ORIGINAL RESEARCH

Optimization of Diagnosis-Related Groups for 14,246 Patients with Uterine Leiomyoma in a Single Center in Western China Using a Machine Learning Model

Yuan Ma [1,2], Li Li[1], Li Yu[1], Wei He[1], Ling Yi[1], Yuxin Tang[1], Jijie Li[1], Zhigang Zhong[3], Meixian Wang[4], Shiyao Huang[5,6], Yiquan Xiong[5,6], Pei Xiao[7], Yuxiang Huang[1]

[1]Department of Medical Record Management, West China Second University Hospital, Sichuan University, Chengdu, Sichuan, People's Republic of China; [2]Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University), Ministry of Education, Chengdu, Sichuan, People's Republic of China; [3]Department of Prevention, Office of Cancer Prevention and Treatment, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, Cancer Hospital Affiliate to University of Electronic Science and Technology of China, Chengdu, Sichuan, People's Republic of China; [4]National Center for Birth Defects Monitoring, West China Second University Hospital, Sichuan University, Chengdu, Sichuan, People's Republic of China; [5]Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, Chengdu, Sichuan, People's Republic of China; [6]NMPA Key Laboratory for Real World Data Research and Evaluation in Hainan, Chengdu, Sichuan, People's Republic of China; [7]Medical Insurance Office, West China Fourth Hospital, Sichuan University, Chengdu, Sichuan, People's Republic of China

Correspondence: Yuxiang Huang, Email 3593213104@qq.com

Background: Uterine leiomyoma (UL) is one of the most common benign tumors in women, and its incidence is gradually increasing in China. The clinical complications of UL have a negative impact on women's health, and the cost of treatment poses a significant burden on patients. Diagnosis-related groups (DRG) are internationally recognized as advanced healthcare payment management methods that can effectively reduce costs. However, there are variations in the design and grouping rules of DRG policies across different regions. Therefore, this study aims to analyze the factors influencing the hospitalization costs of patients with UL and optimize the design of DRG grouping schemes to provide insights for the development of localized DRG grouping policies.
Methods: The Mann–Whitney U-test or the Kruskal–Wallis H-test was employed for univariate analysis, and multiple stepwise linear regression analysis was utilized to identify the primary influencing factors of hospitalization costs for UL. Case combination classification was conducted using the exhaustive chi-square automatic interactive detection (E-CHAID) algorithm within a decision tree framework.
Results: Age, occupation, number of hospitalizations, type of medical insurance, Transfer to other departments, length of stay (LOS), type of UL, admission condition, comorbidities and complications, type of primary procedure, other types of surgical procedures, and discharge method had a significant impact on hospitalization costs (P<0.05). Among them, the type of primary procedure, other types of surgical procedures, and LOS were the main factors influencing hospitalization costs. By incorporating the type of primary procedure, other types of surgical procedures, and LOS into the decision tree model, patients were divided into 11 DRG combinations.
Conclusion: Hospitalization costs for UL are mainly related to the type of primary procedure, other types of surgical procedures, and LOS. The DRG case combinations of UL based on E-CHAID algorithm are scientific and reasonable.
Keywords: uterine leiomyoma, diagnosis-related groups, decision tree

Introduction

Uterine leiomyoma (UL), also known as uterine fibroids, is the most common benign pelvic tumor in women.[1,2] It originates from the smooth muscle cells in the myometrium and is characterized by the presence of an abundant extracellular matrix composed of collagen, fibronectin, and proteoglycans.[3,4] UL is most frequently diagnosed in women between the ages of 30 and 50, with the highest incidence observed in women aged 40 to 50, accounting for approximately 51.2% to 60.9% of cases.[5] Despite being a benign tumor, UL often presents with debilitating symptoms such as heavy menstrual bleeding, pelvic pressure, pain, and anemia.[6] Additionally, UL is associated with infertility,

recurrent miscarriage, and preterm birth.[7] It is worth noting that approximately 70% of uterine sarcomas can be traced back to UL, and it also contributes to a significant proportion of uterine cancer-related deaths.[8] Consequently, the clinical complications of UL have a profound negative impact on women's health.

The diagnosis of UL is established through four clinical methods.[6] First, through a pelvic examination, characterized by swelling, irregularity, firmness, and a non-tender uterus. Second, a transvaginal ultrasound revealing a uterine volume of less than 375mL and fewer than four fibroids, characterized by low echogenicity. Third, saline-infusion sonohysterography. Fourth, Magnetic Resonance Imaging (MRI), which is the most accurate method for mapping and quantifying fibroids. Uterine sarcoma (US) is considered in the differential diagnosis of UL. US represents rare and aggressive uterine tumors. Originating from the endometrial connective tissue or myometrium, it accounts for approximately 3% of all uterine tumors.[9,10] Studies indicate that leiomyosarcoma is the most common form of US.[11] The distinction between leiomyoma and leiomyosarcoma relies on conventional morphological criteria, including mitosis, atypia, and necrosis.[12] In some instances, differentiation may also be based on hormonal status or other treatment modalities.

UL is currently the most common indication for hysterectomy worldwide. In Canada, approximately 30% of women undergo hysterectomy for uterine fibroids, making it the second most common surgical procedure after cesarean section.[13] In the United States, this rate increases to 40%, with nearly 600,000 patients undergoing hysterectomy each year.[14,15] Hysterectomy is associated with significant morbidity, mortality, and has a substantial economic impact on healthcare systems.[16] The estimated annual cost of surgical treatment for UL in the United States alone is as high as $2.2 billion.[17] In China, the incidence of UL has been on the rise, increasing from 3.4% in the 1970s to 5.2% in 2012.[18] A study conducted by Gu et al revealed that over 100,000 Chinese women undergo hysterectomy due to UL annually.[19] Consequently, the economic burden of UL on Chinese society is considerable, and it also significantly affects the quality of life of affected women. To alleviate the burden of this disease on patients, it is imperative for the government to implement measures aimed at controlling medical costs.

Diagnosis-related groups (DRG) are widely recognized as one of the most advanced methods of medical payment management.[20] Numerous studies have shown that it effectively controls medical costs and reduces the financial burden on patients.[21,22] DRG was initially developed at Yale University and was first implemented in the United States in 1983.[23] It is a payment system that categorizes diseases with similar clinical symptoms and resource utilization into specific groups.[24] The grouping results serve as the foundation for medical institutions to bill patients and for medical insurance institutions to establish reimbursement standards.[25] In this system, medical institutions will not receive payment if they provide excessive treatment services, ensuring medical quality and controlling medical costs.[26] However, there are variations in DRG grouping rules across different countries and regions. For instance, in Ireland, appendectomy cases are divided into two DRG groups based on the complexity of the procedure and the age of patients, whereas Germany has 11 DRG groups for such cases.[27] Austria only considers age as a distinguishing factor when grouping appendectomy cases, whereas France incorporates age, comorbidities, complications, length of stay (LOS), and discharge method to establish grouping rules.[27] Furthermore, unlike Finland and Sweden, Poland, the United Kingdom, and other countries utilize LOS as a parameter for DRG grouping.[28]

China has encountered challenges similar to those faced by other countries in implementing DRG. In 2009, the Chinese government initiated a prospective payment reform based on DRG.[29] Through application in various regions, standards such as BJ-DRG, CN-DRG, CR-DRG, and C-DRG have steadily become more prevalent. In an effort to standardize the use of grouping standards in pilot cities, the National Health Care Security Administration introduced the China Health Safety Diagnosis-Related Group (CHS-DRG) in 2019.[30] Nevertheless, due to disparities in regional economic development, the Chinese government permits each region to develop localized grouping rules that reflect their specific circumstances.[31] As a result, there are noticeable variations in the design of DRG policies and grouping rules across different regions of China.

Reasonable DRG grouping is essential for scientific cost payment, facilitating the control of medical costs and the monitoring of unreasonable medical expenses. Most previous studies have gathered research samples from China's developed cities, with limited focus on western cities. For instance, Luo et al analyzed elderly cataract patients using multi-center hospitalization data from Shanghai.[32] Similarly, Wu et al investigated colorectal cancer inpatients at a single center in Beijing.[33] Zhi et al examined acute stroke inpatients across multiple centers in Jinhua, Zhejiang Province.[34] Consequently, we collected data on discharged patients with UL from a single center in a western Chinese city. Initially,

univariate analysis was employed to identify the influencing factors of hospitalization costs for UL patients. Subsequently, multivariate stepwise linear regression analysis screened the predictive factors for the machine learning model. Finally, based on the predictive factors, a decision tree model for predicting UL hospitalization costs was constructed, and an attempt was made to establish a UL case grouping scheme aligned with the region's actual circumstances. This study holds significant practical importance, offering insights for optimizing UL DRG grouping management and cost payment in medical security departments across China's western cities. Furthermore, devising more practical grouping rules can effectively mitigate unreasonable medical expenses and reduce patient burden.

Materials and Methods
Data Source and Processing
Data for this study were obtained from the medical record information management system of a single center in a city in western China. This center, a university-affiliated hospital, serves as a national regional medical center, integrating medical care, education, research, preventive healthcare, and talent development. The study encompassed patients discharged with UL between 2019 and 2022. Inclusion criteria included: a primary diagnosis code, with the medical record's first three digits being D25 (UL), according to the International Classification of Diseases, Tenth Revision (ICD-10); a discharge date between January 1, 2019, and December 31, 2022. Exclusion criteria included: significant missing information in the medical record's primary data; obvious logical errors in the medical record's primary data; a length of stay exceeding 60 days. Ultimately, 14,246 eligible cases were included. Detailed information collected from the medical record's first page included age, occupation, admission and discharge dates, admitting and discharging departments, diagnostic and surgical information, cost details, medical insurance, and discharge method.

Decision Tree Model
This study utilized the Exhaustive Chi-squared Automatic Interaction Detector (E-CHAID) algorithm for decision tree analysis in case combination classification. The E-CHAID algorithm represents an enhancement of the CHAID algorithm. The E-CHAID algorithm's combination principle entails optimally segmenting the research sample according to the target variable and chosen predictive variables, utilizing the chi-squared test's significance level for automatic grouping in a multivariate contingency table.[35] A significant advantage of the E-CHAID algorithm is its ability to process non-linear data and accommodate missing values, thus surmounting the constraints of traditional parametric testing methods. Furthermore, the E-CHAID algorithm facilitates more comprehensive variable merging and grouping compared to the CHAID algorithm, enabling more precise selection of grouping variables.[36]

Parameter Setting
In the practical implementation of the E-CHAID algorithm, certain parameters were established. The maximum number of tree layers was configured to 3, the minimum sample size for parent nodes was designated as 100, and minimum sample size for child nodes was designated as 50. A 10-fold cross-validation model was employed to evaluate the performance of the decision tree. The significance level for tree merging and splitting tests was established at $\alpha=0.05$, with a P value less than 0.05 denoting statistically significant differences at each classification node.

Evaluation Index
The effectiveness of the case combination was evaluated using the Kruskal–Wallis H non-parametric test, coefficient of variation (CV), reduction in variation (RIV), and analytic proportion. A CV value lower than 1 signifies a high homogeneity of inpatient costs within the DRG group. A higher RIV value, coupled with significant cost differences among DRG groups as identified by the Kruskal–Wallis H non-parametric test, indicates greater heterogeneity of inpatient costs between DRG groups. A higher analytic proportion suggests an improved explanatory capability of standard costs on actual costs within the DRG group.

Analysis Setting

Owing to the skewed distribution of inpatient costs, the median inpatient costs served as the reference value for standard costs within the DRG group. Furthermore, the upper limit reference value for inpatient costs within the DRG group was set at the 75th percentile plus 1.5 times the interquartile range. Patients whose costs exceeded the upper limit reference value were classified as outliers. The weight of each DRG group was determined by calculating the ratio of its average cost to the average cost of all cases. This reflects the degree of medical resource consumption.

Statistical Analysis

Data cleaning and the establishment of a UL database were conducted using Microsoft Excel 2016 software (Microsoft Corporation, Redmond, Washington, United States), while data analysis was carried out using SPSS 21.0 software (SPSS Inc., Chicago, IL, United States). Owing to the skewed distribution of inpatient costs, non-parametric tests were employed to identify factors influencing inpatient costs.[37] Specifically, the Mann–Whitney *U*-test was applied to binary variables, including the number of hospitalizations and transfer to other departments, whereas the Kruskal–Wallis *H*-test was utilized for multi-category variables, comprising age, occupation, type of medical insurance, source of patient, admission pathway, LOS, type of uterine leiomyoma, admission condition, comorbidities and complications, type of primary procedure, other types of surgical procedures, and discharge method. Following the identification of statistically significant factors influencing inpatient costs via univariate analysis, these factors served as independent variables, with the logarithm of inpatient costs as the dependent variable, in a multiple stepwise linear regression analysis for UL cases. This was done to select predictive factors for the decision tree model.[38] The significance threshold for statistical analysis was established at α=0.05.

Results

General Information

This study encompassed 14,246 patients diagnosed with UL. The predominant age group was 40–49 years, representing 47.71%, while those aged 30–39 years comprised 27.68%. In terms of occupation, workers constituted 35.41%, totaling 5045 cases. Patients hospitalized only once constituted 89.41%. Urban employee medical insurance was the payment method for 38.70% of the patients. A total of 98.43% of the patients did not necessitate transfer to another department. The LOS for most patients ranged between 4 and 6 days, representing 56.05%. 57.10% of the patients originated from the local city. Outpatient cases constituted 99.28%. Intramural uterine leiomyomas accounted for 15.35%. A total of 89.84% of the patients presented with an admission condition. General complications and comorbidities were present in 74.36% of the patients. Laparoscopic surgery was the predominant surgical method for 69.53% of the patients, while abdominal and hysteroscopic surgeries constituted 14.19% and 10.38%, respectively. Other types of surgical procedures primarily involved adnexal surgery and other surgeries or interventions, representing 46.76%. A total of 99.70% of the patients were discharged upon medical advice. The median hospitalization cost was 19,030.49 yuan, with an average cost of 18,627.26 yuan per case. For more details, please refer to Table 1.

Factors Influencing Inpatient Costs for Uterine Leiomyoma

Table 1 displays the univariate analysis results for inpatient costs among patients with UL. Age, occupation, number of hospitalizations, type of medical insurance, transfer to other departments, LOS, type of UL, admission condition, comorbidities and complications, type of primary procedure, other types of surgical procedures, and discharge method significantly impacted inpatient costs (*P*<0.05). However, the source of patient and admission pathway did not show a statistically significant difference (*P*>0.05).

Table 2 showcases the multivariate analysis findings on inpatient costs for patients with UL. The regression model's adjusted coefficient of determination (R^2) was 0.655, and its variance inflation factor (VIF) was below 3, indicating negligible collinearity among the independent variables. The model demonstrated a robust data fit and was statistically significant (F=2710.297, *P*<0.001). The study revealed that, aside from the type of medical insurance and admission condition, all other variables significantly influenced inpatient costs (*P*<0.001). Comparison of the regression model's standardized coefficients indicates that the type of primary procedure, other types of surgical procedures, and LOS were the foremost factors affecting inpatient costs.

476

DovePress

Risk Management and Healthcare Policy 2024:17

Table 1 Univariate Analysis of Inpatient Costs in Patients with Uterine Leiomyoma (n=14,246)

Variables	N	%	Cost in RMB Yuan (M ± IQR)	Z/H	P value
Age				893.569	<0.001
≤29	1067	7.49	16,513.86±5215.04		
30–	3943	27.68	17,369.64±5972.08		
40–	6797	47.71	19,563.84±6217.33		
50–	2217	15.56	21,055.63±4634.29		
≥60	222	1.56	21,581.04±4649.79		
Occupation				168.446	<0.001
Unemployed	934	6.56	19,436.55±6386.61		
Worker	5045	35.41	18,733.74±6633.62		
Peasant	562	3.94	20,166.21±6361.34		
Individual or freelance practice	1269	8.91	19,402.70±6400.69		
Student	33	0.23	17,896.74±4023.53		
Retirees	694	4.87	21,189.14±4918.01		
Other	5709	40.07	18,716.02±6482.47		
Number of hospitalizations				−2.07	0.038
1	12,737	89.41	19,066.11±6402.93		
≥2	1509	10.59	18,703.08±7979.34		
Type of medical insurance				204.047	<0.001
No	607	4.26	16,058.43±11,269.20		
Urban employee medical insurance	5513	38.70	19,205.71±6461.03		
Urban resident medical insurance	2431	17.06	19,486.51±6338.60		
New rural cooperative medical insurance	542	3.80	20,185.83±6212.53		
Other	5153	36.17	18,809.95±6417.85		
Transfer to other departments				−6.77	<0.001
No	14,022	98.43	18,992.42±6555.52		
Yes	224	1.57	21,479.82±6532.34		
LOS				4758.634	<0.001
1–3d	2803	19.68	11,334.91±8116.66		
4–6d	7985	56.05	19,214.84±5207.17		
≥7d	3458	24.27	22,015.88±5322.26		
Source of patient				1.469	0.480
Local	8135	57.10	18,994.79±6631.59		
Within the province	2855	20.04	19,223.13±6579.78		
Outside the province	3256	22.86	18,887.31±6444.11		
Admission pathway				4.336	0.227
Emergency	100	0.70	20,350.42±8832.61		
Outpatient	14,143	99.28	19,018.94±6552.24		
Transfer from other hospitals	2	0.01	24,741.34±11,396.46		
Other	1	0.01	21,413.45±0		
Type of uterine leiomyoma				2245.916	<0.001
Submucous myoma of uterus	1406	9.87	9412.75±3198.57		
Subserous myoma of uterus	669	4.70	18,447.45±5675.31		
Intramural uterine fibroids	2187	15.35	19,076.19±6032.88		
Undefined type	9984	70.08	19,718.33±5755.12		
Admission condition				71.14	<0.001
No	4	0.03	25,175.38±38,972.34		
Yes	12,798	89.84	19,156.72±6375.20		
Uncertain situation	76	0.53	16,882.85±11,006.62		
Clinical uncertainty	1368	9.60	17,785.92±7664.41		

(Continued)

Table 1 (Continued).

Variables	N	%	Cost in RMB Yuan (M ± IQR)	Z/H	*P* value
Comorbidities and complications				3692.709	<0.001
No	3498	24.55	13,716.04±7656.27		
CC	10,593	74.36	20,229.07±5206.58		
MCC	155	1.09	22,469.57±7061.75		
Type of primary procedure				4924.132	<0.001
No surgical intervention or only other interventions	426	2.99	2360.15±1049.26		
High-intensity focused ultrasound therapy	341	2.39	14,258.39±2456.98		
Hysteroscopic surgery	1479	10.38	9268.53±2711.18		
Total vaginal surgery	26	0.18	10,835.09±2722.09		
Abdominal surgery	2021	14.19	19,800.34±6646.21		
Laparoscopic surgery	9905	69.53	19,996.57±5006.76		
Uterine artery embolization	48	0.34	13,908.93±714.76		
Other types of surgical procedures				6395.449	<0.001
No surgical intervention or only intervention	1975	13.86	10,162.86±7081.93		
Other surgeries or interventions	5447	38.24	17,253.03±5062.11		
Adnexal surgery or intervention	79	0.55	16,040.87±3116.53		
Adnexal surgery and other surgeries or interventions	6662	46.76	21,419.92±4161.86		
Da Vinci robot-assisted surgery and other surgeries or interventions	29	0.20	58,118.70±3275.17		
Da Vinci robot-assisted surgery and adnexal surgery and other surgeries or interventions	54	0.38	61,711.80±4591.73		
Discharge method				27.484	<0.001
Non-medically advised discharge	19	0.13	2265.18±14,509.42		
Medically advised discharge	14,203	99.70	19,036.22±6563.83		
Medically advised transfer	24	0.17	19,914.98±8309.32		

Notes: The significance level for statistical analysis was set at α=0.05. Undefined uterine leiomyoma type refers to lack of histopathology. Admission condition refers to the assessment of the patient's condition at admission. M ± IQR, median ± interquartile range; Z, Mann–Whitney *U*-test; H, Kruskal–Wallis *H*-test.
Abbreviations: LOS, length of stay; CC, general complications and comorbidities; MCC, major complications and comorbidities; N, number; RMB, ren min bi.

Table 2 Multivariate Stepwise Regression Analysis of Hospitalization Costs in Patients with Uterine Leiomyoma

Variables	Non-Standardized Coefficients		Standardized Coefficients	*T*-Statistic	*P* value	Tolerance	VIF
	β±SE	95% *CI*					
Constant	7.365±0.089	(7.192, 7.539)		83.116	<0.001		
Type of primary procedure	0.187±0.002	(0.182, 0.191)	0.518	80.500	<0.001	0.583	1.714
Other types of surgical procedures	0.094±0.003	(0.089, 0.099)	0.235	36.894	<0.001	0.594	1.683
LOS	0.165±0.004	(0.156, 0.173)	0.230	39.730	<0.001	0.724	1.382
Comorbidities and complications	0.064±0.006	(0.052, 0.076)	0.060	10.336	<0.001	0.708	1.412
Types of uterine leiomyoma	−0.028±0.003	(−0.034, −0.023)	−0.058	−10.341	<0.001	0.774	1.292
Occupation	−0.009±0.001	(−0.011, −0.007)	−0.044	−8.906	<0.001	0.985	1.016
Discharge method	0.358±0.043	(0.274, 0.441)	0.041	8.414	<0.001	0.995	1.005
Age	−0.020±0.003	(−0.026, −0.015)	−0.037	−6.887	<0.001	0.832	1.202
Transfer to other departments	0.123±0.019	(0.086, 0.159)	0.032	6.549	<0.001	0.999	1.001
Number of hospitalizations	0.030±0.008	(0.015, 0.045)	0.019	3.881	<0.001	0.986	1.014

Notes: An explanation of the variable types can be clearly found in Table 1. The significance level for statistical analysis was set at α=0.05.
Abbreviations: VIF, Variance Inflation Factor; β, non-standardized coefficients; SE, standard error; CI, confidence interval.

DRG Grouping Prediction Using a Decision Tree Model for Uterine Leiomyoma

Table 3 presents the optimized DRG grouping scheme for patients with UL. A decision tree model was constructed, employing the type of primary procedure, other types of surgical procedures, and LOS as predictive variables. The model was segmented into three layers, yielding a total of 11 case combinations. The results of the grouping are illustrated in Figure 1.

Table 3 The DRG Grouping Scheme of Uterine Leiomyoma

Groups	Grouping Rules	N (%)	Median ± IQR in RMB Yuan	Mean	Standard Deviation
DRG1	No surgical intervention or only other interventions	426(2.99%)	2360.15±1049.26	2875.34	2459.66
DRG2	Uterine artery embolization or high-intensity focused ultrasound therapy	389(2.73%)	14,161.55±2444.01	14,763.58	1522.54
DRG3	Hysteroscopic surgery or total vaginal surgery, LOS 1–3d	1095(7.69%)	8748.81±2128.66	9175.63	2208.39
DRG4	Hysteroscopic surgery or total vaginal surgery, LOS≥4d	410(2.88%)	11,219.12±4442.49	12,916.82	4679.66
DRG5	Laparoscopic surgery or abdominal surgery, no surgical intervention or only intervention or other surgeries or interventions or adnexal surgery or intervention, LOS 1–3d	697(4.89%)	15,969.96±3846.68	16,135.94	2847.28
DRG6	Laparoscopic surgery or abdominal surgery, no surgical intervention or only intervention or other surgeries or interventions or adnexal surgery or intervention, LOS 4–6d	3501(24.58%)	17,337.28±4128.84	17,624.83	3128.04
DRG7	Laparoscopic surgery or abdominal surgery, no surgical intervention or only intervention or other surgeries or interventions or adnexal surgery or intervention, LOS≥7d	1113(7.81%)	19,710.67±5805.85	20,456.05	4965.95
DRG8	Laparoscopic surgery or abdominal surgery, adnexal surgery and other surgeries or interventions, LOS 1–3d	320(2.25%)	19,000.72±3503.68	19,119.55	3025.70
DRG9	Laparoscopic surgery or abdominal surgery, adnexal surgery and other surgeries or interventions, LOS 4–6d	3950(27.73%)	20,964.93±3589.97	21,062.23	2899.67
DRG10	Laparoscopic surgery or abdominal surgery, adnexal surgery and other surgeries or interventions, LOS≥7d	2262(15.88%)	22,800.71±4673.43	23,459.25	4629.93
DRG11	Laparoscopic surgery or abdominal surgery, Da Vinci robot-assisted surgery and other surgeries or interventions or Da Vinci robot-assisted surgery and adnexal surgery and other surgeries or interventions	83(0.58%)	60,163.34±5278.26	59,698.99	4015.16

Abbreviations: N, number; IQR, interquartile range; LOS, length of stay; RMB, ren min bi.

Standard Costs and Weights for DRG Grouping of Uterine Leiomyoma

Table 4 displays the cost standards and instances of exceeding limits in the optimized DRG grouping scheme for patients with UL. The CV for the optimized DRG groups was less than 1, demonstrating a high homogeneity in inpatient costs within each DRG group. The RIV was 72.02%, and the results from the Kruskal–Wallis H non-parametric test revealed statistically significant differences in inpatient costs among the DRG groups (H=8087.635, P<0.001), signifying high heterogeneity in inpatient costs across the DRG groups. Crucially, the analytic proportions of the standardized costs for each DRG group surpassed 80%, demonstrating that the standardized costs of the optimized DRG groups provided a robust explanation for the actual costs. Considering these indicators, the DRG grouping scheme for UL demonstrated effective performance in this study.

In the grouping scheme of this study, DRG 11 exhibited the highest standard costs and weight values, indicative of complex cases and higher consumption of medical resources. Conversely, DRG 1 displayed the lowest standard costs and weight values, reflecting simpler procedures and lesser consumption of medical resources. Additionally, among the study samples, 330 cases exceeded the cost limit, representing 2.32% of all cases. These over-limit cases constituted 3.49% of the total costs. Excluding DRG 1 and DRG 4, the remaining groups had a relatively low proportion of over-limit cases.

Discussion

Drawing from international practices, DRG takes into consideration the complexity of diseases and individual patient characteristics. Through effective medical cost management and quality assessment, DRG balances the interests of governments, medical institutions, and patients.[39] Currently, DRG is recognized internationally as a payment method. Importantly, DRG not only serves as a payment method within the healthcare sector but also prompts doctors to devise reasonable treatment plans, thus mitigating unnecessary medical expenses and alleviating patients' financial burdens.
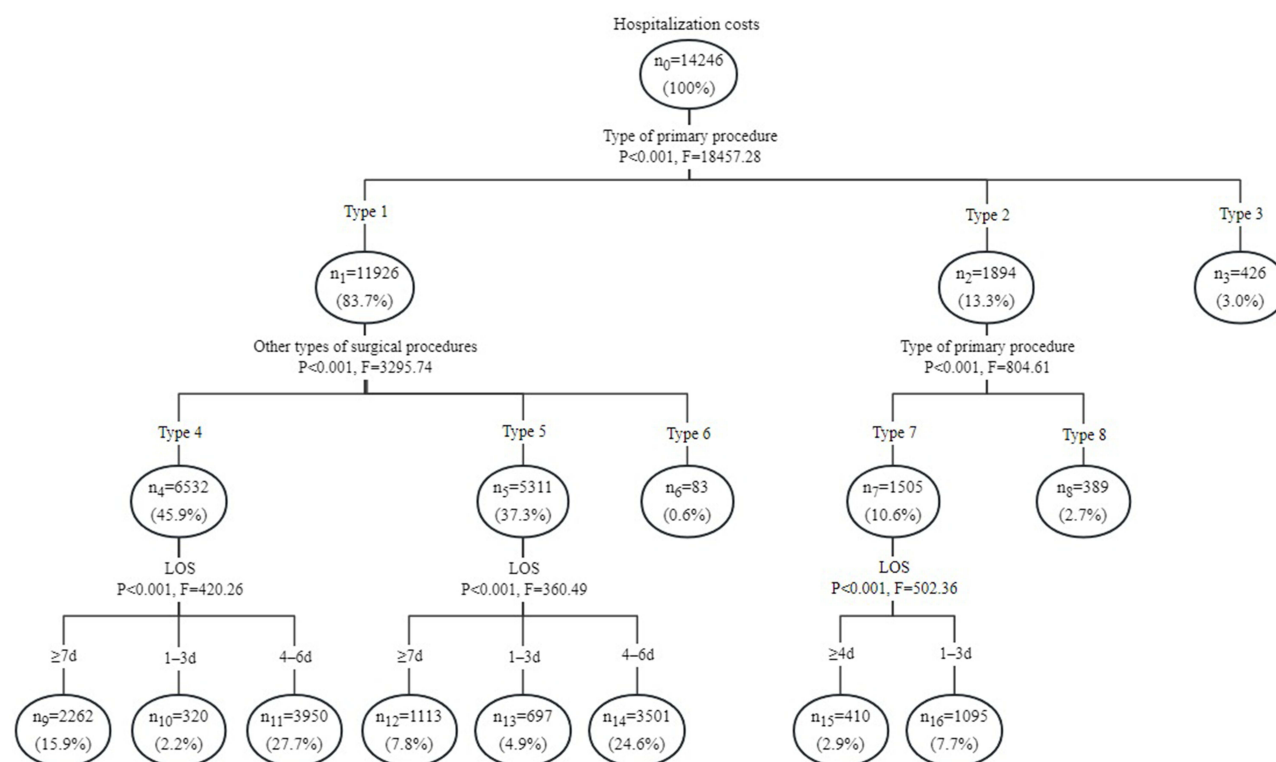
Figure 1 Grouping results based on the decision tree model.

Notes: Type 1, laparoscopic or abdominal surgery; Type 2, hysteroscopic or total vaginal surgery/uterine artery embolization/high-intensity focused ultrasound therapy; Type 3, no surgical intervention or only other interventions; Type 4, adnexal surgery and other surgeries or interventions; Type 5, no surgical intervention or only intervention/other surgeries or interventions/adnexal surgery or intervention; Type 6, Da Vinci robot-assisted surgery and other surgeries or interventions/Da Vinci robot-assisted surgery and adnexal surgery and other surgeries or interventions; Type 7, hysteroscopic surgery/total vaginal surgery; Type 8, uterine artery embolization/high-intensity focused ultrasound therapy. The decision tree model had 3 levels, with 1 parent node and 16 child nodes. The end of the tree model was used as grouping marker, and one branch was a DRG group. Finally, 11 DRG groups were formed.

Additionally, DRG functions as a decision-making tool for healthcare institutions, facilitating enhanced precision in management. In this study, we concentrated on patients with UL at a single center in a city in western China, examining the DRG grouping scheme and payment standards for this condition. By analyzing sample data from 2019 to 2022, we

Table 4 Payment Standard and Exceedance of the DRG Grouping Scheme for Uterine Leiomyoma

Groups	Payment Standard	RW	Upper Limit	Excess Amount N (%)	Percentage of Excess Expenses (%)	CV	Analytic Proportion (%)
DRG1	2360.15	0.15	4519.51	38(8.92%)	26.25	0.86	82.08
DRG2	14,161.55	0.79	19,810.14	2(0.51%)	0.70	0.10	95.92
DRG3	8748.81	0.49	13,230.56	35(3.20%)	6.16	0.24	95.35
DRG4	11,219.12	0.69	20,943.32	37(9.02%)	16.98	0.36	86.86
DRG5	15,969.96	0.87	23,801.57	2(0.29%)	0.44	0.18	98.97
DRG6	17,337.28	0.95	25,650.37	46(1.31%)	2.06	0.18	98.37
DRG7	19,710.67	1.10	31,503.55	40(3.59%)	6.47	0.24	96.36
DRG8	19,000.72	1.03	25,997.47	7(2.19%)	3.24	0.16	99.38
DRG9	20,964.93	1.13	28,190.06	47(1.19%)	1.72	0.14	99.54
DRG10	22,800.71	1.26	32,400.56	76(3.36%)	5.57	0.20	97.19
DRG11	60,163.34	3.20	70,339.74	0(0%)	0.00	0.07	100.78

Abbreviations: RW, relative weight; CV, Coefficient of Variation.

480

Risk Management and Healthcare Policy 2024:17

employed the decision tree model in machine learning, specifically the E-CHAID decision tree model, to develop the UL case grouping scheme for the region. Although the study sample was only from a single center, the cases were well represented in western China. There are three primary reasons. First, it serves as both a national regional medical center and a quality control center for obstetrics and gynecology in Sichuan Province. Second, it encompasses medical treatment, teaching, scientific research, prevention, healthcare, and personnel training as an integral part of the university-affiliated hospital. Third, the gynecology department at this center is recognized as a significant national clinical specialty project. Consequently, the findings of this study may offer valuable insights to regional health administrative departments in optimizing the DRG grouping scheme.

Our study revealed that the E-CHAID decision tree model classified UL patients into 11 DRG groups, utilizing three classification nodes. The type of primary procedure served as the first-level classification node, aligning with findings from domestic tumor-related research.[33,40] Additionally, multivariate stepwise linear regression analysis identified the type of primary procedure as the most significant factor influencing UL patients' hospitalization costs. In the case combinations, type of primary procedure was categorized into three groups: 1) no surgical intervention or only other interventions; 2) uterine artery embolization/high-intensity focused ultrasound therapy/hysteroscopy or total vaginal surgery; 3) laparoscopic or abdominal surgery. Different types of surgery can influence patients' medical resource utilization. This study demonstrated that patients undergoing laparoscopic or abdominal surgery incurred the highest hospitalization costs. The reason is that both laparoscopic and abdominal surgeries involve traumatic incisions, posing a certain infection risk and necessitating anesthesia. Additionally, laparoscopic surgery necessitates specific surgical instruments and, unlike hysteroscopy, permits the examination of other pelvic organs beyond the uterus.[19]

The second-level classification node is divided into two categories: other types of surgical procedures and types of primary procedures. For the primary surgery type 2 at the first level, is further divided into uterine artery embolization or high-intensity focused ultrasound therapy, and hysteroscopy or total vaginal surgery. This is because there is a significant difference in hospitalization costs between patients undergoing uterine artery embolization or high-intensity focused ultrasound therapy and those undergoing hysteroscopy or total vaginal surgery. The former belongs to modern treatment techniques for uterine fibroids and has slightly higher costs. In patients undergoing laparoscopic or abdominal surgery, other types of surgical procedures also have a significant impact on the hospitalization costs of UL patients. After case combinations, other types of surgical procedures are divided into three categories: 1) no surgical intervention or only intervention/other surgeries or interventions/adnexal surgery or intervention; 2) adnexal surgery and other surgeries or interventions; 3) Da Vinci robot-assisted surgery and other surgeries or interventions/Da Vinci robot-assisted surgery and adnexal surgery and other surgeries or interventions. Understandably, the presence of other surgeries or adnexal surgery as well as Da Vinci robot-assisted surgery will increase the consumption of medical resources for UL patients. In particular, research has indicated that the lifespan of robotic instruments is limited, and the cost of replacing instruments is high, with maintenance costs accounting for 10% of the capital acquisition cost.[41,42]

The LOS serves as the third-level classification node. Longer hospital stays correlate with increased hospitalization costs, in alignment with the findings of numerous studies.[33,43] This correlation exists because an increase in LOS leads to greater consumption of medical resources by patients. Interestingly, previous studies have presented divergent views on using LOS as a classification node variable. Wu et al argue that LOS merely represents one aspect of medical output and is unsuitable for analyzing medical cost structures or serving as a classification node.[28] Hasdai highlighted that LOS is a crucial indicator of disease severity and merits consideration as a classification node.[44] Furthermore, in DRG grouping guidelines, developed countries, including the United States, the United Kingdom, and Poland, regard LOS as a significant factor.[45] Referencing the study by Zeng et al, incorporating LOS as a node variable has been shown to facilitate more precise grouping and reduce the proportion of outliers in various groups.[46] Most importantly, the adoption of DRG payment systems incentivizes medical institutions to enhance efficiency and, to some extent, shorten the LOS.[47]

The evaluation indicators from the comprehensive decision tree model demonstrate that the DRG grouping scheme for UL in this study is sound. Our findings indicate significant homogeneity in hospitalization costs within DRG groups and substantial heterogeneity in hospitalization costs between DRG groups. This aligns with the fundamental principle of DRG grouping.[33] Additionally, the subdivided DRG groups possess strong explanatory capability regarding actual hospitalization costs, serving as a valuable reference for medical insurance departments to optimize the UL case grouping scheme.

Analysis of outliers reveals that 330 cases in the overall research sample surpass the cost limit, constituting 2.32% of all cases. Relevant studies suggest that, given specific diagnoses and treatment scenarios, a certain proportion of patients' hospitalization costs within the DRG case mix may exceed the cost limit, but this should not exceed 5%.[48] Overall, our study's outlier proportion aligns with the established requirements. Notably, the outlier proportions in the DRG1 and DRG4 groups stand at 8.92% and 9.02% respectively, surpassing the widely accepted standard. This result suggests that patients in these two DRG groups may consume excessive medical resources, potentially due to over-diagnosis and treatment. Under the DRG payment model, this phenomenon could result in financial losses for medical institutions, impacting their sustainable development.

Evidence suggests that clinical pathways can effectively standardize physicians' medical practices as evidence-based treatment plans.[49] Firstly, clinical pathways reduce unnecessary days of hospitalization and improve bed turnover rates, which not only reduces the financial burden on patients but also promotes a model of quality and efficiency development for healthcare institutions. Secondly, clinical pathways optimize hospitalization cost structures, reflecting the value of healthcare professionals' technical skills and labor. Lastly, clinical pathways can reduce information asymmetry between doctors and patients. Therefore, implementing clinical pathways is recommended to regulate physicians' medical practices, control unreasonable medical expenses, and reduce the occurrence of outliers, benefitting patients, healthcare institutions, and medical insurance departments. Moreover, the quality of DRG grouping directly depends on the data provided in the medical record abstract.[50] It is imperative for healthcare institutions to strengthen standardized training for clinical personnel, medical record staff, and statistical management teams to ensure the medical record abstracts are accurate and complete.[51] Additionally, enhancing electronic medical records and upgrading hospital information system (HIS) functionality is crucial for effective DRG management.[52]

Our sample data were collected between 2019 and 2022. During this period, the COVID-19 pandemic occurred. Fortunately, a series of measures implemented at the center ensured that patients with UL were minimally affected, with no fatalities reported. Following the epidemic's outbreak, the center diligently adhered to government policy instructions and relevant prevention and control measures. For example, it established an Incident Response Service and a command system, identifying responsible departments and members. Clear management protocols were established for the fever clinic, general clinic, wards, COVID-19 nucleic acid testing, and hospital staff. Additionally, employees were tasked with conducting pre-examination, triage, and verifying patients' health codes. In summary, the center maintained normal medical operations during the COVID-19 epidemic and played an active role in the regional epidemic's prevention and control.

Limitations

This study has several limitations. Firstly, the data on UL patients are collected from a single center in a city in western China. Secondly, the absence of standardized data reporting by medical institutions leads to a scarcity of information on various factors influencing hospital costs, including marital status, patient severity, critical conditions, and methods of anesthesia. Thirdly, excessive costs can be attributed to numerous factors, including excessive medical interventions, inadequate clinical pathway management, and the use of various pharmaceutical hygiene materials. However, our current dataset does not support these assertions.

Future Direction

Moving forward, we intend to incorporate a more representative selection of multi-center data, which will be evaluated and refined throughout the survey practice process. Furthermore, it is imperative to enhance the fundamental quality control measures for reported data, and a comprehensive consideration of the factors influencing hospitalization expenses is necessary.

Conclusions

This study introduces a comprehensive DRG grouping scheme for UL, employing a meticulous screening of key node variables influencing hospitalization costs and utilizing the E-CHAID algorithm for case combination classification. Our findings reveal that the hospital costs for UL patients are predominantly associated with the type of primary procedure,

other types of surgical procedures, and LOS. Furthermore, conducting research using real-world data from the western region provides valuable insights for the formulation of a localized DRG grouping scheme. Our study can serve as a reference for other regions or countries aiming to enhance their disease diagnosis grouping systems. Moreover, our research offers a strategic approach to manage hospitalization costs associated with UL, thereby potentially alleviating the financial burden on patients.

Ethics Approval

This study has been approved by the Ethics Committee of West China Second Hospital of Sichuan University (No.2023257) and all patients signed the informed consent form. Ethical principles of the Declaration of Helsinki were adhered to throughout this study.

Disclosure

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Bulletins-Gynecology ACo P. ACOG practice bulletin. Surgical alternatives to hysterectomy in the management of leiomyomas. Number 16, May 2000 (replaces educational bulletin number 192, May 1994). *Int J Gynaecol Obstet*. 2001;73(3):285–293. doi:10.1016/s0020-7292(01)00414-3
2. Stewart EA. Uterine fibroids. *Lancet*. 2001;357(9252):293–298. doi:10.1016/S0140-6736(00)03622-9
3. Sankaran S, Manyonda IT. Medical management of fibroids. *Best Pract Res Clin Obstet Gynaecol*. 2008;22(4):655–676. doi:10.1016/j. bpobgyn.2008.03.001
4. Parker WH. Etiology, symptomatology, and diagnosis of uterine myomas. *Fertil Steril*. 2007;87(4):725–736. doi:10.1016/j.fertnstert.2007.01.093
5. Jiang XB, Liu XK, Han L. Analysis of influencing factors of hospitalization expenses in patients with uterine leiomyoma. *Mod Preventive Med*. 2020;47(12):2202–2205+2232.
6. Sabry M, Al-Hendy A. Medical treatment of uterine leiomyoma. *Reprod Sci*. 2012;19(4):339–353. doi:10.1177/1933719111432867
7. Sunkara SK, Khairy M, El-Toukhy T, et al. The effect of intramural fibroids without uterine cavity involvement on the outcome of IVF treatment: a systematic review and meta-analysis. *Hum Reprod*. 2010;25(2):418–429. doi:10.1093/humrep/dep396
8. Kashani BN, Centini G, Morelli SS, et al. Role of medical management for uterine leiomyomas. *Best Pract Res Clin Obstet Gynaecol*. 2016;34:85–103. doi:10.1016/j.bpobgyn.2015.11.016
9. Giannini A, Golia D'Augè T, Bogani G, et al. Uterine sarcomas: a critical review of the literature. *Eur J Obstet Gynecol Reprod Biol*. 2023;287:166–170. doi:10.1016/j.ejogrb.2023.06.016
10. D'Angelo E, Prat J. Uterine sarcomas: a review. *Gynecol Oncol*. 2010;116(1):131–139. doi:10.1016/j.ygyno.2009.09.023
11. Mallmann P. Uterine sarcoma - difficult to diagnose, hard to treat. *Oncol Res Treat*. 2018;41(11):674. doi:10.1159/000494393
12. Benson C, Miah AB. Uterine sarcoma - current perspectives. *Int J Womens Health*. 2017;9:597–606. doi:10.2147/IJWH.S117754
13. Canadian Institute for Health Information, Statistics Canada. Health indicators. Ottawa: CIHI; 2010. Available from: https://secure.cihi.ca/free_ products/Healthindicators2010_en.pdf. Accessed July 10, 2023.
14. Whiteman MK, Hillis SD, Jamieson DJ, et al. Inpatient hysterectomy surveillance in the United States, 2000–2004. *Am J Obstet Gynecol*. 2008;198 (1):e1–7. doi:10.1016/j.ajog.2007.05.039

15. Wu JM, Wechter ME, Geller EJ, et al. Hysterectomy rates in the United States, 2003. *Obstet Gynecol*. 2007;110(5):1091–1095. doi:10.1097/01. AOG.0000285997.38553.4b

16. Boyd LR, Novetsky AP, Curtin JP. Effect of surgical volume on route of hysterectomy and short-term morbidity. *Obstet Gynecol*. 2010;116 (4):909–915. doi:10.1097/AOG.0b013e3181f395d9

17. Flynn M, Jamison M, Datta S, et al. Health care resource use for uterine fibroid tumors in the United States. *Am J Obstet Gynecol*. 2006;195 (4):955–964. doi:10.1016/j.ajog.2006.02.020

18. Li L, Wang YL. Comparison of clinical efficacy between laparoscopic myomectomy and laparotomy. *Chin J Clin Oncol Rehabil*. 2012;19 (4):363–364.

19. Gu Y, Zhu L, Liu A, et al. Analysis of hysterectomies for patients with uterine leiomyomas in China in 2010. *Int J Gynaecol Obstet*. 2015;129 (1):71–74. doi:10.1016/j.ijgo.2014.10.031

20. Busse R, Geissler A, Aaviksoo A, et al. Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? *BMJ*. 2013;346:f3197. doi:10.1136/bmj.f3197

21. Hunter WG, Hesson A, Davis JK, et al. Patient-physician discussions about costs: definitions and impact on cost conversation incidence estimates. *BMC Health Serv Res*. 2016;16:108. doi:10.1186/s12913-016-1353-2

22. Polyzos N, Karanikas H, Thireos E, et al. Reforming reimbursement of public hospitals in Greece during the economic crisis: implementation of a DRG system. *Health Policy*. 2013;109(1):14–22. doi:10.1016/j.healthpol.2012.09.011

23. Quinn K. After the revolution: dRGs at age 30. *Ann Intern Med*. 2014;160(6):426–429. doi:10.7326/M13-2115

24. Isilak Z, Yalcin M, Un H, et al. Fractional flow reserve-guided lesion or patient management? *Chin Med J*. 2015;128(23):3266. doi:10.4103/0366-6999.170273

25. Thurgar E, Barton S, Karner C, et al. Clinical effectiveness and cost-effectiveness of interventions for the treatment of anogenital warts: systematic review and economic evaluation. *Health Technol Assess*. 2016;20(24):1–486. doi:10.3310/hta20240

26. Goldfield N. The evolution of diagnosis-related groups (DRGs): from its beginnings in case-mix and resource use theory, to its implementation for payment and now for its current utilization for quality within and outside the hospital. *Qual Manag Health Care*. 2010;19(1):3–16. doi:10.1097/ QMH.0b013e3181ccbcc3

27. Quentin W, Scheller-Kreinsen D, Geissler A, et al. Appendectomy and diagnosis-related groups (DRGs): patient classification and hospital reimbursement in 11 European countries. *Langenbecks Arch Surg*. 2012;397(2):317–326. doi:10.1007/s00423-011-0877-5

28. Wu SW, Pan Q, Chen T, et al. Research of medical expenditure among inpatients with unstable angina pectoris in a single center. *Chin Med J*. 2017;130(13):1529–1533. doi:10.4103/0366-6999.208242

29. Liu X, Fang C, Wu C, et al. DRG grouping by machine learning: from expert-oriented to data-based method. *BMC Med Inform Decis Mak*. 2021;21 (1):312. doi:10.1186/s12911-021-01676-7

30. The Office of the National Medical Security Administration. Notice on the issuance of paid national pilot technical specifications and grouping schemes for disease diagnostic related groups; 2019. Available from: http://www.nhsa.gov.cn/art/2019/10/24/art_37_1878.html. Accessed July 10, 2023.

31. Zou K, Li HY, Zhou D, et al. The effects of diagnosis-related groups payment on hospital healthcare in China: a systematic review. *BMC Health Serv Res*. 2020;20(1):112. doi:10.1186/s12913-020-4957-5

32. Luo AJ, Chang WF, Xin ZR, et al. Diagnosis related group grouping study of senile cataract patients based on E-CHAID algorithm. *Int J Ophthalmol*. 2018;11(2):308–313. doi:10.18240/ijo.2018.02.21

33. Wu SW, Pan Q, Chen T. Research on diagnosis-related group grouping of inpatient medical expenditure in colorectal cancer patients based on a decision tree model. *World J Clin Cases*. 2020;8(12):2484–2493. doi:10.12998/wjcc.v8.i12.2484

34. Zhi M, Hu L, Geng F, et al. Analysis of the cost and case-mix of post-acute stroke patients in china using quantile regression and the decision-tree models. *Risk Manag Healthc Policy*. 2022;15:1113–1127. doi:10.2147/RMHP.S361385

35. Yu T, He Z, Zhou Q, et al. Analysis of the factors influencing lung cancer hospitalization expenses using data mining. *Thorac Cancer*. 2015;6 (3):338–345. doi:10.1111/1759-7714.12147

36. Biggs D, De Ville B, Suen E. A method of choosing multiway partitions for classification and decision trees. *J Appl Stat*. 1991;18(1):49–62.

37. Tanguy-Melac A, Aguade AS, Fagot-Campagna A, et al. Management and intensity of medical end-of-life care in people with colorectal cancer during the year before their death in 2015: a French national observational study. *Cancer Med*. 2019;8(15):6671–6683. doi:10.1002/cam4.2527

38. Wu XZ. DRGs case-mix study on hospitalization costs of pneumonia patients. *China Health Statistics*. 2020;37(2):235–238.

39. Wang S, Rao KQ, Liu LH. Current status of DRG payment in china and reference of DRG payment in the world. *Chin J Hospital Admin*. 2015;31 (2):81–83.

40. Zhou LY, Sun MJ, Xia XJ, et al. Study on DRGs grouping of patients with uterine leiomyoma surgery in a 3A hospital in Shanxi Province. *China Health Ser Manage*. 2023;40(2):95–99+135.

41. Newbold RR, DiAugustine RP, Risinger JI, et al. Advances in uterine leiomyoma research: conference overview, summary, and future research recommendations. *Environ Health Perspect*. 2000;108(Suppl 5):769–773. doi:10.1289/ehp.00108s5769

42. The SAGES-MIRA Robotic Surgery Consensus Group. A consensus document on robotic surgery; 2007. Available from: http://www.sages.org/ publications/guidelines/consensus-document-robotic-surgery. Accessed July 10, 2023.

43. Su FY, Fu ML, Zhao QH, et al. Analysis of hospitalization costs related to fall injuries in elderly patients. *World J Clin Cases*. 2021;9 (6):1271–1283. doi:10.12998/wjcc.v9.i6.1271

44. Hasdai D, Garratt KN, Holmes DR, et al. Coronary angioplasty and intracoronary thrombolysis are of limited efficacy in resolving early intracoronary stent thrombosis. *J Am Coll Cardiol*. 1996;28(2):361–367. doi:10.1016/0735-1097(96)00136-2

45. Fox KM. Investigators EUtOrocewPiscAd. Efficacy of perindopril in reduction of cardiovascular events among patients with stable coronary artery disease: randomised, double-blind, placebo-controlled, multicentre trial (the Europa study). *Lancet*. 2003;362(9386):782–788. doi:10.1016/s0140-6736(03)14286-9

46. Zeng S, Luo L, Fang Y, et al. Cost control of treatment for cerebrovascular patients using a machine learning model in Western China. *J Healthc Eng*. 2021;2021:6158961. doi:10.1155/2021/6158961

47. Brügger U, Eichler K. Impact of introducing a DRG reimbursement system in an acute inpatient hospital setting: a literature review. *Rheumatology*. 2010;46(3):556–557.

48. Jin PM, Hua W, Chen J, et al. Study on establishing hospitalization expense standard of single disease based on disease diagnosis correlation group method. *China Health Econo*. 2017;36(2):26–28.

49. Li YW, Xie HY, Zhang B, et al. Effect of clinical pathway on improving the quality of medical service in hysterectomy for uterine leiomyoma. *China Health Qual Manage*. 2013;20(1):28–30.

50. Tanner K, Sabrine N, Wren C. Cardiovascular malformations among preterm infants. *Pediatrics*. 2005;116(6):e833–e838. doi:10.1542/peds.2005-0397

51. Ma Z. Analysis of the influence of quality control of the first page of medical record on DRG. *China Health Econo*. 2018;37(12):94–95.

52. ConwayLenihan A, Ahern S, Moore S, et al. Factors influencing the variation in GMS prescribing expenditure in Ireland. *Health Econ Rev*. 2016;6(1):13. doi:10.1186/s13561-016-0090-x