REVIEW

1105 Role of ferroptosis in esophageal cancer and corresponding immunotherapy
Fan X, Fan YT, Zeng H, Dong XQ, Lu M, Zhang ZY

1119 Core fucosylation and its roles in gastrointestinal glycoimmunology
Zhang NZ, Zhao LF, Zhang Q, Fang H, Song WL, Li WZ, Ge YS, Gao P

1135 Interaction mechanisms between autophagy and ferroptosis: Potential role in colorectal cancer
Zeng XY, Qiu XZ, Wu JN, Liang SM, Huang JA, Liu SQ

1149 Application of G-quadruplex targets in gastrointestinal cancers: Advancements, challenges and prospects
Han ZQ, Wen LN

MINIREVIEWS

1174 Clinical value of serum pepsinogen in the diagnosis and treatment of gastric diseases
Qin Y, Geng JX, Huang B

ORIGINAL ARTICLE

Basic Study

1182 ENTPD1-AS1–miR-144-3p-mediated high expression of COL5A2 correlates with poor prognosis and macrophage infiltration in gastric cancer
Yuan HM, Pu XF, Wu H, Wu C

1200 Clinical significance and potential application of cuproptosis-related genes in gastric cancer
Yan JN, Guo LH, Zhu DP, Ye GL, Shao YF, Zhou HX

Clinical and Translational Research

1215 Integrated analysis of single-cell and bulk RNA-seq establishes a novel signature for prediction in gastric cancer
Wen F, Guan X, Qu HX, Jiang XJ

Case Control Study

1227 Proteomics-based identification of proteins in tumor-derived exosomes as candidate biomarkers for colorectal cancer
Zhou GYJ, Zhao DY, Yin TF, Wang QQ, Zhou YC, Yao SK

Retrospective Cohort Study

1241 Development and validation of a postoperative pulmonary infection prediction model for patients with primary hepatic carcinoma
Contents

Retrospective Study

1253  Clinical association between coagulation indicators and bone metastasis in patients with gastric cancer
  Wang X, Wang JY, Chen M, Ren J, Zhang X

1262  Efficacy of concurrent chemoradiotherapy with thalidomide and S-1 for esophageal carcinoma and its influence on serum tumor markers
  Zhang TW, Zhang P, Nie D, Che XY, Fu TT, Zhang Y

1271  Development and validation of an online calculator to predict the pathological nature of colorectal tumors

1283  Efficacy of continuous gastric artery infusion chemotherapy in relieving digestive obstruction in advanced gastric cancer
  Tang R, Chen GF, Jin K, Zhang GQ, Wu JJ, Han SG, Li B, Chao M

EVIDENCE-BASED MEDICINE

1295  Comprehensive bioinformatic analysis of mind bomb 1 gene in stomach adenocarcinoma

CASE REPORT

1311  Treatment of Candida albicans liver abscess complicated with COVID-19 after liver metastasis ablation: A case report
  Hu W, Lin X, Qian M, Du TM, Lan X
ABOUT COVER
Editorial Board Member of World Journal of Gastrointestinal Oncology, Zhi-Fei Cao, MD, PhD, Assistant Professor, Research Assistant Professor, Department of Pathology, The Second Affiliated Hospital of Soochow University, Suzhou 215004, Jiangsu Province, China. hunancao@163.com

AIMS AND SCOPE
The primary aim of World Journal of Gastrointestinal Oncology (WJGO, World J Gastrointest Oncol) is to provide scholars and readers from various fields of gastrointestinal oncology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJGO mainly publishes articles reporting research results and findings obtained in the field of gastrointestinal oncology and covering a wide range of topics including liver cell adenoma, gastric neoplasms, appendiceal neoplasms, biliary tract neoplasms, hepatocellular carcinoma, pancreatic carcinoma, cecal neoplasms, colonic neoplasms, colorectal neoplasms, duodenal neoplasms, esophageal neoplasms, gallbladder neoplasms, etc.

INDEXING/ABSTRACTING
The WJGO is now abstracted and indexed in PubMed, PubMed Central, Science Citation Index Expanded (SCIE, also known as SciSearch®), Journal Citation Reports/Science Edition, Scopus, Reference Citation Analysis, China National Knowledge Infrastructure, China Science and Technology Journal Database, and Superstar Journals Database. The 2023 edition of Journal Citation Reports® cites the 2022 impact factor (IF) for WJGO as 3.0; IF without journal self cites: 2.9; 5-year IF: 3.0; Journal Citation Indicator: 0.49; Ranking: 157 among 241 journals in oncology; Quartile category: Q3; Ranking: 58 among 93 journals in gastroenterology and hepatology; and Quartile category: Q3. The WJGO’s CiteScore for 2022 is 4.1 and Scopus CiteScore rank 2022: Gastroenterology is 71/149; Oncology is 197/366.

RESPONSIBLE EDITORS FOR THIS ISSUE
Production Editor: Xiang-Di Zhang; Production Department Director: Xiang Li; Editorial Office Director: Jia-Ru Fan.
Chao Lu, Zhi-Xiang Xing, Xi-Gang Xia, Zhi-Da Long, Bo Chen, Peng Zhou, Rui Wang, Department of Hepatobiliary & Pancreaticospleen Surgery, Yangtze University, Jing Zhou hospital, Jingzhou 434020, Hubei Province, China

Corresponding author: Rui Wang, MD, Surgical Oncologist, Department of Hepatobiliary & Pancreaticospleen Surgery, Yangtze University, Jing Zhou hospital, No. 26 Chuyuan Road, Jingzhou District, Jingzhou 434020, Hubei Province, China. wangrui_20222022@163.com

Abstract

BACKGROUND
There are factors that significantly increase the risk of postoperative pulmonary infections in patients with primary hepatic carcinoma (PHC). Previous reports have shown that over 10% of patients with PHC experience postoperative pulmonary infections. Thus, it is crucial to prioritize the prevention and treatment of postoperative pulmonary infections in patients with PHC.

AIM
To identify the risk factors for postoperative pulmonary infection in patients with PHC and develop a prediction model to aid in postoperative management.

METHODS
We retrospectively collected data from 505 patients who underwent hepatobiliary surgery between January 2015 and February 2023 in the Department of Hepatobiliary and Pancreaticospleen Surgery. Radiomics data were selected for statistical analysis, and clinical pathological parameters and imaging data were included in the screening database as candidate predictive variables. We then developed a pulmonary infection prediction model using three different models: An artificial neural network model; a random forest model; and a generalized linear regression model. Finally, we evaluated the accuracy and robustness of the prediction model using the receiver operating characteristic curve and decision curve analyses.

RESULTS
Among the 505 patients, 86 developed a postoperative pulmonary infection, resulting in an incidence rate of 17.03%. Based on the gray-level co-occurrence matrix, we identified 14 categories of radiomic data for variable screening of pulmonary infection prediction models. Among these, energy, contrast, the sum of squares (SOS), the inverse difference (IND), mean sum (MES), sum variance (SUV), sum entropy (SUE), and entropy were independent risk factors for...
pulmonary infection after hepatectomy and were listed as candidate variables of machine learning prediction models. The random forest model algorithm, in combination with IND, SOS, MES, SUE, SUV, and entropy, demonstrated the highest prediction efficiency in both the training and internal verification sets, with areas under the curve of 0.823 and 0.801 and a 95% confidence interval of 0.766-0.880 and 0.744-0.858, respectively. The other two types of prediction models had prediction efficiencies between areas under the curve of 0.734 and 0.815 and 95% confidence intervals of 0.677-0.791 and 0.766-0.864, respectively.

CONCLUSION
Postoperative pulmonary infection in patients undergoing hepatectomy may be related to risk factors such as IND, SOS, MES, SUE, SUV, energy, and entropy. The prediction model in this study based on diffusion-weighted images, especially the random forest model algorithm, can better predict and estimate the risk of pulmonary infection in patients undergoing hepatectomy, providing valuable guidance for postoperative management.

Key Words: Primary hepatic carcinoma; Pulmonary infection; Gray-level co-occurrence matrix; Machine learning; Prediction

©The Author(s) 2023. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Identifying risk factors for postoperative pulmonary infection in patients with primary hepatic carcinoma can improve the level of prevention and clinical treatment, ultimately reducing or even avoiding the occurrence of postoperative infection complications, reducing treatment time and costs, and improving patient efficacy and prognosis. The prediction model developed in our study provides valuable guidance for clinicians in predicting the risk of pulmonary infection and effectively preventing, diagnosing, and treating postoperative infection in patients with primary hepatic carcinoma, leading to an improved patient prognosis.


URL: https://www.wjgnet.com/1948-5204/full/v15/i7/1241.htm

DOI: https://dx.doi.org/10.4251/wjgo.v15.i7.1241

INTRODUCTION
Primary hepatic carcinoma (PHC) is a widespread malignant tumor with high incidence and mortality rates that poses a serious threat to human health worldwide. Surgical treatment remains the most effective treatment option for PHC[1]. However, postoperative infections, including surgical site and pulmonary infections, are among the main complications following surgery[2,3]. Patients with PHC are often accompanied by malnutrition, weakened immunity, and sputum accumulation in the respiratory tract due to prolonged bed rest and increased chronic consumption after surgery[4]. These factors significantly increase the risk of postoperative pulmonary infections in patients with PHC. Previous reports have shown that over 10% of patients with PHC experience postoperative pulmonary infections[5]. Thus, it is crucial to prioritize the prevention and treatment of postoperative pulmonary infections in patients with PHC.

Previous studies have shown that the causes of postoperative pulmonary infection in patients with PHC are related to the condition, respiratory microbiota, anesthesia-related factors, intraoperative procedures, and postoperative care of the patients. Of these factors, intraoperative procedures and the condition of the patients are considered the main triggers[2,6]. However, preoperative indicators that could be effective in preventing postoperative pulmonary infections and predictive models for such infections are still lacking. With the vigorous popularization of information technology in clinical practice, vast amounts of patient imaging data are now available, providing doctors with crucial objective data for clinical diagnosis, disease tracking, and surgical planning[7,8].

In this context, this study aimed to extract the texture features of radiomics of patients with PHC using a gray-level co-occurrence matrix (GLCM) to develop a predictive model to aid doctors in clinical decision-making and medical resource allocation for early interventions and treatments.

MATERIALS AND METHODS

Study population
We conducted a retrospective study on 505 patients with PHC who underwent surgical treatment at Jing Zhou Hospital between January 2015 and February 2023. The study included patients who met the following criteria: (1) Patients who met the diagnostic criteria for PHC (2011 version); (2) Patients with no distant liver cancer metastasis and feasible liver cancer radical surgery; (3) Patients who were confirmed by preoperative imaging to have no pulmonary infection; and (4) Patients whose postoperative adjuvant treatment and follow-up examinations were conducted in the same hospital.
Exclusion criteria included: (1) Patients with Child-Pugh grade C; (2) Patients with preoperative pulmonary disease; (3) Patients with other malignant tumors who have a history of chest trauma, surgery, and radiation chemotherapy; (4) Patients with a history of chronic respiratory diseases in the past; and (5) Patients who had not undergone liver and abdominal magnetic resonance imaging (MRI) examinations or whose imaging data collection was incomplete upon admission. As this was a retrospective chart review, informed written consent was not required in accordance with institutional review board policy, and the research scheme was implemented following the artificial intelligence model training specifications of the unit. All personal information of the patients was encrypted to prevent leakage and exempted from informed consent by the above ethics committee. The process of patient selection and development of the prediction model is shown in Figure 1.

**Diagnostic criteria for postoperative pulmonary infection**
In this study, postoperative pulmonary infection was diagnosed based on the following criteria: (1) Positive sputum culture after surgery; and (2) Identification of abnormal lung inflammatory rales upon lung auscultation. Both criteria need to be met in order for the diagnosis to be made.

**Acquisition of MRI-based radiomics parameters**
A GE Signa HD (3.0T) MR scanner was used in this study. The scanning sequence and parameter settings were as follows: Transverse fast spin echo T1-weighted imaging (repetition time, 400-620 ms, echo time, 10-16 ms); transverse, coronal, and sagittal plane fluid-sensitive proton density-weighted imaging (repetition time, 1800-3400 ms, echo time, 24-34 ms); layer thickness of 3-5 mm, layer spacing of 2 mm, matrix of 320 × (192-320) × 256, and field of view of 16 cm × (16-22) cm × 22 cm. We exported the MRI data of all patients in bitmap image file format from the picture archiving and communication system workstation for analysis. To ensure the reliability of the research results, we unified the window width or window level of T1-weighted imaging and fluid-sensitive proton density-weighted images at 1240/460 and 1050/340, respectively, when exporting the imaging data. Two experienced MRI diagnostic physicians selected a total of three layers of images to display the characteristics of the lesion, including the largest and consecutive layers, for measurement and analysis. We selected a sequence that displayed clear lesion boundaries to determine the lesion range as a reference standard and kept the region of interest (ROI) of other sequences consistent with it.

**Training and verification of the segmentation model**
With the guidance of two senior physicians, one physician manually delineated the ROI along the lesion boundary, covering the maximum layer of the lesion and one layer before and after the lesion. A total of three layers of images were used for calculating texture parameters. The average value was taken, and subsequent analysis was performed using MaZda software (available from: http://www.ele-tel.p.lodz.pl/mazda). To reduce bias caused by changes in brightness and contrast during result analysis, the grayscale of the image was standardized before extracting texture features. The MaZda software was used to calculate histogram and GLCM parameters within the ROI, including energy, contrast, correlation, the sum of squares (SOS), an inverse difference (IND), mean sum (MES), sum variance (SUV), sum entropy (SUE), entropy, and difference variance and entropy.

**Analysis and evaluation of the pulmonary infection prediction model**
All patients included in this study were randomly assigned to a model training set and an internal validation set at a 7:3 ratio. In the early stages of candidate variable selection, least absolute shrinkage and selection operator (LASSO) regression with minimum penalty coefficient and Pearson correlation coefficient were used to select variables that could ultimately be used to develop pulmonary infection prediction models. Three prediction models with different algorithms were developed, namely the artificial neural network model (ANNM), the random forest model (RFM), and the generalized linear regression model (GLRM)[9-12]. Additionally, LASSO was selected to develop a classification model for predicting pulmonary infection based on MRI images[13]. The effectiveness evaluation of each model mainly included decision curve analysis (DCA)[14], area under the receiver operating characteristic curve, and clinical influence curve[15].

**Statistical analysis**
The Wilcoxon rank-sum test or t-test was used to compare continuous variables, while the χ² test was used to compare categorical variables. Pearson’s correlation coefficients were used to assess the correlation between two continuous variables[16]. The charts in this study were analyzed visually using R Studio software (available from: https://www.r-project.org/). Statistical significance was defined as a two-tailed P value of < 0.05.

**RESULTS**

**Clinical characteristics of patients with PHC with or without pulmonary infection**
In the study, a total of 505 patients with PHC were included in the development of the pulmonary infection prediction models. Of these, 86 patients developed postoperative pulmonary infections and were categorized as the infected group, while 419 patients who did not have postoperative infections were included in the non-infected group. The incidence rates of postoperative pulmonary infection in the training and validation sets were 17.28% (61/353) and 16.45% (25/152), respectively. The specific surgical procedures performed included left hepatectomy, left hepatectomy and cholecystectomy, right hepatectomy, right hepatectomy and cholecystectomy, left hepatectomy and biliary drainage, right
hapatetomy, intraoperative radiofrequency therapy, and left hepatetomy. All patients with pulmonary infections improved after receiving treatment such as anti-inflammatory drugs and thoracic puncture drainage, and no other pulmonary complications occurred. The baseline data and radiomics-related extracted variables of the two groups of patients with PHC were summarized in Table 1 and Supplementary Table 1.

Features selected by LASSO regression analysis
As the included candidate variables were inevitably biased and had a non-normal distribution, we added penalty terms to the loss function (i.e., optimization goal) during the training and parameter-solving processes. This allowed the size of the coefficient to be considered. By setting a reduction coefficient (penalty coefficient), the coefficient of features with less impact was reduced to zero, retaining only important features, which is known as LASSO regression. Specifically, cross-validation was performed on all candidate parameters, and a dashed line was drawn at the optimal parameter (i.e., nine for non-zero parameters) to indicate the best-fit LASSO regression model. In the subsequent prediction model analysis, the optimal lambda value selected was substituted into the LASSO coefficient curve containing 18 variables, including 11 independent variables. This was used to develop an independent variable for predicting the risk propensity of postoperative pulmonary infection (Figure 2).

Development of a nomogram using GLRM analysis
As shown in Supplementary Table 2, we conducted a multivariate logistic regression analysis on all included candidate variables and found that seven variables were independent risk factors for postoperative pulmonary infection, including IND, SOS, MES, SUE, SUV, energy, and entropy. Based on the Akaike information criterion, we developed a predictive model for postoperative pulmonary infection and a nomogram (Figure 3). The nomogram showed the overall variables included in the prediction model on the left side, and each variable was assigned a scale value. The total score can be obtained by assigning scores to each variable for the included patients. Finally, the probability of pulmonary infection in patients can be evaluated based on the corresponding risk scale value of the total score. Moreover, the C-index value, which was verified internally by the bootstrap method, was 0.785, indicating good clinical applicability.

Development of the machine learning-based pulmonary infection prediction model
RFM and ANNM are the most widely used machine learning algorithms in various fields, including healthcare[9]. In this study, four supervised learning algorithms were used to develop the pulmonary infection model. As shown in Supplementary Table 3, top-ranking weight values in the RFM prediction model were obtained for energy, entropy, and variance, indicating their potential as candidate variables for RFM-based prediction of pulmonary infection (Figure 4A). Consequently, in ANNM, energy, entropy, SUV, SUE, IND, MES, contrast, and SOS also served as candidate variables for predicting pulmonary infection, and their assignments in the two different algorithm prediction models were inconsistent (Figure 4B, Supplementary Table 4).

Performance of pulmonary infection prediction models
The receiver operating characteristic curves were drawn to evaluate the predictive efficacy of each risk factor and the nomogram model. The receiver operating characteristic curve showed that the RFM model had higher predictive efficacy in both training and verification sets than the ANNM model, with areas under the curve of 0.823 and 0.801 and a 95% confidence interval (CI) of 0.766-0.880 and 0.744-0.858, respectively, compared to areas under the curve 0.815 and 0.787 and 95%CIs of 0.766-0.864 and 0.738-0.836, respectively, for the ANNM model. The predictive performance of preoperative GLCM-based radiomics for pulmonary infection was provided in Table 2. The prediction efficiency of the pulmonary infection prediction model for patients with PHC developed by a machine learning (ML)-based algorithm was
Table 1 Baseline demographic and radiomics of patients with primary hepatic carcinoma

<table>
<thead>
<tr>
<th>Variables</th>
<th>Training cohort</th>
<th></th>
<th>Testing cohort</th>
<th></th>
<th>P value</th>
<th>Training cohort</th>
<th></th>
<th>Testing cohort</th>
<th></th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall, n = 353</td>
<td>Infection, n = 61</td>
<td>Non-infection, n = 292</td>
<td></td>
<td></td>
<td>Overall, n = 152</td>
<td>Infection, n = 25</td>
<td>Non-infection, n = 127</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 60</td>
<td>240 (68.0)</td>
<td>37 (60.7)</td>
<td>203 (69.5)</td>
<td>0.231</td>
<td>109 (71.7)</td>
<td>16 (64.0)</td>
<td>93 (73.2)</td>
<td>0.488</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 60</td>
<td>113 (32.0)</td>
<td>24 (39.3)</td>
<td>89 (30.5)</td>
<td>43 (28.3)</td>
<td>9 (36.0)</td>
<td>34 (26.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>234 (66.3)</td>
<td>39 (63.9)</td>
<td>195 (66.8)</td>
<td>0.78</td>
<td>106 (69.7)</td>
<td>18 (72.0)</td>
<td>88 (69.3)</td>
<td>0.975</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>119 (33.7)</td>
<td>22 (36.1)</td>
<td>97 (33.2)</td>
<td>46 (30.3)</td>
<td>7 (28.0)</td>
<td>39 (30.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 18.5</td>
<td>47 (13.3)</td>
<td>10 (16.4)</td>
<td>37 (12.7)</td>
<td>0.877</td>
<td>20 (13.2)</td>
<td>4 (16.0)</td>
<td>16 (12.6)</td>
<td>0.627</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.5-23.9</td>
<td>121 (34.3)</td>
<td>21 (34.4)</td>
<td>100 (34.2)</td>
<td>54 (35.5)</td>
<td>6 (24.0)</td>
<td>48 (37.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.0-27.9</td>
<td>97 (27.5)</td>
<td>16 (26.2)</td>
<td>81 (27.7)</td>
<td>47 (30.9)</td>
<td>9 (36.0)</td>
<td>38 (29.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 28.0</td>
<td>88 (24.9)</td>
<td>14 (23.0)</td>
<td>74 (25.3)</td>
<td>31 (20.4)</td>
<td>6 (24.0)</td>
<td>25 (19.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>114 (32.3)</td>
<td>19 (31.1)</td>
<td>95 (32.5)</td>
<td>0.952</td>
<td>63 (41.4)</td>
<td>14 (56.0)</td>
<td>49 (38.6)</td>
<td>0.163</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>239 (67.7)</td>
<td>42 (68.9)</td>
<td>197 (67.5)</td>
<td>89 (58.6)</td>
<td>11 (44.0)</td>
<td>78 (61.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drinking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>100 (28.3)</td>
<td>18 (29.5)</td>
<td>82 (28.1)</td>
<td>0.945</td>
<td>56 (36.8)</td>
<td>8 (32.0)</td>
<td>48 (37.8)</td>
<td>0.747</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>253 (71.7)</td>
<td>43 (70.5)</td>
<td>210 (71.9)</td>
<td>96 (63.2)</td>
<td>17 (68.0)</td>
<td>79 (62.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of hepatitis B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>134 (38.0)</td>
<td>27 (44.3)</td>
<td>107 (36.6)</td>
<td>0.332</td>
<td>58 (38.2)</td>
<td>10 (40.0)</td>
<td>48 (37.8)</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>219 (62.0)</td>
<td>34 (55.7)</td>
<td>185 (63.4)</td>
<td>94 (61.8)</td>
<td>15 (60.0)</td>
<td>79 (62.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of hepatitis C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>129 (36.5)</td>
<td>22 (36.1)</td>
<td>107 (36.6)</td>
<td>1</td>
<td>50 (32.9)</td>
<td>9 (36.0)</td>
<td>41 (32.3)</td>
<td>0.898</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>224 (63.5)</td>
<td>39 (63.9)</td>
<td>185 (63.4)</td>
<td>102 (67.1)</td>
<td>16 (64.0)</td>
<td>86 (67.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>110 (31.2)</td>
<td>19 (31.1)</td>
<td>91 (31.2)</td>
<td>1</td>
<td>57 (37.5)</td>
<td>11 (44.0)</td>
<td>46 (36.2)</td>
<td>0.611</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>243 (68.8)</td>
<td>42 (68.9)</td>
<td>201 (68.8)</td>
<td>95 (62.5)</td>
<td>14 (56.0)</td>
<td>81 (63.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>119 (33.7)</td>
<td>21 (34.4)</td>
<td>98 (33.6)</td>
<td>1</td>
<td>52 (34.2)</td>
<td>9 (36.0)</td>
<td>43 (33.9)</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>234 (66.3)</td>
<td>40 (65.6)</td>
<td>194 (66.4)</td>
<td>100 (65.8)</td>
<td>16 (64.0)</td>
<td>84 (66.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cirrhosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>176 (49.9)</td>
<td>32 (52.5)</td>
<td>144 (49.3)</td>
<td>0.76</td>
<td>68 (44.7)</td>
<td>7 (28.0)</td>
<td>61 (48.0)</td>
<td>0.105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>177 (50.1)</td>
<td>29 (47.5)</td>
<td>148 (50.7)</td>
<td>84 (55.3)</td>
<td>18 (72.0)</td>
<td>66 (52.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child-Pugh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>253 (71.7)</td>
<td>39 (63.9)</td>
<td>214 (73.3)</td>
<td>0.187</td>
<td>102 (67.1)</td>
<td>14 (56.0)</td>
<td>88 (69.3)</td>
<td>0.289</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>100 (28.3)</td>
<td>22 (36.1)</td>
<td>78 (26.7)</td>
<td>50 (32.9)</td>
<td>11 (44.0)</td>
<td>39 (30.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFP in mg/L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 100</td>
<td>206 (58.4)</td>
<td>27 (44.3)</td>
<td>179 (61.3)</td>
<td>0.021</td>
<td>89 (58.6)</td>
<td>11 (44.0)</td>
<td>78 (61.4)</td>
<td>0.163</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 100</td>
<td>147 (41.6)</td>
<td>34 (55.7)</td>
<td>113 (38.7)</td>
<td>63 (41.4)</td>
<td>14 (56.0)</td>
<td>49 (38.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HBV DNA
Negative 233 (66.0) 192 (65.8) 0.944 106 (69.7) 36 (28.3)
Positive 120 (34.0) 46 (30.3) 0.003 83 (54.6) 19 (76.0)

Tumor size in cm ≤ 4 203 (57.5) 179 (61.3) 0.003 83 (54.6) 19 (76.0)
> 4 150 (42.5) 69 (45.4) 0.357 106 (69.7) 19 (76.0)

Treatment Laparoscopic 95 (26.9) 69 (45.4) 0.003 45 (29.6) 19 (76.0)
Open 258 (73.1) 113 (34.2) 0.002 106 (69.7) 36 (28.3)

LNM Yes 108 (30.6) 91 (31.2) 0.722 48 (31.6) 16 (58.6)
No 245 (69.4) 201 (68.8) 0.68 104 (68.4) 36 (28.3)

Energy 4.48 (2.97, 5.91) 4.02 (2.62, 5.26) < 0.001 4.52 (2.74, 6.02) 4.01 (2.45, 5.58)
Contrast 351.00 (316.00, 384.00) 360.00 (334.00, 391.75) < 0.001 347.00 (319.50, 381.00) 355.00 (311.00, 388.00)
Correlation 16.39 (12.49, 20.01) 16.70 (13.04, 20.11) 0.006 16.63 (12.02, 21.30) 16.07 (11.96, 20.02)
SOS 0.89 (0.71, 1.08) 0.83 (0.69, 0.97) < 0.001 0.86 (0.73, 1.06) 0.84 (0.69, 0.96)
IND 1.46 (1.17, 1.85) 1.38 (1.12, 1.67) < 0.001 1.52 (1.15, 1.90) 1.44 (1.12, 1.69)
MES 2.77 (1.95, 3.47) 2.50 (1.86, 3.12) < 0.001 2.55 (1.96, 3.47) 2.38 (1.89, 3.04)
SUV 21.80 (16.60, 27.00) 19.85 (16.10, 24.00) < 0.001 21.40 (15.97, 26.40) 20.00 (15.45, 24.35)
Entropy 0.84 (0.62, 1.07) 0.76 (0.58, 0.94) < 0.001 0.82 (0.60, 1.10) 0.73 (0.56, 0.94)
DIV 316.00 (210.00, 404.00) 295.00 (193.00, 391.75) 0.411 296.50 (196.50, 391.75) 296.00 (207.50, 400.50)
DIE 234.00 (188.00, 274.00) 229.00 (188.00, 274.00) 0.774 232.50 (189.75, 283.00) 231.00 (188.50, 271.50)

Data are presented as n (%) or median (interquartile range). AFP: Alpha-fetoprotein; BMI: Body mass index; DIE: Difference entropy; DIV: Difference variance; HBV: Hepatitis B virus; IND: Inverse difference; LNM: Lymph node metastasis; MES: Mean sum; SOS: Sum of squares; SUE: Sum entropy; SUV: Sum variance.

Table 2 Comparison of predictive efficacy of pulmonary infection prediction models via receiver operating characteristic curves

<table>
<thead>
<tr>
<th>Model</th>
<th>Training set</th>
<th>Internal validation set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC mean</td>
<td>AUC 95%CI</td>
</tr>
<tr>
<td>RFM</td>
<td>0.823</td>
<td>0.766-0.880</td>
</tr>
<tr>
<td>ANNMM</td>
<td>0.815</td>
<td>0.766-0.864</td>
</tr>
<tr>
<td>GLRM</td>
<td>0.766</td>
<td>0.717-0.815</td>
</tr>
</tbody>
</table>

1Variables included in the model.
ANNMM: Artificial neural network model; AUC: Area under the curve; CI: Confidence interval; GLRM: Generalized linear regression model; RFM: Random forest model.

better than that of the GLRM.

Furthermore, Figure 5 shows the DCA curve, with the abscissa indicating the threshold probability and the ordinate indicating the net benefit. The black horizontal line indicates a net benefit of zero, indicating that all patients were free of postoperative pulmonary infections. The gray diagonal line represents a scenario where all patients had a postoperative pulmonary infection and received treatment. DCA addresses the practical needs for clinical decision-making by
Predicting postoperative pulmonary infection

Figure 2 Predictor variable selection based on the least absolute shrinkage and selection operator regression method. A: Optimal parameter (lambda) selection in the least absolute shrinkage and selection operator (least absolute shrinkage and selection operator) model; B: Least absolute shrinkage and selection operator coefficient profiles of the candidate features.

incorporating patient or decision-maker preferences in the analysis. The DCA curve shows that the nomogram is more effective in predicting post-hepatectomy pulmonary infection than administering all or no postoperative treatment to patients.

Predictive value of an ML-based pulmonary infection prediction model

After developing three predictive models for pulmonary infection based on candidate predictive factors from preoperative MRI findings, we evaluated the optimal predictive performance of the RFM model. To further evaluate the differentiation efficiency of RFM, we used the clinical influence curve to assess the “classification accuracy” in the training set and internal verification set. As shown in Supplementary Figure 1, RFM effectively distinguished patients with pulmonary infection from those without pulmonary infection, which was consistent with the results of the postoperative pathological examination. Our study suggested that RFM is a reliable tool for the preoperative evaluation of pulmonary infection in patients with PHC and may become a powerful guiding tool for determining postoperative prevention of pulmonary infection. This also demonstrated that RFM was suitable for the preoperative assessment of the risk stratification of pulmonary infection.

DISCUSSION

The advancement in surgical technology has significantly reduced the preoperative mortality rate among patients with PHC. However, postoperative pulmonary infection has become a common and prominent complication, and its incidence rate has not shown significant improvement. Previous studies have reported postoperative pulmonary infection rates ranging from 9.60% to 22.38% [4,17]. The results of this study showed that the incidence rate of pulmonary infection after liver cancer surgery was 17.03%, consistent with previous studies [5,18,19]. The difference in the incidence rate of pulmonary infection can be attributed to the diagnostic criteria used by various diagnosis and treatment centers and the specific types of surgery for distinct populations. Furthermore, anti-infection and expectorant therapies are the main treatment measures for postoperative pulmonary infections in patients with PHC [5,20]. The reasonable selection of antibacterial and expectorant drugs, promotion of sputum excretion, prevention of respiratory difficulties caused by sputum stasis in pulmonary infections, and reduction of the induction of bronchial obstruction and atelectasis are all essential determinants of the incidence of postoperative pulmonary infections. Additionally, preoperative warning and evaluation for patients with primary liver cancer are crucial in reducing the occurrence of postoperative pulmonary infections [21,22]. In this context, our study employed advanced algorithms to explore the risk factors associated with postoperative infection, especially the predictive model developed using preoperative imaging parameters. This model can improve the level of prevention, clinical diagnosis, and treatment, reduce or avoid the occurrence of postoperative infection complications, and decrease treatment time and cost, thereby enhancing patient efficacy and prognosis.

This study used MRI, which has good soft tissue resolution and can perform multiplane and multisequence imaging. Its most significant advantage is its ability to achieve high spatial resolution and large-field scanning, making it the preferred imaging modality for preoperative MRI staging of cervical cancer. A non-inhibitory high-resolution MRI T2-weighted imaging sequence was used, and sagittal and short-axis scans (i.e., perpendicular to the long axis of the liver) were performed to fully display the cross-sectional area of the tumor. Moreover, the higher spatial resolution of MRI was utilized to leverage the natural contrast of pelvic fat and clearly display the depth and extent of invasion of PHC. However, it is important to note that the infiltration of the PHC in T2-weighted imaging may show uneven signal intensity, and most enhancement scans show uneven enhancement. In contrast, inflammatory lesions in the parenchymatous organ show uniform enhancement. Previous studies have highlighted the abundant venous plexus...
around the liver, which can create an atypical enhancement of high signal intensity on MRI, leading to texture analysis errors[23-25]. Additionally, tumors with a larger volume, worsening inflammation and edema, and an abundant venous plexus around the uterus can also become key factors affecting the preoperative evaluation of pulmonary infection.

Currently, image texture analysis is increasingly widely used in various research fields, including tumor diagnosis, efficacy evaluation, and prognosis[26,27]. Our study showed that the texture feature parameters in various sequences of patients with pulmonary and non-pulmonary infection had certain predictive values in texture analysis. It is speculated that the neovascularization, tumor necrosis, and invasive growth patterns in malignant tumors can contribute to complex internal structures that can be perceived and quantified through image texture analysis. Additionally, the heterogeneity of lesions measured by texture analysis feature parameters can precisely reflect the different texture compositions of different tumors, including their potential invasiveness[28]. Through image texture features, we quantitatively described the spatial distribution of pixels in MRI images. Our results indicated that the parameters based on GLCM showed significant statistical differences, making them potential predictive variables for pulmonary infection. For example, this study found a significant positive correlation between entropy differences and pulmonary infections. Previous studies have shown that entropy difference, a characteristic parameter used to measure structural disorder or complexity of images, is higher in tumors with higher malignancy levels[29,30]. Similarly, this study found that patients with pulmonary infections had a higher entropy value. We also included candidate variables with predictive value in the ML-based algorithm model. The results demonstrated that without distinguishing the predictive variables, the GLCM-based prediction efficiency reached its highest value of 0.823. This suggests that MRI examination before hepatectomy and texture analysis using sequence images have significant potential for predicting the risk stratification of pulmonary

Figure 3 Nomogram prediction model for predicting pulmonary infection in patients with primary hepatic carcinoma undergoing hepatectomy. A: Nomogram predicts risk of pulmonary infection; B: Calibration curves for the nomogram. SOS: Sum of squares; IND: Inverse difference; MES: Mean sum; SUV: Sum variance.
This study inevitably has the following limitations. First, the sample size was relatively small and from a single center, as we had strict requirements for MRI scanning parameters and equipment of the included patients. To overcome this, future studies should expand the sample size and conduct a prospective cohort study across multiple centers. Second, as a retrospective study, selection bias may have influenced the inclusion of the study population, leading to potential bias errors caused by the researcher’s personal experience or subjective judgment. Third, although we extracted GLCM-related parameters from MRI, higher-order textures were not included in the analysis. Thus, future studies should focus on expanding the extraction of higher-order texture parameters and exploring more predictive texture features for pulmonary infection prediction. Nevertheless, our pulmonary infection prediction model based on GLCM still has great development value for future clinical practice.

CONCLUSION

In summary, incorporating ML-based algorithms and GLCM radiomics features can facilitate timely and accurate risk stratification of pulmonary infection in patients with PHC before hepatectomy. Specifically, the RFM based on a random forest algorithm can aid clinicians in identifying high-risk patients with pulmonary infections and determining an appropriate surgical scope in a timely manner. This approach has promising applications for improving clinical outcomes and enhancing patient care.
Lu C et al. Predicting postoperative pulmonary infection

**ARTICLE HIGHLIGHTS**

**Research background**
Primary hepatic carcinoma (PHC) is a widespread malignant tumor with high incidence and mortality rates that poses a serious threat to human health worldwide. Surgical treatment remains the most effective treatment option for PHC. However, postoperative infections, including surgical site and pulmonary infections, are among the main complications following surgery.

**Research motivation**
To extract the texture features of radiomics of patients with PHC using a gray-level co-occurrence matrix to develop a predictive model to aid doctors in clinical decision-making and medical resource allocation for early interventions and treatments.

**Research objectives**
To identify the risk factors for postoperative pulmonary infection in patients with PHC and develop a prediction model to aid in postoperative management.

**Research methods**
Radiomics data were selected for statistical analysis, and clinical pathological parameters and imaging data were included in the screening database as candidate predictive variables. We then developed a pulmonary infection prediction model using three different models: An artificial neural network model; a random forest model (RFM); and a generalized linear regression model. Finally, we evaluated the accuracy and robustness of the prediction model using the receiver operating characteristic curve and decision curve analyses.

**Research results**
The RFM algorithm, in combination with sum of squares, inverse difference, mean sum, sum variance, sum entropy, and entropy, demonstrated the highest prediction efficiency in both the training and internal verification sets, with areas under the curve of 0.823 and 0.801 and 95% confidence intervals of 0.766-0.880 and 0.744-0.858, respectively. The artificial neural network model and generalized linear regression model had prediction efficiency areas under the curve of 0.734 and 0.815 and 95% confidence intervals of 0.677-0.791 and 0.766-0.864, respectively.

**Research conclusions**
Postoperative pulmonary infection in patients undergoing hepatectomy may be related to risk factors such as sum of squares, inverse difference, mean sum, sum variance, sum entropy, energy, and entropy. The RFM prediction model in this study based on diffusion-weighted images can better predict and estimate the risk of pulmonary infection in patients undergoing hepatectomy, providing valuable guidance for postoperative management.

**Research perspectives**
Identifying risk factors for postoperative pulmonary infection in patients with PHC can improve the level of prevention and clinical treatment, ultimately reducing or even avoiding the occurrence of postoperative infection complications, reducing treatment time and costs, and improving patient efficacy and prognosis. The prediction model developed in our study provides valuable guidance for clinicians in predicting the risk of pulmonary infection and effectively preventing, diagnosing, and treating postoperative infection in patients with PHC, leading to an improved patient prognosis.

**Figure 5** Prediction performance of pulmonary infection risk based on different supervised algorithms. A: Decision curve analysis for three prediction models in the training set; B: Decision curve analysis for three prediction models in the testing set. ANNM: Artificial neural network model; GLRM: Generalized linear regression model; RFM: Random forest model.
ACKNOWLEDGEMENTS

We thank all R language and R package developers for their assistance as well as all R package selfless contributors used for statistical analysis and visualization in this study.

FOOTNOTES

Author contributions: Lu C and Xing ZX contributed equally to this research; Wang R was the guarantor and designed the study; Lu C, Xing ZX, Xia XG, Long ZD, Chen B, and Zhou P participated in the acquisition, analysis, and interpretation of the data and drafted the initial manuscript; Lu C, Xing ZX, and Wang R revised the article critically for important intellectual content.

Institutional review board statement: The study was reviewed and approved by the Institutional Review Board of Jingzhou Hospital (Approval No. 2023-JH019).

Informed consent statement: All personal information of the patients was encrypted to prevent leakage and exempted from informed consent by the above ethics committee.

Conflict-of-interest statement: The authors declare that they have no conflicts of interest.

Data sharing statement: No additional data are available.

STROBE statement: The authors have read the STROBE Statement—checklist of items, and the manuscript was prepared and revised according to the STROBE Statement—checklist of items.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

Country/Territory of origin: China

ORCID number: Lu Chao 0000-0002-6236-1550; Rui Wang 0000-0002-6992-2287.

S-Editor: Chen YL

L-Editor: Filipodia

P-Editor: Zhang XD

REFERENCES


Pricing postoperative pulmonary infection


