days (IQR: 3). Features including age, sex, diagnosis type, and percutaneous coronary intervention (PCI) were used to predict LOS. The predictive performance of the models was evaluated using mean absolute error (MAE). Although zero-truncated negative binomial regression (ZTNBR) outperformed other count regression models, support vector machine (SVM) achieved the lowest MAE (2.6), compared to artificial neural network (ANN) (MAE: 2.9) and the ZTNBR (MAE: 3.0). The type of cardiovascular disease was the most significant predictor of LOS, followed by PCI, age, and sex. These findings highlight the importance of patient-specific factors in designing healthcare strategies to optimize hospital resource allocation and improve patient outcomes.

Keywords: Cardiovascular disease, Machine learning (ML), Count data, Overdispersion, Length of stay (LOS)

Article info:

Received: 27 Dec 2024

Accepted: 25 Jun 2025

Publish: 01 Mar 2026

*** Corresponding Author:**

Samaneh Sabouri, Assistant Professor.

Address: Department of Biostatistics, School of Health, Mashhad University of Medical Sciences, Mashhad, Iran

Phone: +98 (51) 31892707

E-mail: sabourism@mums.ac.ir

Copyright © 2026 The Author(s);

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-BY-NC: <https://creativecommons.org/licenses/by-nc/4.0/legalcode.en>), which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Introduction

Cardiovascular diseases remain the leading cause of death worldwide, affecting both developing and developed countries, with over three-quarters of fatalities occurring in low- and middle-income regions [1-3]. In Iran, the adoption of a Western lifestyle, combined with advancements in healthcare services and increased life expectancy, has contributed to a rise in non-communicable diseases, including cardiovascular diseases [1].

Heart failure is often associated with frequent hospital readmissions and extended length of stay (LOS) [4]. LOS is a critical indicator of resource utilization and quality of care because shorter hospital stays can reduce medical costs and overall social expenses. Although earlier approaches primarily linked LOS to specific medical conditions, it is now recognized that multiple factors, such as demographic features, treatment complexity, and social circumstances, affect hospitalization duration. Accurate LOS prediction enables hospitals to minimize unnecessary delays, which is crucial since patients generally prefer home recovery. In addition, prolonged hospital stays increase the risk of complications like infections. The COVID-19 pandemic has further highlighted the significance of hospital bed capacity, emphasizing the need for reliable models to predict LOS based on admission data. Consequently, precise LOS prediction is essential for managing healthcare costs, optimizing hospital capacity, and improving service efficiency [5-7].

Traditional statistical models face challenges when applied to response variables, such as LOS. Although widely used, ordinary least squares regression is unsuitable for positively skewed data because it violates the assumption of normality in the response variable. This often results in biased coefficient estimates and inaccurate standard errors, leading to unreliable conclusions [7, 8]. Poisson regression (PR), commonly used for count data, assumes that the mean and variance are equal. This condition is rarely met in real-world data. Overdispersion, where the variance exceeds the mean, can invalidate PR results if unaddressed [7, 9]. To tackle this problem, it is recommended to use negative binomial regression (NBR) as an alternative [10]. Additionally, zero-truncated PR (ZTPR) and zero-truncated negative binomial regression (ZTNBR) provide additional improvements when count data do not include zero values [11, 12]. The accuracy of the results would be significantly enhanced by employing a model that fits the data well.

The issue of overdispersion can also be addressed using modern data science techniques. While count regressions are commonly applied to model LOS, machine learning (ML) methods have gained considerable attention for their predictive capabilities [5, 8]. ML algorithms learn from data to make accurate predictions [13] without relying on predefined distributional assumptions or strict relationships between variables [14, 15]. In this study, support vector machine (SVM) and artificial neural network (ANN) models were utilized due to their effectiveness in prediction and classification tasks [8, 14, 16]. ANN is inspired by the human brain and consists of interconnected nodes that process data and recognize patterns through learning, making it ideal for handling complex, non-linear relationships, particularly in medical research. SVM, a supervised learning algorithm, is widely used for classification and regression tasks, focusing on identifying the optimal hyperplane that separates different data classes [17].

Given the high prevalence of cardiovascular diseases, this study aimed to predict LOS and identify key factors affecting hospitalization duration in northeastern Iran. Various count regression models (PR, NBR, ZTPR, and ZTNBR) and ML approaches (ANN and SVM) were applied to address data overdispersion and improve LOS prediction accuracy.

Methods

This cross-sectional study assessed LOS in patients diagnosed with cardiovascular disease. Patients were admitted to 43 hospitals affiliated with [Mashhad University of Medical Sciences](#) in northeastern Iran between March 20, 2020, and February 14, 2021. Data were retrospectively extracted from the SEPAS Electronic Health Record system. As anonymized data were used, informed consent was not required.

Variables

Patients aged 18 years and older with a diagnosis of cardiovascular disease were included. The exclusion criteria comprised patients with LOS < 1 day, those who were voluntarily discharged, or transferred to other medical centers. LOS was defined as the number of days from hospital admission to discharge and regarded as the dependent variable.

Features, such as age (years), sex (male/female), percutaneous coronary intervention (PCI) status (yes/no), re-admission (yes/no), and diagnosis, were extracted from

the dataset. The diagnosis variable represents the reasons for a patient's hospital admission.

This variable was categorized based on the international classification of diseases (ICD) codes, with I-codes specifically for heart diseases [18, 19]. It was divided into five categories: Cerebrovascular diseases (CVD), cardiopulmonary and pulmonary circulation diseases (CPD and PCD), ischemic heart disease (ISD), hypertension diseases (HTND), and others (including cardiomyopathy, atrioventricular and left bundle branch block, cardiac arrest, conduction disorders, paroxysmal tachycardia, heart failure, and ill-defined heart disease descriptions).

Statistical analyses

Data preprocessing was performed to minimize the impact of noise on the results. This included the removal of duplicate values and a quality check to ensure data consistency and integrity. The dataset was complete for all investigated variables, and duplicate entries were removed. Quantitative variables were summarized using Mean \pm SD and median (interquartile range [IQR]), while qualitative variables were described by frequency and percentage. The Kolmogorov–Smirnov test confirmed a non-normal distribution of LOS due to its positive skewness. The Mann–Whitney and Kruskal–Wallis tests were applied to compare LOS across categories of independent variables. Variables with a $P < 0.05$ were selected for model fitting. In this study, various count regression models and ML methods were fitted. Data were analysed using R software, version 4.4.0. The VGAM package was used for count regressions, while the `radant` model and `e1071` packages were employed for ANN and SVM, respectively.

Count regressions

LOS was treated as the target outcome variable, and it was rounded to the nearest whole number. The initial model applied was PR, and the deviance statistics is a jargon here and is not plural. By dividing the deviance residual by its degrees of freedom, a quantity can be obtained to check the adequacy of the model. The closer to one, the assumption of the PR would be met [11]. Overdispersion was further tested using the ARE package in R.

Additional count regression models were fitted to the data, including NBR, ZTPR, and ZTNBR. Model performance was compared using the Akaike information criterion (AIC) and Bayesian information criterion

(BIC) [7, 11]. The model with the lowest AIC and BIC values was selected as the best one for predicting LOS in cardiovascular disease patients.

ML methods

The dataset was divided into training (70%) and testing (30%) sets using simple random sampling. The training set was used to develop the ANN and SVM models, while the testing set was employed to evaluate their predictive performance. A three-layer feed-forward back-propagation ANN was applied with 11 nodes (explanatory variables) in the input layer, 6 to 13 nodes in the hidden layer, and one node for the LOS in the output layer. It is recommended that the number of nodes in the hidden layer should be between 70% and 90% of the input layer's size, ensuring it does not exceed twice the number of nodes in the input layer [14]. Additionally, an SVM model with a Gaussian radial basis kernel was used for prediction. During the training process, a 10-fold cross-validation was employed to determine the optimal hyperparameters and also the number of hidden layer nodes in the ANN, reducing the risk of overfitting. The best architecture for the ANN was selected with 9 nodes in the hidden layer and a decay of 0.1. The SVM model with a cost parameter of one was chosen as the best one.

Assessing the models' performance

Given the extreme skewness of the LOS distribution, criteria, such as mean squared error do not perform well due to their sensitivity to outliers. The error increases rapidly when a long LOS is predicted to be short or vice versa [14]. Instead, mean absolute error (MAE) measures the mean absolute difference between the predicted and observed values ($MAE = \sum_{i=1}^n |y_i - \hat{y}_i| / n$). Therefore, it does not have this limitation when dealing with dispersed data. In the current study, MAE was used to evaluate the predictive performance of the models. The one with the lowest MAE was selected as the most suitable model for modeling dispersed count data.

Results

A total of 12,752 patients were included in the study. The Mean \pm SD age was 62.7 \pm 13.6 years. The mean LOS was 4.4 \pm 5.7 days, while the median was 3 days (IQR: 3). Approximately 9% of patients died before discharge. LOS was significantly longer ($P < 0.001$) for patients who died (8.7 \pm 10.5 days) compared to those who survived (4.0 \pm 4.9 days). Angiography was performed on 69.6% of patients, and 18.3% underwent PCI. Table 1 presents

Table 1. Characteristics of patients admitted with cardiovascular diseases in the study sample

Variables	No. (%)	LOS		P	
		Mean±SD	Median (IQR)		
Sex	Male	6848(53.7)	4.5±5.9	2.8 (3.5)	<0.001*
	Female	5904(46.3)	4.3±5.5	2.6 (3.1)	
Diagnosis	CVD	2226(17.5)	8.1±10.2	4.8 (6.4)	<0.001**
	CPD and PCD	133(1.0)	7.4±6.5	5.6 (6.4)	
	ISD	7691(60.3)	3.4±3.1	2.2 (2.4)	
	HTND	968(7.6)	3.1±3.9	2.1 (2.2)	
	Other	1734(13.6)	4.7±5.4	3.0 (3.7)	
Readmission	Yes	1126(8.8)	4.3±4.6	2.8 (4.3)	0.437*
	No	11626(91.2)	4.4±5.9	2.7 (3.3)	
PCI	Yes	2336(18.3)	2.4±2.1	1.8 (1.6)	<0.001*
	No	10416(81.7)	4.9±6.2	2.9 (3.9)	



Abbreviations: LOS: Length of stay; ISD: Ischemic heart disease; HTND: Hypertension diseases; CPD and PCD: Cardiopulmonary diseases and pulmonary circulation diseases; CVD: Cerebrovascular diseases; PCI: Percutaneous coronary intervention; IQR: Interquartile range.

*Mann-Whitney test, **Kruskal-Wallis test.

the descriptive statistics for patient characteristics. LOS was compared across different patient groups and variables. Significant features ($P < 0.05$) were selected for model fitting.

ISD was the most common reason for admission, with 2.6% mortality before discharge. The highest percentage of deaths was observed in CVD patients (22.4%). Patients with CPD and PCD, HTND, and other cardiovascular conditions accounted for 19.5%, 9.5%, and 17.2% of deaths, respectively.

A standard PR was initially applied to assess overdispersion. The overdispersion test was significant ($P < 0.001$), with deviance and dispersion values of 3.3 and 5.1, respectively. Alternative count regression models, including NBR, ZTPR, and ZTNBR, were fitted to predict LOS. The result showed that ZTNBR and NBR (Table 2).

All variables were significantly associated with LOS in the ZTNBR model. A 10-year increase in age was associated with a 4% increase in LOS. Men had a 14% longer LOS compared to Women. Patients who underwent PCI had shorter LOS. The diagnosis was also a significant

predictor; patients diagnosed with ISD, HTND, and other heart diseases had shorter LOS than those with CVD.

Model performance was assessed using MAE on the test dataset. The SVM model achieved the lowest MAE, outperforming ANN and ZTNBR (Table 3). In the SVM model, the most important variables were diagnosis, PCI, age, and sex, with weights of 1.07, 1.03, 1.01, and 1.00, respectively. The ANN model identified the same variables in the same order of importance: Diagnosis (0.11), PCI (0.03), age (0.02), and sex (0.01).

Discussion

LOS is an indicator of healthcare system efficiency and quality, and researchers are actively exploring different methods to improve it [7]. This study aimed to predict LOS and identify factors related to the duration of hospitalization using a large dataset on patients with cardiovascular diseases. This study revealed that LOS was highly overdispersed and exhibited significant right skewness. Among various count regression and ML models, SVM selected as the best model handling overdispersed data.

Table 2. Summary of count regression models

Variables	Exp (β) (95% CI)				
	PR	NBR	ZTPR	ZTNBR	
Age (y)	1.001 (1.001, 1.002) [†]	1.003 (1.001, 1.004) [†]	1.001 (1.000, 1.002) [†]	1.004 (1.002, 1.005) [†]	
Sex	Male	1.088 (1.070,1.107) [†]	1.109 (1.078,1.142) [†]	1.095 (1.076,1.115) [†]	1.157 (1.110,1.206) [†]
	Female	-	-	-	-
Diagnosis	ISD	0.473 (0.463,0.482) [†]	0.472 (0.454,0.490) [†]	0.461 (0.452,0.471) [†]	0.395 (0.374,0.417) [†]
	HTND	0.409 (0.394,0.425) [†]	0.406 (0.382,0.433) [†]	0.391 (0.375,0.407) [†]	0.316 (0.288,0.346) [†]
	CPD and PCD	0.933 (0.876,0.995) [*]	0.939 (0.824,1.071)	0.934 (0.876,0.995) [*]	0.935 (0.773,1.133)
	Other	0.584 (0.569,0.600) [†]	0.581 (0.553,0.610) [†]	0.579 (0.564,0.595) [†]	0.518 (0.483,0.556) [†]
PCI	CVD	-	-	-	-
	Yes	0.608 (0.591,0.626) [†]	0.610 (0.584,0.636) [†]	0.543 (0.525,0.563) [†]	0.473 (0.445,0.503) [†]
	No	-	-	-	-
AIC	80143.2	62087.4	79236.9	57733.4	
BIC	80202.8	62154.5	79296.6	57800.5	



Abbreviations: ISD: Ischemic heart disease; HTND: Hypertension diseases; CPD and PCD: Cardiopulmonary diseases and pulmonary circulation diseases; CVD: Cerebrovascular diseases; PCI: Percutaneous coronary intervention; PR: Poisson regression; NBR: Negative binomial regression; ZTPR: Zero-truncated Poisson regression; ZTNBR: Zero-truncated negative binomial regression; AIC: Akaike information criterion; BIC: Bayesian information criterion.

*P<0.05, †P<0.001.

Hachesu et al. examined various ML classification techniques for analyzing LOS in cardiac patients and identified SVM as the most accurate model [20]. Similarly, count regressions and ML methods were used to model ecological count data in a previous study. The result showed that ANN underperformed data, while SVM was recommended as the most suitable model [14]. These findings align with our results, which demonstrated the superior performance of SVM.

After selecting the optimal model, the most important variable in predicting LOS was identified. Diagnosis emerged as the key predictor for LOS in patients with cardiovascular diseases. A previous study reported that patients with ISD experienced longer LOS compared to those with hypertensive heart disease [4]. In our study, ISD was the most common reason for admission, while patients with CVD had the longest LOS compared to other diagnostic groups.

Table 3. Evaluation of predictive performance of the models

Model	MAE
SVM	2.6
ANN	2.9
ZTNB	3.0



Abbreviations: MAE: Mean absolute error; ANN: Artificial neural network; SVM: Support vector machine; ZTNBR: Zero-truncated negative binomial regression.

We found that PCI also played a significant role in predicting LOS. As a minimally invasive procedure, PCI often results in faster recovery and shorter hospital stays. If stable, patients undergoing successful PCI are typically discharged within 48 to 72 hours [8]. Our results corroborate this, showing a reduced LOS among patients who underwent PCI, consistent with prior research.

The influence of age and sex on LOS was also evaluated. Bender et al. hypothesized that women experience longer LOS due to older age and greater disease severity, but found no significant association between sex and LOS in univariate analysis, whereas age remained a significant predictor in the adjusted model [21]. The SVM analysis indicated that age was a more important predictor than sex, supporting existing evidence that older patients tend to have longer hospital stays [22, 23].

This study's strengths include using a large multi-center registry from Khorasan Razavi Province, Iran, which enhanced the generalizability of the findings. However, several limitations should be acknowledged. The study lacked detailed information on hospital-specific factors, such as variations in healthcare quality, staffing expertise, and resource availability, which could influence LOS. Additionally, our data did not include laboratory results, comorbidities, or socio-demographic characteristics, limiting the ability to control for these potentially confounding factors. Other variables, such as hospital bed availability, rehabilitation services, and the nature of the admission (elective vs. urgent), may also impact LOS but were not accounted for in this analysis [24, 25].

Conclusion

Diagnosis emerged as the most important factor in predicting LOS. These findings offer valuable insights for enhancing hospital services and optimizing resource management. This study also highlighted the importance of selecting an appropriate model for data analysis because it directly influences the accuracy of the results. ML methods do not rely on distributional assumptions and can accurately predict outcomes by capturing complex relationships between variables. In this study, SVM demonstrated the highest predictive performance among the models analyzed, particularly in handling overdispersed data.

Ethical Considerations

Compliance with ethical guidelines

This study was approved by the Ethics Committee of [Mashhad University of Medical Sciences](#), Mashhad,

Iran (Code: IR.MUMS.FHMPM.REC.1402.214). As anonymized data were used, informed consent from patients was not required.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' contributions

Writing the original draft: Samaneh Sabouri; review and editing: Maryam Salari; Conceptualization, formal analysis, methodology and supervision: All authors.

Conflict of interest

The authors declared no conflicts of interest.

Acknowledgments

The authors thank the Vice-Chancellor for Research at [Mashhad University of Medical Sciences](#) and all individuals involved in data collection.

References

- [1] Baeradeh N, Ghoddusi Johari M, Moftakhar L, Rezaeianzadeh R, Hosseini SV, Rezaianzadeh A. The prevalence and predictors of cardiovascular diseases in Kherameh cohort study: A population-based study on 10,663 people in southern Iran. *BMC Cardiovascular Disorders*. 2022; 22(1):244. [DOI:10.1186/s12872-022-02683-w] [PMID]
- [2] Saki N, Karandish M, Cheraghian B, Heybar H, Hashemi SJ, Azhdari M. Prevalence of cardiovascular diseases and associated factors among adults from southwest Iran: Baseline data from Hoveyze Cohort Study. *BMC Cardiovascular Disorders*. 2022; 22(1):309. [DOI:10.1186/s12872-022-02746-y] [PMID]
- [3] WHO. Cardiovascular diseases (CVDs). Geneva: World Health Organization; 2021. [Link]
- [4] Tigabe Tekle M, Bekalu AF, Tefera YG. Length of hospital stay and associated factors among heart failure patients admitted to the University Hospital in Northwest Ethiopia. *PLoS One*. 2022; 17(7):e0270809. [DOI:10.1371/journal.pone.0270809] [PMID]
- [5] Jain R, Singh M, Rao AR, Garg R. Predicting hospital length of stay using machine learning on a large open health dataset. *BMC Health Services Research*. 2024; 24(1):860. [DOI:10.1186/s12913-024-11238-y] [PMID]
- [6] Stone K, Zwigelaar R, Jones P, Mac Parthaláin N. A systematic review of the prediction of hospital length of stay: Towards a unified framework. *PLOS Digital Health*. 2022; 1(4):e0000017. [DOI:10.1371/journal.pdig.0000017] [PMID]

- [7] Zeleke AJ, Moscato S, Miglio R, Chiari L. Length of stay analysis of COVID-19 hospitalizations using a count regression model and quantile regression: A study in Bologna, Italy. *International Journal of Environmental Research and Public Health*. 2022; 19(4):2224. [DOI:10.3390/ijerph19042224] [PMID]
- [8] Lan T, Liao YH, Zhang J, Yang ZP, Xu GS, Zhu L, et al. Mortality and readmission rates after heart failure: A systematic review and meta-analysis. *Therapeutics and Clinical Risk Management*. 2021; 17:1307-20. [DOI:10.2147/TCRM.S340587] [PMID]
- [9] Verburg IW, de Keizer NF, de Jonge E, Peek N. Comparison of regression methods for modeling intensive care length of stay. *PLoS One*. 2014; 9(10):e109684. [DOI:10.1371/journal.pone.0109684] [PMID]
- [10] Dormann CF. Overdispersion, and how to deal with it in R and JAGS. 2016. [Link]
- [11] Hossain Z, Akter R, Sultana N, Kabir E. Modelling zero-truncated overdispersed antenatal health care count data of women in Bangladesh. *PLoS One*. 2020; 15(1):e0227824. [DOI:10.1371/journal.pone.0227824] [PMID]
- [12] Worku G, Tadesse G, Arega A, Tesfaw D. Determinants of the number of children born in Ethiopia, evidenced from 2019 miniEDHS: Using zero-truncated count regression models. *Research Square*. 2022. [DOI:10.21203/rs.3.rs-1983782/v1]
- [13] Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. *Journal of Global Health*. 2018; 8(2):020303. [DOI:10.7189/jogh.08.020303] [PMID]
- [14] Sidumo B, Sonono E, Takaidza I. Count regression and machine learning techniques for zero-inflated overdispersed count data: Application to ecological data. *Annals of Data Science*. 2024; 11(3):803-17. [DOI:10.1007/s40745-023-00464-6]
- [15] Tiruneh SA, Vu TTT, Rolnik DL, Teede HJ, Enticott J. Machine learning algorithms versus classical regression models in pre-eclampsia prediction: A systematic review. *Current Hypertension Reports*. 2024; 26(7):309-23. [DOI:10.1007/s11906-024-01297-1] [PMID]
- [16] Sun W, Zhang P, Wang Z, Li D, Chen J. Prediction of cardiovascular diseases based on machine learning. *ASP Transactions on Internet of Things*. 2021; 1(1):30-5. [Link]
- [17] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: With applications in R. 2nd ed. New York: Springer; 2021. [DOI:10.1007/978-1-0716-1418-1]
- [18] Alkhalfan F, Gyftopoulos A, Chen YJ, Williams CH, Perry JA, Hong CC. Identifying genetic variants associated with the ICD10 (International Classification of Diseases10)-based diagnosis of cerebrovascular disease using a large-scale biomedical database. *PLoS One*. 2022; 17(8):e0273217. [DOI:10.1371/journal.pone.0273217] [PMID]
- [19] Porras JPR, Garrido FB. Algorithm for predicting the most frequent causes of mortality by analyzing age and gender variables. *Journal of Positive Psychology and Wellbeing*. 2021; 6:1419-29. [Link]
- [20] Hachesu PR, Ahmadi M, Alizadeh S, Sadoughi F. Use of data mining techniques to determine and predict length of stay of cardiac patients. *Healthcare Informatics Research*. 2013; 19(2):121-9. [DOI:10.4258/hir.2013.19.2.121] [PMID]
- [21] Bender U, Norris CM, Dreyer RP, Krumholz HM, Raparelli V, Pilote L. Impact of sex- and gender-related factors on length of stay following non-ST-segment-elevation myocardial infarction: A multicountry analysis. *Journal of the American Heart Association*. 2023; 12(15):e028553. [DOI:10.1161/JAHA.122.028553] [PMID]
- [22] Kang JH, Bae HJ, Choi YA, Lee SH, Shin HI. Length of hospital stay after stroke: A Korean nationwide study. *Annals of Rehabilitation Medicine*. 2016; 40(4):675-81. [DOI:10.5535/arm.2016.40.4.675] [PMID]
- [23] Pinaire J, Azé J, Bringay S, Cayla G, Landais P. Hospital burden of coronary artery disease: Trends of myocardial infarction and/or percutaneous coronary interventions in France 2009-2014. *PLoS One*. 2019; 14(5):e0215649. [DOI:10.1371/journal.pone.0215649] [PMID]
- [24] Walsh B, Smith S, Wren MA, Eighan J, Lyons S. The impact of inpatient bed capacity on length of stay. *The European Journal of Health Economics*. 2022; 23(3):499-510. [DOI:10.1007/s10198-021-01373-2] [PMID]
- [25] Harerimana G, Kim JW, Jang B. A deep attention model to forecast the Length of Stay and the in-hospital mortality right on admission from ICD codes and demographic data. *Journal of Biomedical Informatics*. 2021; 118:103778. [DOI:10.1016/j.jbi.2021.103778] [PMID]

This Page Intentionally Left Blank