Clinical efficacy of new sedative drugs based on machine learning big data statistics Analysis and prediction

Zhuo Zeng [1], Hao Fang[2], Wenkang Fan[2], Zizhen Li[1], WeiJie Pan[1], Shixun Li[1]

(1: Xiamen University Jiageng College, Fujian, China

2: Xiamen University, Fujian, China)

ABSTRACT

In order to solve the analysis and prediction of the clinical efficacy of sedative drugs and the problems of adverse drug reactions, various methods of differential analysis, correlation analysis, regression prediction and data sampling are used to explore the connection between the clinical efficacy analysis and prediction and adverse reactions. The new drug group R and traditional drug group B were subdivided into intraoperative adverse reactions (cough, body movement and other intraoperative reactions) and 24 hours postoperative adverse reactions (dizziness, headache, drowsiness, fatigue, abdominal distension and abdominal pain, and other postoperative reactions). After Spearman medical statistical correlation analysis, the significance scale and correlation coefficient thermal map were obtained, and the important factors with P value <0.05 were selected. A normality analysis of the underlying signs of patients under different agents then confirmed that subject selection was without sample selection bias. Considering the paired relationship between the two sedative groups and the three corresponding vital signs (e. g., heart rate, pulse, blood pressure), the paired T test was used to assess significant differences. Finally, a multiple linear regression model was used to analyze the cause of the difference and predict the IPI within 3 minutes of drug administration based on medication information and patient information. R drugs had a positive effect on patient satisfaction, while the total dose of sedatives, the total dose of analgesics, anesthesiologist satisfaction, and endoscopist satisfaction had a negative impact on patient satisfaction. These three vital signs differ between the new drug and traditional drug groups. Based on the development of artificial intelligence in the new era, taking new drugs as the background and combining with medical clinical experiments, the application prospect of clinical experiments and points out the current shortcomings, aiming to provide basis for clinical experiments of new sedative drugs in the future.

Keywords: Clinical Experiment of New Sedative Drugs; Efficacy Analysis and Prediction; Machine Learning; Big Data Statistics; Analysis and Prediction

1 INTRODUCTION

In recent years, with the aggravation of the aging trend in China, the number of patients with gastrointestinal diseases continues to rise. Advanced and effective treatment plans and nursing methods are crucial for patients with gastrointestinal diseases, which can significantly improve the clinical treatment effect and prognosis of patients [1]. In minimally invasive gastrointestinal surgery, we need to use local sedative and analgesic drugs. The traditional sedative drug is "B drug", and a drug research and development center has developed a new drug "R drug". In order to fully understand the drug efficacy of new drugs, the hospital conducted clinical experiments on new drugs and collected data on new sedatives and traditional sedatives in clinical experiments [2]. Using sedative drugs during surgery, different adverse effects are likely, and sedative drugs also have a significant effect on vital signs. By comparing the occurrence of adverse reactions of new drugs and

traditional drugs, and evaluating the data of vital signs, the efficacy evaluation of new drugs can be obtained to provide a reference for drug selection [3]. And based on the era of artificial intelligence data background, mathematical statistical knowledge and machine learning integration algorithm, supervised learning thought data specimens of scientific classification, in line with the current society the direction of artificial intelligence and medical guidance strategy, will greatly reduce the clinical diagnosis of case sample analysis of the world, so as to improve the efficiency of clinical diagnosis. This question mainly studies the efficacy of new sedative drugs in clinical trials and predicts the possible consequences of drug use [4]. The aim is to grasp the efficacy of new sedative drugs, which has significant reference significance for subsequent drug development, improvement, and promotion of use. This study aims to develop a new type of analgesic drug and evaluate its analgesic effect and safety through clinical trials in response to the research and clinical trials of analgesic drugs. This study will adopt a combination of laboratory research and clinical trials to screen and optimize substances with analgesic effects, and ultimately determine new analgesics with clinical application potential. Analgesics have important clinical significance, but the commonly used analgesics in the market currently have some limitations, such as significant side effects and unsatisfactory efficacy. Developing a new type of analgesic drug has important research value [5]. The purpose of this study is to develop a novel analgesic drug through laboratory research and clinical trials, and to evaluate its analgesic effect and safety.

2 DATA SOURCES AND ANALYSIS METHODS

2.1 Data sources

The data of this study were statistically analyzed based on the survey results of surgical clinical patients in Zhongshan Hospital of Xiamen University in 2023 and combined with 7 related physiological indicators provided in question A of 2023 National College Students' Mathematical modeling Competition in 2023. Select a certain number of patients and randomly divide them into an experimental group and a control group. The experimental group patients were given new analgesics, while the control group patients were given conventional analgesics. Compare the differences in analgesic efficacy and safety between two groups of patients. Collect basic information about patients, including age, gender, medical history, etc. At the same time, data on the analgesic effect and adverse reactions of patients were collected through case record forms and questionnaire surveys. Use appropriate statistical methods to analyze clinical trial data and compare the differences in analgesic effects and adverse reactions between two groups of patients [6]. Based on the results of laboratory research and clinical trial data, this study will organize and analyze the collected data. Evaluate the analgesic effect and safety of new analgesics by comparing the differences between different groups.

2.2 Analysis methods

Test the effects of R and B on adverse reactions and predict the existing variables; this study will classify intraoperative and 24h postoperative adverse reactions according to type, pre-treatment through data cleaning and screening, and conduct differential analysis by chi-square test. Corresponding binary Logistic models were constructed for predicting the different adverse reactions. And to establish a reasonable analytical model to investigate whether there are significant differences between the new and original drug groups in different vital signs, and to try to determine whether this difference is due to the new drug [7]. The independent variable was set to the name of the sedative drug, and the dependent variable included seven vital signs, and the

significant differences were determined by analyzing the sign data corresponding to different drugs at different time points. Based on this, a differential analysis model was established to assess whether the magnitude of the differences between variables evaluated the significant effects. For "causality": the factors of the significant difference are not necessarily the cause of the phenomenon. On the basis of "correlation", other potential factors are introduced to establish regression models. The regression parameter values were calculated for the preliminary analysis. Then the significant correlation indicators were combined for secondary regression to determine whether the factors caused significant effects. Comprehensive analysis and comprehensive evaluation system were established. By constructing the difference analysis model and regression model, the drug name and different vital signs were determined according to the significance level and the regression parameter value. Predict IPI data within 3 minutes of dosing based on medication information and patient information. First, IPI data were collated for the whole procedure and found large deletions in IPI 15 and IPI 20. Since the whole process IPI data has influence on prediction within three minutes, considering only data within three minutes may introduce error because part of the information is lost. To address this issue, after cleaning the data, the IPI data within ten minutes was processed using the weighted averaging method to obtain a representative indicator. Then, partial least squares regression was used to predict the IPI data within three minutes [8]. Moreover, the BP neural network model was used to evaluate the prediction model by calculating the mean square error of the partial least squares regression model. Based on the existing data, patient information, operation time, medication dose and other data were screened, potential factors were selected for Spearman correlation analysis, the influencing factors of patient satisfaction were judged according to the significance level, and the relationship with postoperative satisfaction was analyzed according to the correlation coefficient.

3 EXPERIMENT PROCESSING

Increase "postoperative satisfaction evaluation", the "postoperative within 24 patients satisfaction evaluation very satisfied" into "1", "postoperative 24 patients satisfaction evaluation satisfaction" into "2", "general" postoperative 24 patients satisfaction evaluation "into" 3 "," postoperative 24 patients satisfaction evaluation unsatisfactory "into" 4 "," postoperative 24 patients satisfaction evaluation is very unsatisfactory "into" 5 ".

Add "total time", the calculation method is "camera time" minus "camera time". Add "total operation time", calculated as "PACU" minus "start time". The words "points" in "anesthesiologist satisfaction" and "endoscopist satisfaction" were removed and only the numbers were retained.

Spearman Correlation analysis based on the available data, eight potential influencing factors of patient satisfaction were selected from the aspects of patient information and surgery-related indicators. They are gender, height, age, name of sedative drug, total dose of sedative drug, total dose of analgesics, satisfaction of anesthesiologist, and satisfaction of endoscopist [9]. The scale and the correlation heat map were obtained by Spearman correlation analysis as follows:

4 EXPERIMENTAL RESULT

4.1 Logistic Regression model establishment

In this paper, a binary Logistic regression model was constructed to predict whether adverse reactions would occur during and 24h after surgery. Taking the focused adverse reaction (nausea and vomiting) as an example, we discarded the variables with P value greater than 0.05 and obtained the following results:

Table1: Binary Logistic regression results

Item	Regression Coefficient	Standard Error	Wald	P
B_0	-3.153	0.33	96.279	0.000
Gen	0.379	0.198	3.859	0.051
Smo	-0.439	0.221	4.081	0.044
RB	0.631	0.181	12.714	0.000

Model evaluation: Conduct likelihood chi-square test of the model. $P < 0.05$ rejected the null hypothesis; the accuracy, recall and precision are above 75%, indicating that the constructed model can better predict the occurrence of postoperative adverse reactions (nausea and vomiting).

In the construction of adverse reaction prediction models, the method of integrating models is used to predict adverse reactions, integrating the advantages of multiple models and effectively improving the prediction accuracy of the models. The use of data rank for analysis reduced the impact of outliers and outliers, and screened for significant differences in vital sign indicators, making regression analysis results more accurate [10]. In predicting IPI data, multiple models and evaluation metrics are used to select the optimal model, making the prediction more accurate.

In building an integrated model, the amount of data needs to be large enough, otherwise the results will be distorted. At this time, by merging some categories, the number of samples for each category can be increased to be closer to the situation where the sample size is large enough. When there are a large number of identical values in the problem, making predictions can lead to information loss and affect the selection of life indicators. In this case, the same values can be combined into one rank and the average rank can be calculated. This can avoid losing too much information.

5 CONCLUSION

This experiment comprehensively uses Pearson, chi-square detection, paired T detection and other significant detection models. Based on the basic principles of big data statistics, it analyzes the influence relationship between different factors from different focuses, and analyzes more angles to make the detection results more accurate. Using the method of sampling survey, the significance level and effect size of the average grouping data of patients are tested, and the sample data of partial least squares regression model are used to classify the sample data. Based on the small sample data capacity, multicollinearity can be processed. The partial least squares regression model concentrates the characteristics of principal component analysis, canonical correlation analysis and linear regression analysis to make the model test more reasonable. The weighted average will predict the IPI at 3min after medication, making the results more generally regular. Using machine learning random forest (Random Forest) algorithm realizes the average group screening of massive patient data, and establishes the binary regression model to select the relevant grouping of patients with postoperative adverse reaction control CP (<0.05); effectively reduces the experimental error. The core

idea of this experiment is to use machine learning feature engineering screening, using supervised learning and integration algorithm, using logical regression (Logistic Regression) to fit sample parameters, set preoperative postoperative, take B drugs and R drug group control experiment, more accurate and powerful proved the new sedative drugs (R) for gastrointestinal surgery, provide many convenience for medical clinical diagnosis and treatment.

This experimental study is committed to the development and effect prediction of new sedative drugs, in order to adapt to the use of sedative drugs in gastrointestinal clinical surgery, especially for special groups, has a far-reaching impact and research significance. The innovation of the experiment lies in the research and development of new drugs for clinical diagnosis of gastrointestinal medicine, and the topic selection perspective is novel and unique. In addition, big data statistical analysis and artificial intelligence were deeply used to compare the clinical efficacy of new drugs, thus proving that the new drugs play a significant role in the clinical treatment of gastrointestinal tract, and promoting the development of clinical medicine and pharmaceutical field.

What this experiment needs to be improved is that the detection value of vital signs in the paired T-test is related to the basic information of human body. It is considered that the average value of the sedative drug name is not rigorous. Here, the distribution of the corresponding data sets of the two drugs is approximately the same, and no rigorous proof is made. It is hoped that in the future, under the guidance of artificial intelligence, clinical medical surgery can be improved, with better, more cost saving and reduce side effects, which will play a vital role in the whole clinical experiment and the improvement and progress of the medical system.

The clinical trial efficacy analysis and prediction of the two sedative drugs have important practical significance. The model proposed in this article can provide reference and suggestions for clinical doctors to better prevent and reduce adverse reactions when using sedatives, enabling them to better grasp the changes in various vital signs, adjust the dosage and type of sedatives in a timely manner, and achieve better clinical treatment effects.

In addition, the methods and models of this study can also provide an effective research approach and approach for related disciplines, such as pharmacy, statistics, and data mining. We can refer to the data preprocessing methods and model frameworks we use, as well as experience in model evaluation indicators.

Finally, we hope that the research findings of this article can provide useful references and insights for the clinical treatment and drug development of related diseases, and make a modest contribution to the cause of human health.

REFERENCES

[1] SPSSPRO. (2021). Scientific Platform Serving for Statistics Professional (Version 1.0.11) [O nline Application Software]. Retrieved from https://www.spsspro.com

[2] Mao, S., Wang, J., Pu, X., et al. (2006). Advanced Mathematical Statistics (2nd ed.). Beiji ng: Higher Education Press.

[3] SPSSPRO. (2021). Scientific Platform Serving for Statistics Professional (Version 1.0.11) [O nline Application Software]. Retrieved from https://www.spsspro.com

[4] Xu, W. (2012). Review of the correlation coefficient studies. Journal of Guangdong Univ ersity of Technology, 29(3), 12-17.

[5] Hong, M. (2021). Effect of postoperative care intervention on early functional recovery i n patients undergoing gastrointestinal surgery. Chinese Medical Guide, 19(35).

[6] Shangguan, W. G. (2021). Clinical study of dexmedetomidine and oxycodone on agitatio n during general anesthesia in elderly patients undergoing gastrointestinal surgery. Rura l Health in China, 13(03).

[7] Ma, Q.,$ & Cao$, Q. (2018). Effect of systematic care in patients undergoing gastrointest inal surgery. Clinical Medicine Research and Practice, 3(34).

[8] Tian, S.,$ & Zhang$, Y. (2023). A review of machine learning algorithms and their appli cations. Software, 44(07).

[9] Lu, G., Zhu, Y.,$ & Xin$, Y. (2008). Difference between statistical significance and actual significance. Statistical Education, 06.

[10] Wang, Z., Han, L., Li, X., Wang, X., Zhang, Z.,$ & Wang$, J. (2023). Progress in machi ne learning in drug screening. The Chinese Journal of Hospital Pharmacy, 09-22.