Logistic Model and Gradient Boosting Machine Model for Physical Therapy of Lumbar Disc Herniation

Ping Zhao, Jin Xue, Xiaomei Xu, Lifei Wang, and Dan Chen

1 Nursing Department, The Third People’s Hospital of Hefei, Hefei Third Clinical College of Anhui Medical University, Hefei 230022, China
2 Department of Orthopaedics, The Third People’s Hospital of Hefei, Hefei Third Clinical College of Anhui Medical University, Hefei 230022, China
3 Department of Neurosurgery, The Third People’s Hospital of Hefei, Hefei Third Clinical College of Anhui Medical University, Hefei 230022, China

Correspondence should be addressed to Lifei Wang; wlf680802@sina.com and Dan Chen; cd810424@126.com

Received 21 January 2022; Revised 26 March 2022; Accepted 19 April 2022; Published 11 May 2022

Academic Editor: Kelvin Wong

Copyright © 2022 Ping Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objective. Physical therapy is a common clinical treatment for patients with lumbar disc herniation. The study is aimed at exploring the feasibility of mathematical expression and curative effect prediction of physical therapy in patients with lumbar disc herniation using a logistic model and gradient boosting machine (GBM). Methods. A total of 142 patients with lumbar disc herniation were treated with physical therapy. The pain was evaluated by the visual analogue scale (VAS) before each treatment. The logistic model was used to conduct a global regression analysis on patients with lumbar disc herniation. The final results of the whole course of treatment were predicted by the measured values of 2-9 times of treatment. The GBM model was used to predict and analyze the curative effect of physical therapy. Results. The mathematical expression ability of the logistic regression model for patients with lumbar disc herniation undergoing physical therapy was sufficient, and the global determination coefficient was 0.721. The results would be better for more than five measurements. The AUC of GBM model logistic regression analysis was 0.936 and 0.883, and the prediction effect is statistically significant. Conclusion. Both the logistic and GBM model can fully express the changes in patients with lumbar disc herniation during physical therapy.

1. Introduction

Lumbar disc herniation, the most common degenerative spine disease, can cause low back pain and sciatica, causing great distress to patients’ daily life [1]. Timely physical treatment can achieve traction disc, joint loosening, and other effects to relieve the patient’s lumbar pain and sciatica. Physical therapy has many advantages, such as easy and reliable efficacy. Thus, it is widely used in the clinic and plays an important therapeutic role [2]. In clinical rehabilitation practice, how to optimize the prescription and further improve the efficacy is one of the important topics of concern for clinical rehabilitation workers. There are numerous reports on analyzing rehabilitation efficacy in lumbar disc herniation. However, due to interindividual variability and the diversity of physical therapy, physical therapy prescriptions often have distinct individualized characteristics, which also poses some difficulties for efficacy prediction [3]. In clinical practice, predicting efficacy by an investigational treatment is widely adopted. In this process, three aspects have become the key to optimizing the configuration of medical resources and improving the efficiency of rehabilitation treatment: how to determine the appropriate length of treatment time, accurately estimate the efficacy of the entire course of therapy; timely termination of the treatment regimens that may be less effective [4–9].

Machine learning is a multidomain interdisciplinary emerging over the last 20 years, involving several disciplines such as probability theory, statistics, approximation, convex analysis, and algorithmic complexity theory [10–14]. It is an essential branch of computer science mainly used to model complex relationships between predictor and response...
variables and has the potential to transform modern methods of epidemiological research. With "big data" attracting much attention, machine learning can provide epidemiologists with new tools and ideas for solving problems that traditional methods cannot solve [15]. The gradient boosting machine (GBM), a machine learning algorithm for regression and classification, can train many models in order. Then, each new model updates the prediction using the gradient descent method [16–20]. GBM has shown great success in a wide range of practical applications and can be highly customized to the application’s specific needs, gradually minimizing the loss of function. The logistic regression model is a classical linear regression model that can mine the linear relationship before variables and establish an effective model to predict future variable results.

The logistic curve model was used for the mathematical expression on the treatment process of cervical radicular pain. The prediction of the effect of physical therapy on cervical radicular pain was achieved by further mathematical feature analysis and predictive ability analysis [21–23]. During physical therapy, patients with cervical spondylosis show a closer course of remission than patients with lumbar disc herniation. Based on the above background, this study attempted to apply the GBM model and logistic model to describe the process of physical therapy and the prediction of efficacy in patients with lumbar disc herniation.

2. Material and Method

2.1. General Data. A total of 142 patients with lumbar disc herniation, aged 17 to 87 years (53.1 ± 15.2) years, who were treated at the Department of physical medicine and rehabilitation of our hospital from December 2018 to June 2021, were selected, including 89 males and 53 females. Inclusion criteria were as follows: (1) they met the diagnostic criteria of lumbar disc herniation according to the diagnosis and treatment code of China Rehabilitation Medicine based on the symptom presentation, physical examination, and auxiliary examinations, accompanied by radicular radiation pain; (2) they can understand and apply visual analog scale for pain rating; (3) they can adhere to completion of more than ten physical therapy sessions; (4) the patient’s condition was stable.

The patient after admission determined the corresponding physical therapy prescription according to his actual condition. The treatment contents included traction, joint loosening, and physical factor therapy. These treatments were used once daily for five consecutive days each week, ten times for one course. The visual analogue scale (VAS) pain assessment was performed simultaneously before each treatment. The specific method is to draw a 10 cm long line segment in front of the patient and inform the patient that the left end of the line segment represents no pain, and the right end of the line segment represents intolerable severe pain and asked patients to label the average degree of pain in the past 24 h on the line segment. The researcher then used a ruler to measure the distance from the left endpoint of the line segment to the marked point and recorded each patient to the end of the course to obtain ten pain VAS score values. Throughout treatment, every effort was made to guarantee that each patient’s treatment time point was relatively fixed.

2.2. Mathematical Expression and Predictive Power Analysis. Using the logistic model as a regression analysis model, y represents the degree of pain during treatment, and x represents the time point to complete treatment. The corresponding x values at the 10 measurement time points in this study are 1, 2, ..., 9, 10. The a, b is the parameter, and the H is the distance between the upper and lower asymptotes representing the maximum possible relief during treatment as determined by the pain level at the initial visit. According to the vas minimal clinically significant difference (MCSD) theory proposed by Kelly [24], the position of the upper asymptote is defined as the level 1.2 above the 1st pain score value, i.e., $K = 1st \text{VAS} + 1.2$. Since the maximum value of the VAS score is 10, the $K$ value was taken as 10 when the $\text{VAS} \geq 8.8$ at the first assessment.

To explore the predictive ability of different treatment duration, the predictive models were constructed with the measurement results of 1st to 2nd, 1st to 3rd, 1st to 4th, 1st to 5th, 1st to 6th, 1st to 7th, 1st to 8th, and 1st to 9th. The logistic model curve fitting analysis were completed. The predicted value of each model for the degree of pain at the 10th rating time point was then evaluated, and the difference between this expected value and the actual measurement was calculated.

Based on the actual measured values and MCSD theory, cases with a decrease of more than 1.2 VAS score on the whole course were defined as clinically effective cases, and the cases with less than 1.2 points were defined as invalid cases. The magnitude of decline predicted by each model was calculated simultaneously, and cases with a decrease of more than 1.2 points were defined as predicted valid cases. Cases with a decrease of less than 1.2 points were defined as predicted null cases. When the expected outcome was consistent with the actual measured outcome, it was defined as the predicted successful case and vice versa as the expected unsuccessful case. The predictive ability of different predictive models was compared between groups.

Remission magnitude was defined as predicting a successful case and vice versa an unsuccessful case when the difference between the predicted value and the actual measured last score value was within 1.2 points. The success rate predicted by the remission magnitude was calculated for each length of treatment, and the difference between the models in predicting remission magnitude was compared.

2.3. GBM Model Construction. Data acquisition, data preprocessing, training, model evaluation, and prediction for GBM models were implemented using Python 3.8 (https://www.python.org/downloads/Guido van Rossum). The data were randomly divided into training and test sets in a 7:3 ratio, and the optimal parameters were trained in the GBM model using the data from the training set, and the simple algorithm for GBM proceeds as follows:

Step 1. Construct a training dataset \( \{y_i, x_i\}^N \) trained on N samples, where \( y \) is the outcome variable and \( x \) is the
3. Result

3.1. Efficacy Analysis. Before treatment, the pain VAS score of 142 patients with lumbar disc herniation in this study was 6.8 (3.0-10.0) points, and the posttreatment pain VAS score was 4.3 (0.1-7.8) points. After physical therapy, the patient’s pain was significantly relieved. The symptoms of lumbar disc herniation were also considerably improved by comparing the MRI of the lumbar disc before and after treatment (Figure 1).

3.2. Mathematical Expressivity Analysis. The regression curves of each case showed some differences, and the regression curves of some patients showed "S" type characteristics (Figure 2(a)). Some additional patients showed regression lines that resembled straight lines (Figure 2(b)). The coefficient of determination of the 142 patients in this study was 0.731 (0.001-0.979). The global coefficient of determination was 0.721, suggesting that this model can better express the course of physical therapy for patients with lumbar disc herniation.

3.3. Effective Case Prediction Power Analysis. The predictive ability of the length of treatment for effective vs. ineffective cases gradually improved with the number of treatment sessions, and specific data are detailed in Table 1. The difference between them was statistically significant by the chi-square test for the data in the table. Their success rates were all predicted to be above 80% based on validity completed with more than five measurements, and they were relatively close to each other. However, the predictions obtained based on 2-4 measurements are relatively close to each other and have low predictive power. After merging the cells, the predicted outcomes of measurements more than 5 times were compared with those of measurements within 4 times, and the difference between them was found to be statistically significant.

3.4. Remission Magnitude Predictive Power Analysis. The predictive ability of the remission range of the whole course of treatment varied in different treatment duration. With the increase in treatment times, the difference between the predicted and measured values gradually decreases. The prediction success rate of remission amplitude obtained by each prediction model increases with treatment duration (Table 2). The predictive power of 2-4 measurements was low, and all of them had less than 50% correct. The accurate rates of prediction results of 5-7 measurements were all at 60% sufficient, and the accurate rates of 8-9 measurements were relatively high. All of them reached more than 80%.

According to the characteristics of prediction ability, the treatment duration was combined into 2-4 times, 5-7 times, and 8-9 times. It was found that the difference between the groups of 2-4 times and 5-7 times was statistically significant (P < 0.01). The difference between the groups of 8-9 times and the group of 5-7 times was also statistically significant (P < 0.01).

3.5. Prediction Model of Short-Term Prognosis Based on GBM. The optimal parameters of GBM model were obtained through grid search and manual debugging. The final parameters are adjusted to (learning_rate = 0.1, n_estimators = 1000, max_depth = 4, min_samples_split = 2, min_samples_leaf = 1, subsample = 1, max_features = 'sqrt', random_state = 10). The test set data verify the prediction model constructed by the training set. The influencing factors in the GBM model were ranked by importance, and
the top five are the sagittal diameter of protrusion, the degeneration level of the surgical segment, age, the time from initial symptom to operation, and the degeneration level of the adjacent segment.

The area under the ROC curve (AUC) of the GBM prediction model was 0.936 [95% CI (0.782, 0.973)]. The sensitivity, specificity, and Youden index were 93.3%, 87.5%, and 0.808, respectively. The AUC of logistic regression analysis was 0.883 [95% CI (0.756, 0.914)]. The sensitivity, specificity, and Youden index were 88.6%, 67.8%, and 0.624, respectively. The prediction effects of the two prediction models were statistically significant ($P < 0.001$) (Figure 3).

4. Discussion

Rehabilitation therapy plays an essential role in many treatment fields, including bone and joint system diseases, nervous system diseases, and circulatory and respiratory...
In the actual rehabilitation treatment, it is found that the true efficacy of rehabilitation prescription can be understood to a certain extent by observing the patient’s treatment feedback information, which can predict the treatment effect [29, 30]. However, how to choose a reasonable length of treatment time so that we can save medical resources to the greatest extent terminates the treatment scheme with poor curative effect in time, accurately predicts the curative effect, and prevents unnecessary prescription adjustment. Thus, we need to have a sufficient understanding of the characteristics of the treatment process. Through the long-term observation of the treatment process, this study proposed to use the logistic curve model and GBM model to predict and analyze the curative effect of physical therapy and nursing of lumbar disc herniation.

In clinical treatment, the remission process of different patients is different. The main difference is the remission speed and onset time [31, 32]. A mathematical model that can fully reflect the above characteristics must be constructed [33–35]. The logistic model is an “S” curve with upper and lower asymptotes. In the model construction, we pay special attention to the parameter design: for example, the characteristics of onset time and descent rate are fully expressed by parameters $a$ and $b$. Relevant research shows that GBM has a strong classification ability in many machine learning. GBM is an improved forward learning machine learning enhancement algorithm for regression and classification problems. Its fundamental theory is to produce a prediction model composed of weak learners or essential learners and gradually obtain approximately accurate prediction results from vulnerable learners to solid learners.

From the analysis of mathematical expression ability, the logistic regression model is satisfied with the expression ability of the physical treatment process of lumbar disc herniation. The global determination coefficient is 0.721. In the analysis of the predictive ability of different treatment practice lengths on the efficacy of the whole course of treatment, with the extension of treatment time, both effectiveness prediction and remission amplitude prediction show a good change prediction trend. In the effectiveness prediction analysis, when the forecast is made according to the measured values more than five times (because the pain evaluation is completed before each treatment, that is, more than four experimental treatments), the prediction ability is close. The accuracy rate is stable at more than 80% in the prediction analysis of lumbar disc herniation. When 8-9 measurements are used for prediction in the mitigation amplitude analysis, the accuracy is more than 80%. GBM usually takes the weak prediction model as the benchmark model to generate a complex prediction model. It constructs the model in a phased manner and optimizes it by selecting any differentiable function as the loss function, which can better fit the data. However, there are still some limitations to this study. The sample size of this study was not large enough, which may cause the overfitting of the model. Moreover, the model’s prediction accuracy requires the physical therapy process of patients many times, the experimental data are more strictly adopted, and subsequent studies can simplify the model and take a more convenient approach.

### Table 2: Analysis of ability to predict remission amplitude for different treatment time lengths (cases, %)

<table>
<thead>
<tr>
<th>Length of treatment</th>
<th>Successful prediction</th>
<th>Unsuccessful prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st–2nd</td>
<td>46/32.4</td>
<td>96/67.6</td>
</tr>
<tr>
<td>1st–3rd</td>
<td>59/41.6</td>
<td>83/58.5</td>
</tr>
<tr>
<td>1st–4th</td>
<td>64/15.1</td>
<td>78/64.9</td>
</tr>
<tr>
<td>1st–5th</td>
<td>93/65.5</td>
<td>49/34.5</td>
</tr>
<tr>
<td>1st–6th</td>
<td>82/57.7</td>
<td>60/42.3</td>
</tr>
<tr>
<td>1st–7th</td>
<td>92/64.8</td>
<td>50/35.2</td>
</tr>
<tr>
<td>1st–8th</td>
<td>114/80.3</td>
<td>28/19.7</td>
</tr>
<tr>
<td>1st–9th</td>
<td>124/87.3</td>
<td>18/12.7</td>
</tr>
</tbody>
</table>

**Figure 3:** ROC curve of GBM model and logistic regression analysis model.
5. Conclusions

In conclusion, it is feasible and accurate to use the logistic curve model and GBM model to predict rehabilitation treatment’s mathematical expression and curative effect for patients with lumbar disc herniation. The treatment effectiveness and remission range can be effectively expected by analyzing more than five measured values obtained from more than four experimental treatments.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

All authors declare no conflicts of interest in this paper.

References


patients with frailty complicated by intractable dizziness,” *Neuropeptides*, vol. 90, no. 1, article 102189, 2021.


[34] Z. Guo, L. Meng, S. Tian et al., “A five-parameter logistic model to predict the possibility of misdiagnosis for improving the specificity of Lugol chromoendoscopy in the diagnosis of esophageal neoplastic lesions,” *Frontiers in Oncology*, vol. 11, no. 1, article 763375, 2021.