Study of clinical experiments of novel sedative drugs based on chi-square test and logistics regression model

Hongye Wang *

School of Mathematics and Statistics, Central China Normal University, Wuhan, China

* Corresponding author: 1467953072@qq.com

Abstract. Local sedation and analgesic drugs are often indispensable in medical trauma surgery. They can not only assist doctors to complete the operation, but also greatly reduce the pain of patients. Therefore, constantly looking for better effective sedative drugs is an extremely important field in medical research. For the new sedative drug "R drug", this paper uses the data analysis of patient intraoperative and postoperative adverse reactions, patient vital signs, patient satisfaction and related problems to establish an appropriate mathematical model and evaluate its clinical efficacy. First, Data cleaning was performed on the data[1], Quantification and normalized dimensionless pretreatment; Next, We calculated K2 values using the chi-square test and Yates's correction, By comparing the K2 values to the standard values, Then to judge whether there are significant differences between each index; Then, When predicting adverse reactions, we used a logistics regression model to predict the probability that each reaction may occur, The accuracy of the model is tested by the heat map of the confusion matrix and the ROC curve; Last, We performed the data analysis of each index in 0-5 minutes, By mapping the nuclear density of the different indicators under the B and R drugs, Mann-Whitney U test using independent samples, P values were calculated to determine whether the new drug group and the original drug group showed significant differences in vital sign data.

Keywords: Chi-square test, logistics Regression model, A Mann – Whitney U test for independent samples.

1. Introduction

In medical surgery, we often need to use local sedative and analgesic drugs to assist in completing the operation and relieve the pain of patients. The traditional sedative drug is "B drug". Now a drug research and development center has developed a new drug "R drug". Clinical research is a key link in the research of new drugs, and the use of new drugs usually needs to go through two stages: biological trial and clinical trial. This study is a non-interventional study of "R drug" implemented in the hospital.

2. Data preprocessing

2.1. Normalized to the dimensionless treatment

The specific index is normalized by the following formula:

$$r_{i,j} = \frac{d_{i,j} - mind_{i,j}}{maxd_{i,j} - mind_{i,j}}, (i = 1,2,3,\cdots,m) \quad (1)$$

Normalization eliminates the errors caused by different indexes with different indexes, which makes the data more comparable and interpretable, and then improves the accuracy and stability of the model, thus improving the prediction effect of the model.

2.2. Quantification of preprocessing

We quantify the data of some indicators in the indicators and change them into 0,1,2 which are easy to calculate and data processing. Because the sample size of PONV and halo history is too small,

we do not consider them. The specific treatment is shown in the following table: in particular, whether the additional injection is indicated by 0,1,2.

Table 1. Quantitative instructions

index	illustrate
gender	man=1, woman=0
History of surgery	yes=1, not=0
Have a past history	yes=1, not=0
Whether you smoke or not	never smoked=0, occasional smokers=1, regular smokers=2
Whether or not to drink heavily	yes=1, not=0
Is there any additional sedation?	yes=1, not=0
Is there a choking cough?	yes=1, not=0
Is there no body-free movement?	yes=1, not=0
Intraoperative other	yes=1, not=0
Description of past history	not=0, yes but not serious=1, yes, but seriously=2

3. Chi-square test model

3.1. Chi-square test principle

The chi-square test is a method employed to examine whether there are notable connections between discrete-type variables. The chi-square test can measure the difference between the observed and expected values, and then determine whether the difference is sufficient to overturn the null hypothesis. Suppose there are two categorical variables X and Y, whose values are X_1, X_2 and Y_1 and Y_2, respectively, and the sample frequency listing table 2 is:

Table 2. 2×2, contingency table

	Y_1	Y_2	Total
X_1	a	b	a+b
X_2	c	d	c+d
Total	a+c	b+d	a+b+c+d

If the argument is H1: "X is related to Y", the independence test can be used to investigate whether the two variables are related, and the reliability of this judgment can be given more accurately. Specifically, the value of the test statistic K^2 is calculated from the data in the table 1.

$$K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}, n = a+b+c+d(a,b,c,d>5) \quad (2)$$

By comparing the actual value with the standard value according to the following table 3, the credibility of the conclusion "X is related to Y" can be determined.

Table 3. Control table of chi-square standard values

$P(K\geq k_0)$	0.05	0.01	0.005	0.001
K_0	3.841	6.635	7.879	10.828

3.2. Intraoperative adverse effects

Table 4. Four intraoperative indicators with B medicine and R medicine 2×2 contingency table

	1	0	Total
B medicine	29	446	475
R medicine	54	716	770
Total	112	1133	1245
Physical movement			

	1	0	Total
B medicine	7	468	475
R medicine	46	724	770
Total	53	1192	1245
Choking			

	1	0	Total
B medicine	29	446	475
R medicine	54	716	770
Total	112	1133	1245
Belch			

	1	0	Total
B medicine	7	468	475
R medicine	46	724	770
Total	53	1192	1245
Snore			

From the contingency table, we can calculate from the square formula of the body movement, cough, hiccup, snoring K^2 is: 7.8393,14.5980,13.2780,1.4048.

3.3. Postoperative adverse effects

Table 5. Four indicators after surgery with B medicine and R medicine 2×2 contingency table

	1	0	Total
B medicine	72	403	475
R medicine	50	720	770
Total	122	1123	1245
Disgusting			

	1	0	Total
B medicine	5	470	475
R medicine	22	748	770
Total	27	1218	1245
Vomit			

	1	0	Total
B medicine	38	437	475
R medicine	54	716	770
Total	92	1153	1245
Dizzy			

	1	0	Total
B medicine	17	458	475
R medicine	8	762	770
Total	25	1220	1245
Dizziness			

After surgery, we take the four indexes in Table 6 as examples to calculate the K^2 of each index as shown in the table below:

Table 6. The chi-square value of each indicator

index	K^2
disgusting	24.9511
vomit	4.5088
dizzy	0.4182
dizziness	9.6320
headache	0.0561
Drowsiness	8.2855
Weak	8.8144
Abdominal distension	22.2125
bellyache	7.2129
other	0.4361

3.4. The Yates correction method for Chi-squared test

In the chi-square test, the Yates correction method is a method to correct for the continuity correction error. It is used to avoid frequency deviations from significant levels when category contingency table data is very small (about 5 or less). Yates, the method of the correction method:

a. For 2x2 contingency tables of any size, the chi-square statistics are calculated using the formula:

$$K^2 = \frac{(|ad-bc|-n/2)^2}{(a+b)(c+d)(a+c)(b+d)} \quad (3)$$

Where a, b, c and d are the observed counts in four cells and n is the sum of all observed counts.

b. For cases with 1 F, the chi-square value should be compared to the p-value to determine whether to reject the null hypothesis or not. After using the Yates correction, the new adjusted card-square value can be obtained using the following formula:

$$K^2Y = \frac{[|ad-bc|-\frac{n}{2}-0.5sign(ad-bc)]^2}{(a+b)(c+d)(a+c)(b+d)} \quad (4)$$

c. By calculating the new adjusted chi-square value K^2Y, you can find a new p-value together with the K^2 distribution table with a degree of freedom of 1.

In the intraoperative adverse reactions, the frequency of snoring and hiccups was very small, and the corrected chi-square values were changed to 0.7386 and 12.0486.

In the postoperative adverse reactions, the frequency of drowsiness was very small, and the chi-square value was changed to 6.9708 after correction.

Considering the above analysis, we found that during the operation, body movement, hiccup, cough, postoperative, nausea, vomiting, dizziness, drowsiness, fatigue, abdominal distension and abdominal pain were significant.

4. logistics Regression model

The logistics regression model was built on the data after normalized dimensionless processing.

Considering the vector t= (t_1, t_2, ..., t_n) of an independent variable, let the conditional probability P (y = 1|x) = P according to the probability of the observed variable relative to t event, then the logistics regression model can be expressed as:

$$P(y = 1|t) = f(t) = \frac{1}{1+e^{-g(t)}} \quad (5)$$

Here $f(t) = \frac{1}{1+e^{-g(t)}}$ is called the logistics function, where g (t) =W_0 + W_1t_1 +... + W_nt_n, then in the x condition, the probability that the next y does not occur is:

$$P(y = 0|t) = 1 - \frac{1}{1+e^{-g(t)}} = \frac{1}{1+e^{g(t)}} \quad (6)$$

Therefore, the ratio of event occurrence and non-occurrence probability is:

$$\frac{P(y = 1|t)}{P(y = 0|t)} = \frac{P}{1-P} = e^{g(t)} \quad (7)$$

This ratio is called the occurrence ratio, simply odds, log odds:

$$ln\frac{P}{1-P} = g(t) = W_0 + W_1t_1 + \cdots + W_nt_n \quad (8)$$

According to the regression coefficient in the table, we can predict each adverse reaction. The following table shows the regression results of the injection [3] [5].

Table 7. The regression results of the spray

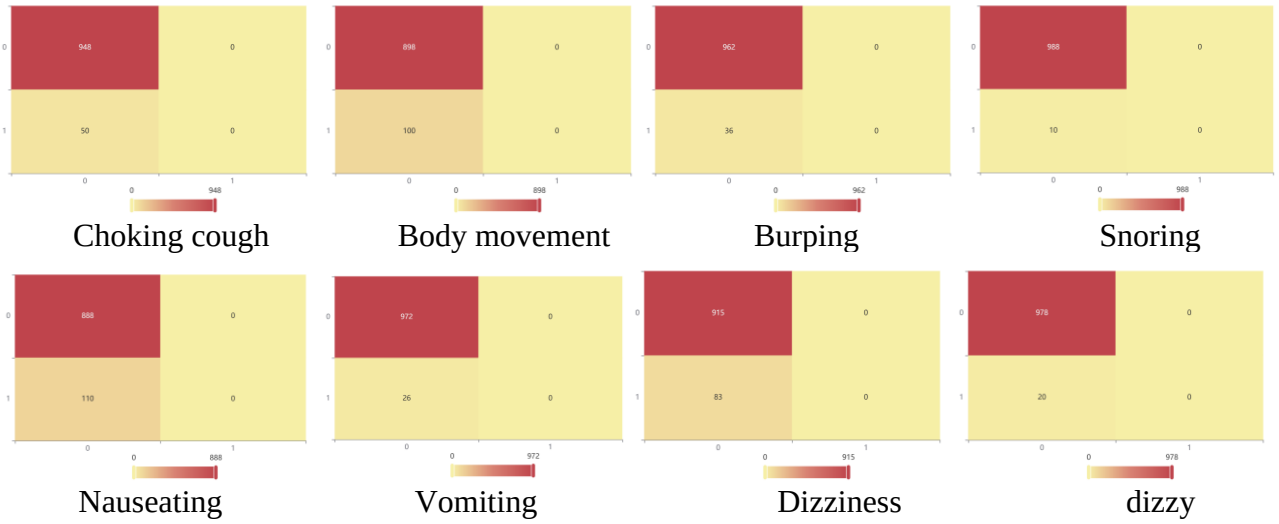
item	Regression coefficient	standard error	Wald	P	OR	Upper limit of OR value	95% lower confidence interval
constant	-2.873	1.000	8.109	0.004***	0.057	0.008	0.408
gender	0.647	0.525	1.517	0.218	1.91	0.682	5.348
age	0.037	0.855	0.002	0.966	1.038	0.194	5.542
height	0.76	1.858	0.168	0.682	2.139	0.056	81.538
weight	-1.925	1.447	1.769	0.183	0.146	0.009	2.489
History of surgery	-0.256	0.737	0.121	0.728	0.774	0.183	3.282
Have a past history	-0.203	0.594	0.117	0.732	0.816	0.255	2.612
Whether you smoke or not	0.183	0.304	0.36	0.548	1.2	0.661	2.178
Whether or not to drink heavily	-0.641	0.613	1.094	0.296	0.527	0.158	1.752
Sedative name	-2.349	1.019	5.319	0.021**	0.095	0.013	0.703
There is no additional sedation	-0.2	0.353	0.52	0.571	0.819	0.409	1.637
Dependent variable				Squirting			

Note: ****, **, * represent 1%, 5%, 1 respectively 0% significance level

5. model testing

5.1. The heat map of the confusion matrix

The confusion matrix heat map is used to show the performance of the classification model on different categories. It reflects the classification performance of the model by combining the predicted categories and the true categories into a two-dimensional matrix. Typically, rows in this matrix represent the true category and columns represent the predicted category. In the heat map of the confusion matrix, each cell is filled with a color to reflect information about the number or accuracy of samples under the corresponding category. Some thermal maps of the confusion matrix will also indicate the recall rate, precision rate, F1 score and other indicators, so as to evaluate the model performance more comprehensively. The heat maps of each indicator are as follows:



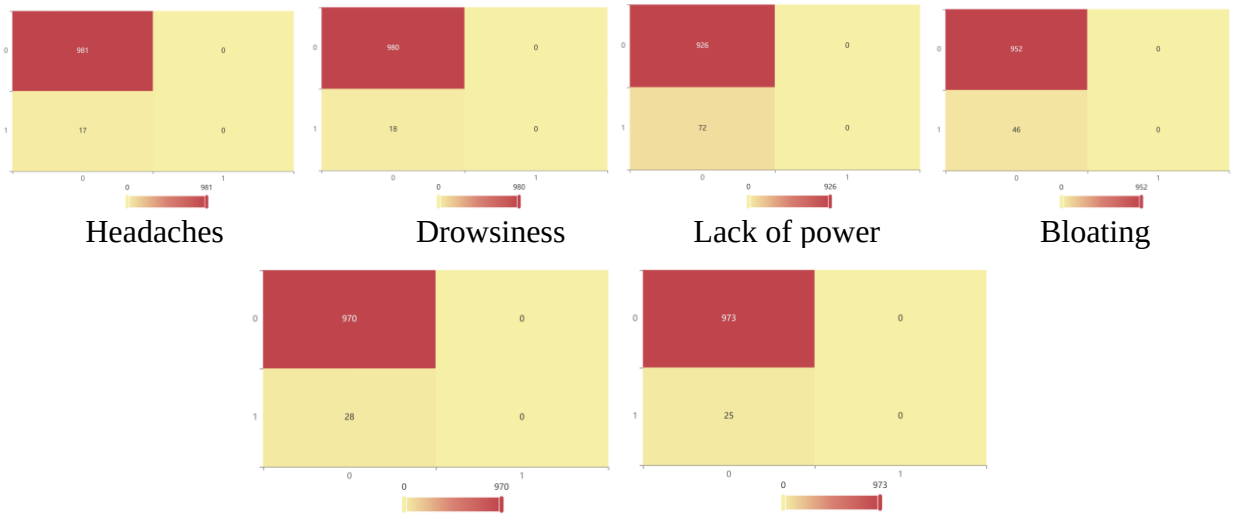


Figure 1. The heat map of the confusion matrix

By observing the color depth of the heat map, we can get that the two-taxonomic logic model in this paper is standard and reasonable.

5.2. ROC curve

The following figure shows the ROC graph, used to measure the classification effect of the logistic regression. The ROC graph combines sensitivity (TPR) and specificity (FPR) to measure the relationship simultaneously. Ideally, TPR should be close to 1 and FPR close to 0.

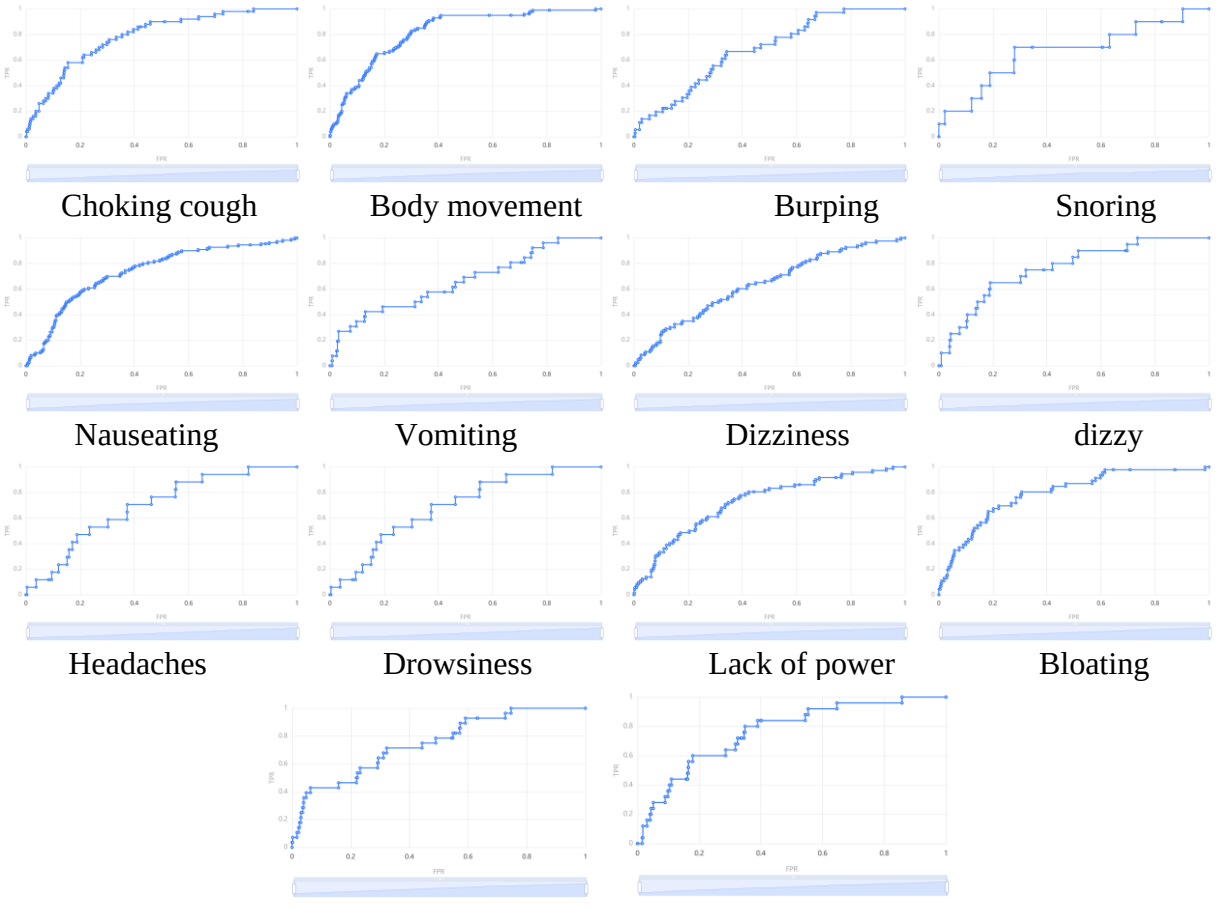


Figure 2. ROC curve

After calculating the AUC, we found that the AUC values were all close to 1, indicating that our assumed binary classification logic model performs well.

6. Nuclear density map model

A kernel density map is a visualization tool used to represent a probability density function of a set of data. These data often represent measurements of continuous variables of a certain phenomenon. The kernel density maps show the data by using a curve or a region. The nuclear density map can better display the main features of the data distribution and density, and can help us quickly find the peaks, outliers, and the overall shape in the dataset. It can also be used to compare the distribution between different datasets. In the following figure, the nuclear density map at 1,2,3 and 5 minutes after medication:

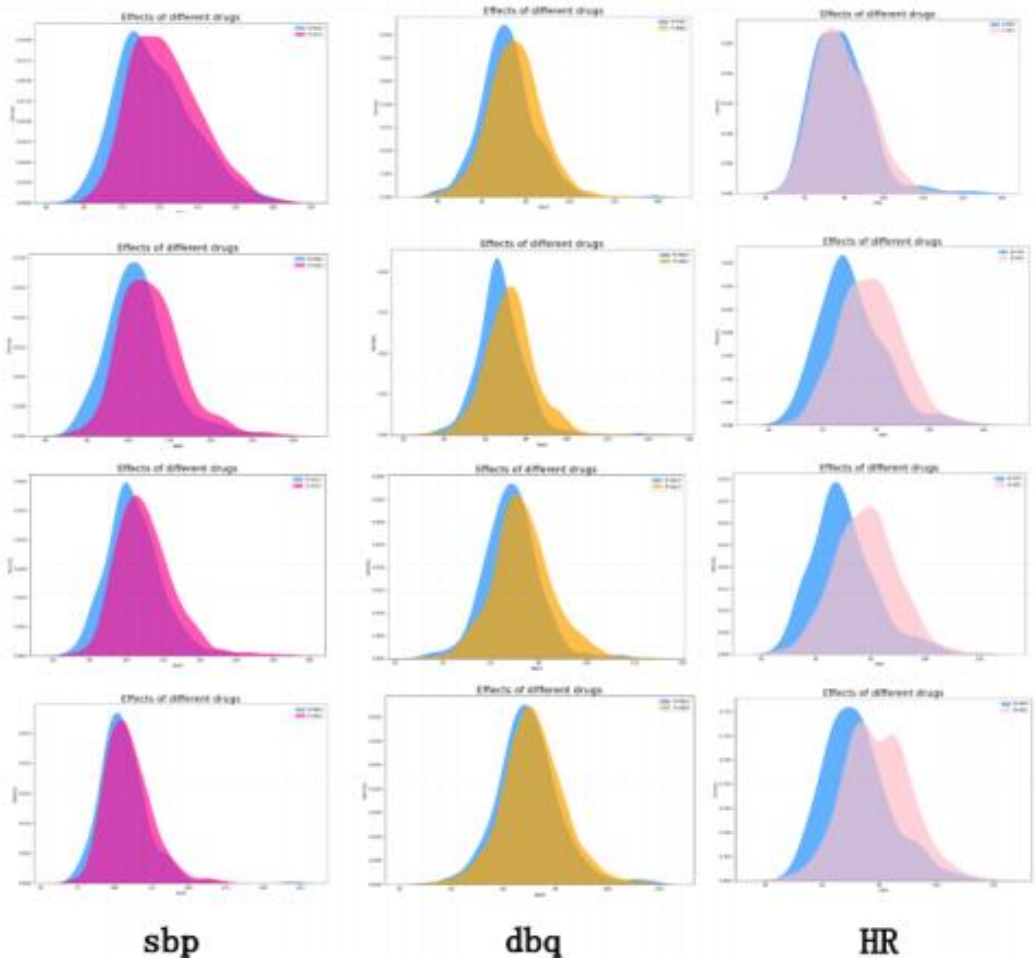


Figure 3. 1,2,3, and 5 min of nuclear density

Through this figure, we found that the vital signs data showed a non-normal distribution, and we predicted that the new drug group and the original drug group did show significant differences in the vital sign data.

7. A Mann – Whitney U test for independent samples

This is a non-parametric hypothesis test for comparing two independent samples, which is applicable where two independent samples do not satisfy a normal distribution or equal variances and two independent samples are from the same distribution.

The U threshold, based on the theoretical formula or with reference to the Mann-Whitney test table, was calculated as the minimum allowed U value, at the significance level. (The smaller the U value,

the greater the difference between the two groups of data.) Comparing the calculated U value with the U critical value; If the U value is less than the U critical value, the null hypothesis is rejected, and consider that the two independent samples do not come from the same distribution; otherwise, accept the null hypothesis.

The results of the independent sample Mann-Whitney U test usually output statistical indicators such as U-values, p-values and z-values. Where the p-value represents the magnitude of the difference between the two data groups, while the z-value is used to determine the direction of difference. If the p-value was less than the significance level, the two data sets were considered significantly different.

7.1. Sbp, dbp, SpO2, HR, IPI

Table 8. The P-value table of the Mann-Whitney U test

Sbp00	Dbp00	Sbp1	Dbp1	Sbp2	Dbp2	Sbp3	Dbp3	Sbp5	Dbp5
0.935	0.952	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.001***	0.003***
Spo200	Spo2005	Spo21	Spo2015	Spo22	Spo2025	Spo23	Spo25		
0.032**	0.048**	0.024**	0.004***	0.000***	0.000***	0.001***	0.229		
HR00	HR005	HR1	HR2	HR3	HR5				
0.904	0.968	0.164	0.000***	0.000***	0.000***				
IPI00	IPI005	IPI1	IPIjinjing	IPI015	IPI2	IPI025	IPI3	IPI5	
0.794	0.205	0.751	0.000***	0.000***	0.000***	0.002***	0.000***	0.000***	

According to the Mann-Whitney U test analysis table described above:

a. For the sbp and dbp metrics, from the p-values in the table, varying with time varying from 0 to 5 minutes, the two drugs showed significant differences under the sbp and dbp metrics.

b. For the SpO_2 index, from the p-value in the table, varying with time changes from 0 to 5 minutes, the two drugs showed significant differences under the SpO_2 indicator.

c. For the HR index, from the p-value in the table, varying from 0 to 5 minutes over time, the two drugs showed a significant difference under the HR indicator.

d. For the IPI index, from the p-value in the table, varying from 0 to 5 minutes over time, the two drugs showed a significant difference under the IPI index.

8. Conclusion

According to the chi-square test results, the symptoms of cough, movement, body, and hiccups between the new drug group and the original drug group, with no significant difference in postoperative adverse reactions, nausea, vomiting, dizziness, drowsiness, fatigue, abdominal pain, but dizziness, headache and other symptoms. Prediction of adverse reactions (regression equation):

$$y = -0.4272 + 0.723x_1 + 0.243x_2 - 1.96x_3 + 1.573x_4 - 0.716x_5 + 0.713x_6 + 0.017x_7 + 0.076x_8 - 1.533x_9 + 1.966x_{10} \quad (9)$$

Then we tested the prediction model of each adverse reaction by the XGBoost model [2], and found that the model worked well.

Through the Mann-Whitney U test of independent samples, we found that the two sets of data from medications B and R, Before the medication, the vital signs (sbp00, dbp00, SpO200, HR00, IPI 00); however, after the medication changed from 0 to 5 minutes, sbp, dbp, SpO_2, HR, and IPI. It is worth noting that the IPI index is the core indicator to depict the vital signs, and the two drugs showed significant differences in the five vital signs indicators including the IPI index. Therefore, we believe

that the new drug group and the original drug group showed a significant difference in the vital signs data, and the difference was due to the new drug [4].

9. Model evaluation

9.1. Advantage

The normalized dimensionless treatment improves the accuracy and accuracy of the model, improves the comparability and interpretability and reduces the data of the complexity of the data.

We found a non-normal distribution of data using the independent sample Mann-Whitney U test instead than T test.

Reasonable trade-off of data is used to avoid the impact of missing data and 0 value on the model.

9.2. Weakness

When using logistics regression models, the collinearity of the data can have an impact on the effect of the model.

References

[1] Wang Yufen, Zhang Chengzhi, Zhang Beibei. Review of data cleaning studies [J]. Modern Library and Information Technology, 2007: 54 - 60.

[2] Wang Kunzhang; Jiang Shubo; Zhang Hao; Chao Zheng. Regression-classification-regression lifetime prediction model based on XGBoost [J]. ChinaTest, 2022: 8.

[3] Wang Huiwen, Meng Jie. Predictive modeling approach to multivariate linear regression [J]. Journal of Beijing University of Aeronautics, 2007: 125 - 129.

[4] Shen Qu, Li Zheng, Gwen Sherwood. Research on the satisfaction status and influencing factors of pain control in patients after surgery [J]. centreChina Nursing Journal, 2007: 6 - 11.

[5] Wang Zhenyou, Chen Li'e. Application of a statistical prediction model for multiple linear regression [J]. Statistics and Decision Making, 2008: 48 - 49.