Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 2
CII 2023

A Decision Tree Model-based Study on the Grouping of Hospitalization Expenses of Type 2 Diabetes Patients in a Traditional Chinese Medicine Hospital

Xue Zhang [1], Jianke Ren [2], Xin Wang [2], Lijun Liang [2, *]

[1] Graduate School of Tianjin University of Traditional Chinese Medicine, China

[2] School of Management, Tianjin University of Traditional Chinese Medicine, China

* Corresponding author Email: lljun216@126.com

Abstract. Objective To discover the case-mix approach for TCM treatment of dominant disease type 2 diabetes patients and to analyze the hospitalization expense standards for different combinations of patients to provide theoretical references for better participation of TCM hospitals in the DRG payment method as well as reasonable regulation of hospitalization expenses for TCM type 2 diabetes patients in the future. Methods Case data were collected from 1,171 patients discharged with a principal diagnosis of type 2 diabetes from September 2021 to April 2023 from a tertiary care Chinese medicine hospital in Tianjin. Univariate and multivariate stepwise regression analyses determined categorical node variables. The decision tree model was used to perform case combinations and to analyze the expense standards and weights of different groups of disease types. Results The number of days of hospitalization, conditions in admission, and whether complicated cardiovascular and cerebrovascular diseases are the main factors affecting the hospitalization expense of type 2 diabetes patients in Chinese medicine hospitals, and the decision tree model was used to group the patients to form nine groups of patient case combinations. The standard of hospitalization expense for each combination was set reasonably. Conclusion The grouping results are more satisfactory, and the grouping scheme and hospitalization expense standards can provide a scientific basis for DRG grouping, adopting differentiated expense management, and improving the medical resource efficiency in TCM hospitals.

Keywords: TCM Hospitalization Expenses; Type 2 Diabetes Mellitus; Decision Tree Model; DRG.

1. Introduction

Traditional Chinese medicine (TCM) emphasizes evidence-based treatment and has unique advantages and crucial clinical value in preventing and treating diabetes and its complications [1]. However, due to the unique nature of TCM services, lack of sufficient evidence-based evidence, and other reasons, TCM healthcare organizations may not implement the DRG payment method for the time being [2]. However, the reform of the DRG payment method is a general trend. At the same time, hospitalization expense is one of the essential factors closely watched by patients and healthcare-related departments. Hence, the study on analyzing hospitalization expenses for treating type 2 diabetes mellitus in TCM medical institutions is critical. The study takes type 2 diabetes patients in a tertiary care TCM hospital in Tianjin as an example, uses the decision tree model to classify type 2 diabetes patients, and formulates the hospitalization expense standards for different groups of patients, which provides theoretical references for the patients to understand the expense of treatment, the promotion of DRG payment method in TCM hospitals as well as for the related healthcare institutions to reasonably regulate the hospitalization expense of type 2 diabetes patients in TCM hospitals.

2. Data Source and Methods

2.1. Data Source

The case home page data of type 2 diabetes patients from September 2021 to April 2023 were collected from the case management system of a tertiary care TCM hospital in Tianjin. The inclusion

criteria were principal diagnosis of type 2 diabetes mellitus, discharge for improvement, removal of samples with incomplete information on the case home page, and unequal total hospitalization expenses and various expenses. A total of 1171 valid cases were obtained.

2.2. Data Pre-processing

Factors such as gender, age, conditions in admission, treatment category, and other factors were used as independent variables, and categorical assignments were made, as shown in Table 1. Hospitalization expense was used as the dependent variable, and by the K-S normality test, the hospitalization expense was non-normally distributed, which was processed by a logarithmic transformation [3], and by the normality test, it was approximated to be normally distributed. The logarithmic-transformed expense was used as the dependent variable for the data analysis.

Table 1. Variable assignment table

Variable name	Variable	Quantitative method
Hospitalization expense(yuan)	Y	LogY
Gender	X1	Male = 1, Female = 2
Age (years)	X2	≦45=1, 46-60=2, 61-75=3, ≧76=4
Conditions in admission	X3	Critical = 1, Emergency = 2, General = 3
Treatment category	X4	Chinese medicine = 1, Chinese and Western medicine = 2, Western medicine = 3
Hospitalization days	X5	≦3=1, 4-7=2, 8-14=3, ≧15=4
Whether complicated with cardiovascular and cerebrovascular diseases	X6	Yes = 1, No = 0
The number of comorbidities and complications	X7	≦4=1, 5-8=2, ≧9=3

2.3. Study Methods

SPSS 26.0 software was used for statistical analysis. The Mann-Whitney U or Kruskal-Wallis H test was first used for univariate analysis. The key factors affecting hospitalization expenses were extracted by multivariate step-by-step linear regression. Finally, the selected factors were used as classification nodes for case grouping using the decision tree CHAID method. Decision tree grouping evaluation indexes: the coefficient of variation (CV) was used to evaluate intra-group heterogeneity; the smaller the CV, the smaller the intra-group variation and the better the grouping effect. The Kruskal-Wallis H test was statistically significant, and the larger the reduction in variance (RIV) was, the more influential the between-group variation was, and the better the grouping effect was [4]. The difference was considered statistically significant with $P<0.05$.

3. Results

3.1. Analysis of Influencing Factors of Hospitalization Expenses

Comparison of hospitalization expenses among patients with different genders, ages, conditions in admission, treatment categories, hospitalization days, whether or not they have combined cardiovascular and cerebrovascular diseases, and number of comorbidities and complications showed statistically significant differences ($P<0.05$), as shown in Table 2. Multivariate stepwise linear regression was performed with log-transformed hospitalization expenses as the dependent variable and the factors statistically significant in the univariate analysis as the independent variables. The results showed that three factors, hospitalization days, conditions in admission, and the presence of

combined cardiovascular and cerebrovascular diseases, entered the model (F=416.614, P<0.001, R^2=0.516). All VIFs are less than 3, indicating no collinearity issue, as shown in Table 3.

Table 2. Results of a univariate analysis of factors influencing hospitalization costs in type 2 diabetes patients

Variable	Category	Sample size n (%)	Hospitalization expenses (yuan)		Z/χ^2	P
			Median	Quartile range		
Gender	Male	598 (51.07)	13094.34	6897.83	-2.034	0.042
	Female	573 (48.93)	13814.74	6432.24		
Age (years)	≦45	92 (7.86)	11071.07	5889.42	40.659	0.000
	46-60	277 (23.65)	12820.49	6134.93		
	61-75	578 (49.36)	13577.88	6398.11		
	≧76	224 (19.13)	15136.24	8204.63		
Conditions in admission	Critical	154 (13.15)	19213.90	12,571.85	163.525	0.000
	Emergency	183 (15.63)	16632.98	11093.16		
	General	834 (71.22)	12782.03	5230.33		
Treatment category	Chinese medicine	23 (1.96)	15668.47	10693.23	11.780	0.003
	Chinese and western medicine	1130 (96.33)	13420.70	6658.86		
	Western medicine	18 (1.54)	14514.49	12055.19		
Hospitalization days	≦3	50 (4.27)	6099.01	4057.09	399.086	0.000
	4-7	236 (20.15)	10096.47	4793.41		
	8-14	747 (63.79)	13829.78	5208.96		
	≧15	138 (11.78)	22939.78	13529.11		
Whether complicated with cardiovascular and cerebrovascular diseases	Yes	846 (72.25)	14495.08	7444.19	-11.445	0.000
	No	325 (27.75)	11223.20	5107.93		
The number of comorbidities and complications	≦4	92 (7.86)	10562.11	7343.17	68.640	0.000
	5-8	483 (41.24)	12738.01	6243.16		
	≧9	596 (50.90)	14432.19	6489.39		

3.2. Decision Tree Model Classification Results

The influencing factors extracted from the multiple stepwise regression analysis were used as the classification nodes of the CHAID decision tree. The decision tree was utilized to classify 1,171 type 2 diabetes patients, and the results showed a total of 9 terminal nodes. The CV value of each category was less than 1, indicating that the intragroup variation of hospitalization expenses of patients in each category was small; the hospitalization expenses of each category were tested by the Kruskal-Wallis H test, and the results showed that the intergroup differences of hospitalization expenses of patients in 9 categories were statistically significant (H=560.489, P<0.05), and the RIV value was 0.58, which showed an excellent heterogeneity between groups. Considering all the indicators, the classification scheme is more reasonable, as shown in Table 4.

Table 3. Results of multivariate stepwise linear regression analysis of hospitalization costs for type 2 diabetes patients

Variable	Non-standardized coefficient	Standard error	Standardized coefficient	t	P	Tolerance	VIF
Constant	3.776	0.026		146.861	< 0.001		
Hospitalization days	0.175	0.006	0.589	28.817	< 0.001	0.991	1.009
Conditions in admission	-0.076	0.006	-0.268	-12.689	< 0.001	0.926	1.079
Whether complicated with cardiovascular and cerebrovascular diseases	0.084	0.010	0.186	8.769	< 0.001	0.923	1.083

Table 4. Decision tree case classification scheme for type 2 diabetes patients in TCM hospitals

Groups	Case-mix classification	Sample size n (%)	Average (yuan)	Standard deviation (yuan)	CV
Group 1	Days of hospitalization≦3 days	50 (4.3)	7188.49	5123.77	0.71
Group 2	4-7 days of hospitalization No cardiovascular or cerebrovascular disease	81 (6.9)	8301.43	2283.90	0.28
Group 3	Days of hospitalization 4-7 days Conditions in admission(General) Combined cardiovascular and cerebrovascular disease	100 (8.5)	11170.19	4701.92	0.42
Group 4	Days of hospitalization 8-14 days Conditions in admission(General) No cardiovascular or cerebrovascular disease	181 (15.5)	12442.81	3078.40	0.25
Group 5	4-7 days of hospitalization Conditions in admission(Critical or emergency) Combined cardiovascular and cerebrovascular disease	55 (4.7)	14437.50	5285.12	0.37
Group 6	Days of hospitalization 8-14 days Conditions in admission(General) Combined cardiovascular and cerebrovascular disease	373 (31.9)	14480.16	4012.17	0.28
Group 7	Days of hospitalization 8-14 days Conditions in admission(Critical or emergency)	193 (16.5)	19239.05	8697.38	0.45
Group 8	Days of hospitalization ≧15 Conditions in admission(General)	74 (6.3)	20476.56	7649.41	0.37
Group 9	Days of hospitalization ≧15 days Conditions in admission(Critical or emergency)	64 (5.5)	32083.40	15460.05	0.48

3.3. Development of Relevant Standards for Hospitalization Expenses of Case Combinations

The median hospitalization expense was used as the standard expense for each case combination, and P75 plus 1.5 times the interquartile spacing for each group was used as the upper limit of

hospitalization expenses [5]. Excess expenses were defined as patient expenses exceeding the upper limit of hospitalization expenses. The weight of the disease category was the average expense per case/cost per capita for all cases in that category, with the larger value reflecting the greater consumption of healthcare resources by patients in that category [5]. In this study, the number of patients exceeding the limit was 40 cases, accounting for 3.42% of the total cases, and the proportion of total exceeding cost was 7.69%. Combination 9 had the highest disease weight, i.e., combinations with ≧15 hospitalization days and critical or emergency admission consumed the most medical resources, as shown in Table 5.

Table 5. Hospitalization expense standard of type 2 diabetes patients in different groups

Groups	Number of samples	Standard expenses (yuan)	P_{75}	IQR	Upper limit of cost(yuan)	The number of overrun cases	Proportion of out-of-limit patients (%)	Over-limit cost ratios (%)	Case weight
Group 1	50	6099.01	8518.85	4057.09	14604.49	3	6.00	19.73	0.47
Group 2	81	8219.40	9953.33	3188.09	14735.47	0	0	0	0.54
Group 3	100	10108.30	12655.50	3970.83	18611.75	4	4.00	10.11	0.73
Group 4	181	12339.05	14058.46	3548.57	19381.32	5	2.76	4.91	0.81
Group 5	55	12928.33	16253.42	5158.54	23991.23	4	7.27	13.39	0.95
Group 6	373	13731.25	16110.96	4206.55	22420.79	15	4.02	7.55	0.95
Group 7	193	17607.06	22969.55	9160.85	36710.83	3	1.55	4.86	1.26
Group 8	74	18105.69	24846.17	9185.97	38625.13	3	4.05	8.50	1.34
Group 9	64	29299.67	36645.25	13534.78	56947.42	3	4.69	12.53	2.10
Add up the total	1171	-	-	-	-	40	3.42	7.69	-

4. Discussions and Suggestion

In this study, 1171 patients were divided into nine groups using the decision tree CHAID algorithm as an example of Chinese medicine treatment for the dominant type 2 diabetes patient. The results showed that including influencing factors was reasonable, the heterogeneity within the group was small, the heterogeneity between the groups was significant, and the grouping effect was more satisfactory. The number of hospitalization days was the most critical factor affecting the hospitalization expense of type 2 diabetes patients in Chinese medicine hospitals, followed by admission status and whether or not there was a combination of cardiovascular and cerebrovascular diseases. Patients with three-day hospitalization had the lowest hospitalization expense and disease weight, and patients with ≧15 days of hospitalization and critical or emergency admission condition had the highest hospitalization expense and disease weight, and the combination consumed the most medical resources. The number of days of hospitalization has the most significant impact on the hospitalization expenses of type 2 diabetes patients. The longer the number of days of hospitalization, the higher the series of the costs incurred by the patient, such as material expenses, nursing expenses, bed expenses, etc., increase accordingly, and the hospitalization material expenses, nursing expenses, bed expenses, of the patient will inevitably increase. Hospitals should strengthen the clinical pathway management of type 2 diabetes to reduce the number of invalid hospitalization days. Secondly, the hospitalization material expenses, nursing expenses, and bed expenses of type 2 diabetes patients whose admission situation is critical or acute are higher, consistent with the research results [6]. Most patients ignore glycemic control due to insufficient knowledge and ideological attention to diabetes, which leads to various complications such as nephropathy and podiatry, resulting in severe conditions that necessitate multiple hospitalizations, surgical treatments, or prolonged hospitalization days. In addition, it is necessary to focus on whether type 2 diabetes patients are combined with cardiovascular and

cerebrovascular diseases, and the hospitalization material expenses, nursing expenses, and bed expenses of patients with combined cardiovascular and cerebrovascular diseases are relatively high, consistent with the results of existing studies [7]. The incidence of cardiovascular and cerebrovascular diseases in diabetic patients is 80% [8]. Diabetic patients should strictly control their blood glucose and blood pressure and go to the hospital for regular physical examinations to prevent the acceleration of disease progression.

In this study, referring to previous literature, the median hospitalization expense was used as the standard cost for each category of patients, and the upper-cost limit was determined by 75% of the hospitalization expense plus 1.5 times IQR. Due to special diagnostic and treatment needs, a few patients can exceed the upper limit of the hospitalization expense, and the overlimit percentage is generally controlled within 5% [4]; if the overlimit rate is too high, there may be overdiagnosis and treatment. The calculated overlimit percentage of patients in each group is within 5%, and the overall overlimit percentage is 3.42%, also within 5%. The hospitalization expense standard set by this method is reasonable. However, in-depth analysis of the grouping results revealed that there were two categories (Group 1 and Group 5) with an over-limit proportion more significant than 5%, and a higher over-limit proportion generally implies that the proportion of over-limit costs is also higher, especially Group 1 (hospitalization days ≦ three days) is the most obvious, which has an over-limit ratio of 6% and an over-limit cost proportion of 19.73%, suggesting that there is a small number of patients consuming a higher amount of healthcare resources within this category. The combination with a higher proportion of costs exceeding the upper limit should focus on strengthening control, standardizing diagnostic and treatment behaviors, reducing unnecessary examinations and operations, and controlling the unreasonable growth of costs.

Due to the uniqueness of Chinese medicine, previous studies have rarely examined hospitalization expenses in Chinese medicine hospitals, and the DRG payment method has yet to be implemented in Chinese medicine hospitals. In this study, type 2 diabetes patients hospitalized in TCM hospitals were classified from three aspects: the number of days of hospitalization, conditions in admission, and whether or not they were combined with cardiovascular and cerebrovascular diseases, and the hospitalization expense standards, cost ceilings, and disease weights were calculated for different categories of patients. The grouping of the patients was reasonable, and the development of the cost standards was accurate, which is significant to healthcare institutions' regulation of healthcare costs and the realization of the DRG payment method in TCM.

Acknowledgments

This work was supported by the National Natural Science Foundation of China Project, Project Number: 72104180.

References

[1] TaoYao, Jiangyan Xu, WenliLiu, et al. Modern medical interpretation of treating diabetes nephropathy from the spleen [J]. Chinese Journal of Traditional Chinese Medicine, 2023,38 (8): 3619-3623.

[2] XiaoyingZhu, KanTian, PengqianFang, et al. Discussion on Several Issues and Countermeasures of DRG Payment in Traditional Chinese Medicine Hospitals in China [J]. Chinese Hospitals, 2022, 26 (5): 10-12.

[3] XuezhiWu. Analysis of hospitalization costs for acute appendicitis patients based on DRGs [J]. China Health Statistics, 2020,37 (3): 465-467.

[4] YijingPeng, YanSun, RenqiangYu, et al. A grouped study on hospitalization costs of premature infants based on decision tree models [J]. China Health Economics, 2023,42 (4): 30-33.

[5] XiaolinYu. Case combination study of DRGs in patients with upper urinary tract stones based on CHAID decision tree [J]. China Health Statistics, 2022, 39 (6): 925-928.

[6] YongbingCheng, XinzhouFan, Li Hua, et al. Analysis of hospitalization expenses and influencing factors of type 2 diabetes in a county of Shaanxi Province from 2012 to 2017 [J]. China Public Health Management, 2019, 35 (3): 429-432.

[7] YawenZhang, XiaoningHe, JingWu. Research on incidence rate of chronic complications and related medical expenses of newly diagnosed type 2 diabetes patients in Tianjin [J]. China Health Statistics, 2021,38 (1): 47-50.

[8] YingtingZuo, ShoulingWu, ShuohuaChen, et al. The risk of cardiovascular and cerebrovascular diseases in different subtypes of hypertension in individuals with abnormal glucose metabolism [J]. Journal of Capital Medical University, 2023,44 (1): 42-48.