87687_Auto_Edited.docx

Retrospective Study

Development of a machine learning-based model for predicting the risk of early postoperative recurrence of hepatocellular carcinoma

Machine learning-based Model of HCC

Abstract

BACKGROUND

Surgical resection is the primary treatment for HCC. However, studies indicate that nearly 70% of patients experience HCC recurrence within five years following hepatectomy. The earlier the recurrence, the worse the prognosis. Current studies on postoperative recurrence primarily rely on postoperative pathology and patient clinical data, which are lagging. Hence, developing a new pre-operative prediction model for postoperative recurrence is crucial for guiding individualized treatment of HCC patients and enhancing their prognosis.

AIM

This study aims to identify key variables in pre-operative clinical and imaging data using machine learning algorithms to construct multiple risk prediction models for early postoperative recurrence of hepatocellular carcinoma (HCC).

METHODS

The demographic and clinical data of 371 HCC patients were collected for this retrospective study. These data were randomly divided into training and test sets at a ratio of 8:2. The training set was analyzed, and key feature variables with predictive

value for early HCC recurrence were selected to construct six different machine learning prediction models. Each model was evaluated, and the best-performing model was selected for interpreting the importance of each variable. Finally, an online calculator based on the model was generated for daily clinical practice.

RESULTS

Following machine learning analysis, eight key feature variables (age, intratumoral arteries, alpha-fetoprotein (AFP), pre-operative blood glucose, number of tumors, the glucose-to-lymphocyte ratio (GLR), liver cirrhosis, and pre-operative platelets) were selected to construct six different prediction models. The XGBoost model outperformed other models, with the area under the ROC curve in the training, validation, and test datasets being 0.993 (95%CI: 0.982-1.000), 0.734 (0.601-0.867), and 0.706 (0.585-0.827), respectively. The XGBoost model also had an excellent calibration curve and DCA, indicating good predictive performance and clinical application value.

CONCLUSION

The XGBoost model exhibits superior performance and is a reliable tool for predicting early postoperative HCC recurrence. This model may guide surgical strategies and postoperative individualized medicine.

Key Words: Machine Learning; Hepatocellular carcinoma; Early recurrence; Risk prediction models; Imaging features; Clinical features

Zhang YB, Bo Y, Lei P, Yang G, Zhang W, Zhang D. Development of a machine learning-based model for predicting the risk of early postoperative recurrence of hepatocellular carcinoma. *World J Gastroenterol* 2023; In press

Core Tip: This current study aimed at employing machine learning techniques to select imaging and preoperatively clinical characteristic variables, to which the clinicians were

easily accessible, to develop six different risk prediction models for early postoperative recurrence of HCC. We compared the sensitivity and specificity of these models in detecting patients at high risk of early postoperative recurrence of HCC. In addition, to increase the feasibility and applicability of the constructed model, we generated a calculator online based on the predictive model to help clinicians apply it in their daily medical practice.

INTRODUCTION

Compared to other models, the XGBoost model demonstrated superior performance and emerged as the best predictive measurement model. This predictive model is easily accessible in daily clinical practice and may serve as a crucial tool in guiding postoperative follow-up and individualized medicine for HCC patients.

MATERIALS AND METHODS

Early HCC recurrence is one of the main factors shortening the survival of postoperative HCC patients. Given that most HCC patients experience HCC recurrence within 5 years post-resection of their primary tumors [2], prevention, early diagnosis, and precise treatment will be critical for limiting postoperative recurrence and improving the survival of HCC patients. This study found an early HCC recurrence rate of 59.02% in this population. While most cases with early HCC recurrence usually come from the hidden micrometastasis of the initial HCC in the body, the late HCC recurrence appears two years after resection. It is commonly from new polyclonal origins of HCC lesions that may completely differ from the initial liver cancer. Actually, early HCC recurrence is usually associated with more aggressive features and a worse prognosis. A previous study has suggested that adjuvant therapy for HCC patients at risk of early recurrence may prolong their survival [17]. However, identifying the patients at high risk of postoperative recurrence before surgery is complicated because of the lack of specific tumor biomarkers and clinical features. Therefore, the clinical prediction model will be valuable for predicting early HCC recurrence. However, most

available clinical prediction models are often constructed based on the assumption of linear relationships and usually fail to reflect the nature of complex, multidimensional, nonlinear relationships between different predictor variables in the pathogenesis of HCC to affect their predictive performance. Machine learning has shown great potential in analyses of data with nonlinear relationships, and in general, machine learning models outperform traditional regression models in predicting the development of HCC [18].

In this study, we analyzed pre-operative imaging and clinical indicators that were easily accessible and non-invasive and selected eight important feature variables using machine learning algorithms, including pre-operative blood glucose value, platelet count, intratumoral artery mass, methemoglobin, cirrhosis, GLR, age, and number of tumors. Subsequently, we generated six machine learning prediction models and validated their values in predicting the risk of early postoperative HCC recurrence. After randomly dividing the independent test set, we applied hyperparameter tuning in the training dataset with five-fold cross-validation and validated them in the test dataset. The results indicated that the XGBoost model exhibited better discriminative and predictive values than other models constructed. The calibration plots for the XGBoost model also displayed an excellent consistency between predictions and actual observations, and the DCA unveiled that the XGBoost model exhibited excellent clinical application. The model was, therefore, selected as the optimal model with an optimal cut-off value of 55% for the prediction probability. If a patient has a predicted probability of greater than 55%, the patient will have an increased risk of early HCC recurrence after primary tumor resection. Therefore, the prediction model may be valuable for predicting HCC patients for their risk of early postoperative HCC recurrence. Accordingly, to make the prediction model more convenient for clinical practice, we generated the model-based calculator online for physicians to use this model to predict early postoperative HCC recurrence. This will significantly improve the efficiency of the practical application of the prediction model.

To ensure better performance and clinical interpretation of the predictive model, we used the shape model to analyze its interpretation after determining the importance of the relationship between each characteristic variable and early postoperative HCC recurrence. The results indicated that the pre-operative blood glucose value was the variable with the highest importance, consistent with previous studies [19-21]. Hyperglycemia has now been recognized as a risk factor for the development and prognosis of HCC. Shi DY *et al* [19] found that HCC cells employed glycolysis to support their rapid growth in hypoxic and oxidative stress in the body. On the other hand, hyperglycemia can enhance the growth, invasion, and metastasis of HCC cells by promoting the expression of vascular endothelial growth factor (VEGF) [20]. Similarly, blood platelet counts also influenced the early HCC recurrence more, possibly by a positive feedback regulation [22]. Elevated pre-operative blood platelet counts are independently associated with increased tumor load, extrahepatic recurrence and metastasis, and poor prognoses of HCC [23]. In turn, HCC cells can promote platelet generation by secreting platelet-generating hormones, and the increased blood platelet counts further provide favorable conditions for HCC cell proliferation and metastasis [24, 25]. Intratumoral arterial features on pre-operative imaging are a significant risk factor for early postoperative HCC recurrence. A previous study has shown that the distribution and number of blood vessels supplying the tumor can impact the prognosis of HCC patients [26]. Multiple small arteries supplying blood to the tumor may reflect high pressure inside the tumor. The high-pressure environment tends to cause the tumor cells to shed and form microvascular invasion (MVI) in the surrounding vessels, which may even spread to other locations in the liver, increasing the risk of early postoperative HCC recurrence. Moreover, inflammatory factors are essential for the tumor microenvironment, and high levels of γ-glutamyl transferase (GGT) will lead to high cysteine levels in vivo, which influence tumor development regarding redox regulation and drug resistance [27]. GGT is a new potential prognostic indicator of inflammation in HCC [28], while lymphocytes are crucial for the systemic and local immune responses, including anti-tumor immunity [29]. A decrease in the number of

lymphocytes can result in a weakened immune function, which may enhance angiogenesis and extracellular matrix remodeling, creating a mutagenic environment for tumor cells and promoting tumorigenesis and development [30]. Indeed, GLR has shown to be valuable for predicting early postoperative HCC recurrence in a population of 606 patients [31]. Therefore, these vital feature variables included in the best mode were critical for predicting early postoperative HCC recurrence.

In summary, the predictive model constructed in this study may help clinicians predict the risk of early postoperative HCC recurrence in individual HCC patients and better manage them before, during, and after surgery. Our findings may open new avenues for investigating therapeutic strategies for postoperative recurrence of HCC.

We acknowledge that the current study has limitations. First, our study was conducted with a relatively small sample size in a single center and did not include external validation from other centers. Therefore, future studies are needed to validate the value of this model in a larger patient population across multiple centers. Second, the optimal thresholds for the inflammation-related ratio in this study may need to be re-evaluated due to the limited sample size. Considering these limitations, we plan to expand the sample size through a multicenter study in the future and compare our model with other prediction models to further verify its reliability.

RESULTS

Patient characteristics

According to the exclusion criteria, 207 out of 578 cases were excluded, and the remaining 371 patients were finally included for this study. Their demographic and clinical characteristics are shown in Table 1. Among them, 219 (59.02%) out of 371 patients suffered from early HCC recurrence. Those 371 patients were randomized at a ratio of 8:2 into the training and test sets. As a result, 296 patients were in the training set, 221 males and 75 females, of whom 172 (58.11%) experienced early HCC recurrence; 75 patients in the test set, 59 males and 16 females, of whom 47 had early HCC

recurrence (62.67%). No statistically significant difference was in any measure tested between the training and test groups of patients (P > 0.05 for all, Table 1).

Feature filter

A total of 26 potential characteristic variables were selected for this study. Their importance was analyzed using lasso regression, extreme gradient boosting (XGBOOST), and random forest (RF). The ten variables with the highest importance were ranked in descending order. Eight key variables were identified after screening the intersection of the variables using the three models. In no particular order, these variables were age, intratumoral arteries, AFP, blood glucose, number of tumors, GLR, liver cirrhosis, and platelets (Figure 3).

Model construction and performance validation

 We eventually constructed six machine learning models and plotted the training set and validation set AUC-ROC curve (FIG. 6A, 6B), as well as the forest plots of the AUC values of six predictive models in the training set (FIG. 5). The XGBoost analyses indicated larger and better AUC-ROCs in the training, validation, and test sets of 0.993 (95%CI: 0.982-1.000), 0.734 (0.601-0.867), and 0.706 (0.585-0.827), respectively in Figure 7. After the five-fold cross-validation, the AUC score had an SD of 0.002 for the XGBoost model, the smallest value among the six models, indicating that the XGBoost model had the most stable performance (Figure 5) and that XGBoost achieved lower and better Brier scores in Figure 6C. The DCA results revealed that the XGBoost model was the best diagnostic tool and had good clinical utility compared to other models in Figure 6D. In comparison with individual models (Tables 2 and 3), although the random forest model had better indicators for each evaluation in the training dataset, the XGBoost model was an optimal model after a five-fold cross-validation, considering that the random forest model might have overfitting. The XGBoost model achieved an AUC=0.706 (95%CI:0.585-0.827), accuracy=0.64, sensitivity=0.857, specificity=0.545, F1 score=0.766 in the test dataset. The optimal critical value for the prediction probability of the XGBoost model was 55.0 %, according to the Jorden index.

The XGBoost model was further analyzed using the SHAP software package, and the results unveiled that the relationship between the essential features and the impact of each feature in the sample had both positive and negative impacts, and the importance from the highest to lowest ranked as blood glucose, platelets, intratumoral arteries, AFP, liver cirrhosis, GLR, age, number of tumors.

Model demonstrations and applications

Analysis of the XGBoost model by the SHAP package indicated the impact of each feature in the sample and had both positive and negative impacts. The relationship between the magnitude of the eigenvalues and the predicted impact is shown in Figure 8. For applying the XGBoost model, the optimal cut-off point for the model's prediction probability was 55%. If the model predicted >55%, it would indicate that patients undergoing hepatectomy are at higher risk of early postoperative HCC recurrence. In addition, we generated a calculator based on the model online for clinicians to apply this model in their daily practice (Figure 9) (available at: http://152.136.184.184:5050/).

DISCUSSION

Research subjects

The demographic and clinical data of age, gender, pathology, radiological images, hepatectomy, and postoperative follow-up of 578 patients with primary liver cancer at Ningxia Medical University General Hospital from January 2017 to June 2021 were retrospectively collected according to the inclusion and exclusion criteria. After excluding those with non-primary liver cancer, incomplete data, receiving neoadjuvant therapy, or liver cancer rupture, the remaining cases were selected for this study, randomized by a simple number generated by a computer, and divided into the training and test sets at a ratio of 8:2. The training set of data was used to construct the machine learning model, and the test set of data was used to validate the model performance. The study was approved by the Ethics Committee of the General Hospital of Ningxia Medical University and individual patients provided written informed consent. The study was performed per the Declaration of Helsinki, and the study design

followed the TRIPOD statement [15] (Multifactorial predictive reporting norms for individual prognosis or diagnosis). The process of this study is shown (Figure 1).

Inclusion criteria included (1) Postoperative pathologically confirmed primary HCC; the first treatment option was radical hepatectomy. (2) All pre-operative enhanced abdominal MRI examinations. (3) Complete clinical, pathological, and imaging data and postoperative follow-up. Exclusion criteria were (1) Postoperative pathological diagnosis of secondary hepatocellular carcinoma. (2) Patients undergoing emergency surgery for the ruptured bleeding of liver cancer. (3) Receiving pre-operative anti-tumor therapy, such as interventional, ablative, immunotherapy, or targeted therapy. (4) Acute or chronic infection within two weeks prior to surgery.

The early liver cancer recurrence was defined as a previous report [16]. Briefly, a new tumor appeared in the liver within two years after surgical resection of the primary liver cancer. The new tumors were evaluated by radiological imaging or pathological examination. The time to recurrence (TTR) was defined from the surgical resection of primary liver cancer to the appearance of a new tumor in the liver.

Patients were followed up post-operatively in our outpatient service from monthly to every three months in the first-year post-surgical resection of the primary liver cancer. The patients were tested for serum liver functions and alpha-fetoprotein (AFP) levels, chest CT scanning, and enhanced abdomen CT or MRI. Subsequently, the patients who survived were followed up every six months starting from the second year after surgery. This follow-up continued until April 2023 or until tumor recurrence, loss of patient follow-up, or death.

Feature variable filtering

The characteristic variables included (1) patient's demographic characteristics and pre-operative laboratory tests, such as gender, age, AFP, PLT (blood platelet counts), blood glucose, HBeAg, HBsAg, MLR (monocyte to lymphocyte ratio), GLR (GGT to lymphocyte ratio), PLR (platelet count to lymphocyte count ratio), ALR (alkaline phosphatase to lymphocyte count ratio), LMR (lymphocyte to monocyte count), FAR (fibrinogen to albumin level ratio), SII (neutrophil* platelet/Lymphocyte), ALP

(alkaline phosphatase), and ALBI (albumin-bilirubin) score; (2) Pre-operative imaging features (abdominal enhancement MRI): Tumor size, tumor number, homogeneous tumor enhancement, intratumoral necrosis, tumor envelope, whether tumor margin is regular, intratumoral arteries, tumor adjacent to major vessels, ascitic fluid, and liver cirrhosis (Figure 2).

Data pre-processing

(1) The data in the training set were analyzed by receiver operating characteristic (ROC) curves to determine the best cut-off points for GLR, PLR, ALR, LMR, FAR, and SII continuous variable values to divide the patients into two groups (as shown in Figure 4), where the best cut-off value was 0.285 (95%CI:0.482-0.601) for MLR; 32.16 (95%CI:0.526- 0.642) for GLR; 138.18 (95%CI:0.494-0.61) for PLR; 80.29 (95%CI:0.465-0.583) for ALR; 5.33 (95%CI:0.40-0.517) for LMR; 0.07 (95%CI. 0.484-0.6) for FAR; and 312.71 (95%CI: 0.507-0.624) for SII in this population. (2) z-score normalization of continuous variables in the data. (3) When the proportion of missing values was less than 10%, the mean value was used for interpolation, and the case was deleted when the proportion of missing values of the variable was greater than 10%.

Statistical analysis

All statistical analyses were performed using the R 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria) and Python language. Categorical variables are expressed as n (%) and analyzed by the chi-square test or Fisher's exact test. Continuous variables are expressed as the mean ± standard deviation (SD) or the median and interquartile range (IQR) and analyzed by the independent-sample t-test or the Mann–Whitney U test based on their distribution. A p-value of <0.05 was considered statistically significant.

The most significant feature variables were chosen from the intersection of the three methods used for important feature analysis. Six different machine-learning models were constructed based on these feature sets. The hyperparameters for each model were determined through a grid search, selecting the best parameters and performing a five-fold cross-validation on the training data set. Cross-validation provided a more accurate

evaluation of model performance through metrics from multiple experiments. The predictive performance of the model was validated using discriminative and calibration methods, and the clinical usefulness of the model was evaluated by clinical decision curve analysis (DCA). After selection, the importance of each characteristic in the best model was interpreted using the SHAP package in Python, providing insight into the relationships among them.

CONCLUSION

Liver cancer is one of the most common malignant tumors globally, and its incidence is 5th in various malignant tumors. According to statistics, there were about 905,700 new cases of liver cancer worldwide in 2020, and about 830,200 patients with liver cancer died. The incidence and mortality of liver cancer are expected to increase by >50% in 2040 [1]. Hepatocellular carcinoma (HCC) is the primary type of liver cancer. Currently, surgical resection is still the main treatment for HCC. However, nearly 70% of patients experience HCC recurrence within five years after hepatectomy [2, 3], which is the main factor affecting patients' survival rate after surgery. HCC recurrence can be classified into early (≤2 years) or late recurrence (>2 years after HCC hepatectomy). A study has shown that the earlier the recurrence of HCC, the worse the prognosis, and the overall survival (OS) rate of patients with early recurrence of HCC tends to be lower than those with late recurrence [4]. For patients with recurrent HCC, radical treatment is more effective than palliative therapies in prolonging their survival [5, 6]. For patients at high risk of potentially postoperative recurrence, neoadjuvant therapy, appropriate intraoperative margin expansion, and early postoperative use of targeted therapies and immunotherapies may improve their prognosis [7]. However, the lack of specific tumor biomarkers and clinical features makes it challenging to identify the patients at high risk of postoperative recurrence before surgery. Therefore, developing pre-operative non-invasive predictive methods will be highly significant in identifying patients at high risk of postoperative recurrence and precise management of those patients by closely monitoring and individualized treatment on time.

Currently, there are many prediction models to predict the early recurrence of HCC in the clinic [8-10]. However, these models are usually constructed based on linear assumptions, which may not model the complex, multidimensional, and nonlinear relationships among different predictor variables that may exist. Their predictive value is greatly limited [11, 12]. In recent years, machine learning has great potential for research in the field of disease prediction because machine learning can better capture the relationships between nonlinearities, deeply mine the hidden relationships in massive data through various algorithms and build treatment models. These models can efficiently and accurately predict the prognosis of diseases and guide clinical treatment. Previous studies have used artificial intelligence (AI) techniques and machine learning algorithms to extract imaging features from pre-operative CT or MRI images and clinical data to predict the risk of liver cancer recurrence [13, 14]. Although these studies make superior predictions, these constructed models remain not highly interpretable and could not be widely applied in most medical scenarios, a common problem with machine learning research.

This study aimed to utilize machine learning techniques to select easily accessible imaging and pre-operative clinical characteristic variables. The goal was to develop six different risk prediction models for the early postoperative recurrence of HCC. We compared the sensitivity and specificity of these models in identifying patients at high risk of early postoperative HCC recurrence. Furthermore, to enhance the feasibility and applicability of the constructed model, we created an online calculator based on the predictive model. This tool is designed to assist clinicians in their daily medical practice.

ARTICLE HIGHLIGHTS

Research background

A multicenter study with large samples should be conducted in the future, and comparing our model with other prediction models is needed to further verify its reliability.

Research motivation

The XGBoost model exhibits superior performance and is a reliable tool for predicting early postoperative HCC recurrence. This model may guide surgical strategies and postoperative individualized medicine.

Research objectives

Following machine learning analysis, eight key feature variables were selected to construct six different prediction models. The XGBoost model outperformed other models, with the area under the ROC curve in the training, validation, and test datasets being 0.993 (95%CI: 0.982-1.000), 0.734 (0.601-0.867), and 0.706 (0.585-0.827), respectively. The XGBoost model also had an excellent calibration curve and DCA, indicating good predictive performance and clinical application value.

Research methods

The demographic and clinical data of 371 HCC patients were collected and analyzed, the key feature variables were selected to construct six different machine learning prediction models. Each model was evaluated, and the best-performing model was selected for interpreting the importance of each variable. Finally, an online calculator based on the model was generated for daily clinical practice.

Research results

This study aims to identify key variables in pre-operative clinical and imaging data using machine learning algorithms to construct multiple risk prediction models for early postoperative recurrence of hepatocellular carcinoma (HCC).

Research conclusions

To develop a new risk prediction model for the early postoperative recurrence of HCC and enhance the feasibility and applicability of the constructed model.

Research perspectives

Surgical resection is still the main treatment for HCC. HCC recurrence is the main factor affecting patients' survival rate after surgery. Developing pre-operative non-invasive predictive methods will be highly significant in identifying patients at high risk of postoperative recurrence and precise management of those patients by closely monitoring and individualized treatment on time.

87687_Auto_Edited.docx

8   Yuxian Kuai, Hui Huang, Xiaomei Dai, Zhongyue Zhang et al. "In PICU acute kidney injury stage 3 or mortality is associated with early excretion of urinary renin", Pediatric Research, 2021

Crossref

13 words — < 1%

EXCLUDE QUOTES   ON   EXCLUDE SOURCES   < 12 WORDS
EXCLUDE BIBLIOGRAPHY ON   EXCLUDE MATCHES   < 12 WORDS