

90033.docx

**Name of Journal:** *World Journal of Gastroenterology*

**Manuscript NO:** 90033

**Manuscript Type:** EDITORIAL

**leveraging machine learning for early recurrence prediction in hepatocellular carcinoma: a step towards precision medicine**

Ravikulan *et al.* Artificial Intelligence in predicting HCC recurrence

Abhimati Ravikulan, Kamran Rostami

## Abstract

The high rate of early recurrence in Hepatocellular carcinoma (HCC) post curative surgical intervention poses a substantial clinical hurdle, impacting patient outcomes and complicating postoperative management. The advent of machine learning provides a unique opportunity to harness vast datasets, identifying subtle patterns and factors that elude conventional prognostic methods.

Machine learning models, equipped with the ability to analyse intricate relationships within datasets, have shown promise in predicting outcomes in various medical disciplines. In the context of HCC, the application of machine learning to predict early recurrence holds potential for personalized postoperative care strategies.

This editorial comments on the study carried out by Zhang Y-B *et al* exploring the merits and efficacy of random survival forests (RSF) in identifying significant risk factors for recurrence, stratifying patients at low and high risk of HCC recurrence and comparing this to traditional COX proportional Hazard models (CPH).

In doing so, the study demonstrated that the RSF models are superior to traditional CPH models in predicting recurrence of HCC and represent a giant leap towards precision medicine.

**Key Words:** Machine learning; Artificial intelligence; Hepatocellular carcinoma; Hepatology; Early recurrence; Liver resection

Ravikulan A, Rostami K. leveraging machine learning for early recurrence prediction in hepatocellular carcinoma: a step towards precision medicine. *World J Gastroenterol* 2024; In press

**Core Tip:** This study addresses the crucial issue of early recurrence in HCC, emphasizing the significance of aggressive tumour characteristics. RSF, a machine learning model, surpasses conventional CPH models, offering improved prediction, clinical usefulness,

and overall performance. The model's ability to stratify risk facilitates targeted postoperative strategies, showcasing its potential as a guide for personalized patient care.

## **INTRODUCTION**

Developing a reliable pre-operative prediction model for postoperative recurrence of hepatocellular carcinoma (HCC) is essential in guiding individualized treatment and prognostication process of HCC.

In this issue of WJG, Zhang Y-B *et al* endeavour <sup>2</sup> 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 HCC<sup>[1]</sup>.

Hepatocellular carcinoma (HCC) remains a significant challenge in the realm of oncology, particularly due to its propensity for early recurrence following curative resection<sup>[2,3]</sup>. It is the sixth most common cancer worldwide<sup>[4]</sup>. Though surgical resection remains the mainstay of curative therapy for HCC, early recurrence of HCC (within 1 year) stands as a substantial barrier to positive patient outcomes<sup>[3,5]</sup>. Survival rates in early recurrence of HCC can be as low as 25% at 3-5 years post resection<sup>[5]</sup>. There are no current approved therapeutic regimens available for the recurrence of HCC<sup>[6]</sup>.

This raises a significant need to improve models for the early detection of patients at risk of recurrence. Many factors have been identified in predicting risk of recurrence of HCC<sup>[2]</sup> and this editorial explores the promising avenue in the quest for precision medicine<sup>[1]</sup> — the development of a machine learning model as highlighted by the authors aimed at predicting early recurrence after surgical intervention<sup>[11]</sup>. The advent of machine learning provides a unique opportunity to harness vast datasets, identifying subtle patterns and factors that elude conventional prognostic methods<sup>[7,8]</sup>.

Machine learning models, equipped with the ability to analyse intricate relationships within datasets, have shown promise in predicting outcomes in various medical disciplines<sup>[7]</sup>. In the context of HCC, the application of machine learning to predict early recurrence holds potential for personalized postoperative care strategies<sup>[9]</sup>.

Traditionally, predictive models, such as COX proportional hazard (CPH) models, have been employed, but their limitations have spurred the exploration of innovative methodologies<sup>[8-10]</sup>. This study undertakes a critical examination, comparing the efficacy of random survival forests (RSF) with CPH models in forecasting early recurrence for HCC patients following curative resection.

Drawing from a comprehensive cohort of 4,758 patients across two medical centres, this study utilized 15 key features to construct the RSF model. Features encompassed demographic, clinical, and tumour-specific factors. The RSF model was rigorously evaluated for discrimination, calibration, clinical utility, and overall performance, benchmarked against traditional models.

Out of 5,686 patients with HCC undergoing definitive surgical therapy at Eastern Hepatobiliary Surgery Hospital (January 2008 to December 2015), 4,376 met inclusion criteria. The study included patients with Child-Pugh A cirrhosis or B7 Liver function, without extrahepatic metastases, and complete resection of macroscopic tumour with histological evidence of tumour free margins. Exclusions ( $n = 1,310$ ) were due to preoperative anticancer treatment, history of other malignancies, palliative surgery, loss to follow up within 2 months of surgery, and perioperative death. The training cohort comprised 3,370 patients (January 2008 to December 2013), internal validation cohort 1,006 patients (January 2014 to December 2015), and external validation cohort 382 patients from Mengchao Hepatobiliary Hospital of Fujian Medical University.

The RSF model was constructed and used as a regression algorithm with faster training and lower estimation bias. This was achieved by using techniques of random forests such as feature and sample bagging. The model was constructed using fifteen factors including age, gender, aetiology, platelet count, albumin, total bilirubin, alpha-fetoprotein, tumour size, tumour number, microvascular invasion, macrovascular invasion, Edmondson-Steiner grade, tumour capsular, satellite nodules and liver cirrhosis. As 200 survival trees were built, the prediction error was significantly low and at 500 trees constructed, the variable importance (VIMP) for all 15 features was also generated. Utilizing cut-off values (50th and 85th centiles) from the training cohort's risk

index, RSF classified patients into low-risk, intermediate-risk, and high-risk groups, providing valuable insights for postoperative follow-up and adjuvant therapy. Kaplan-Meier analysis validated the stratification in all cohorts ( $P < 0.0001$ ). (Figure 1)

Model performance was assessed across several methods including model discrimination, model calibration, clinical usefulness and overall performance. In training, internal, and external validation cohorts, RSF outperformed existing models with C-index values of 0.725, 0.762, and 0.747, respectively. Overall performance time-dependent Brier (2 years) were 0.147, 0.129, and 0.156. RSF excelled against five other models that follow CPH. Decision curve analysis (DCA) affirmed RSF's superior net benefit over other models. (Table 1)

The RSF model, employing 500 survival trees, showcased superior predictive power. Key factors influencing recurrence were tumour size (which was the most significant risk factor for early recurrence), followed by macrovascular invasion, microvascular invasion, tumour number, and AFP levels.

The limitations to this study include selection bias as the cohort of patients largely had liver disease secondary to hepatitis B, which leaves a large space to question the applicability of these outcomes to other aetiologies of liver disease, and indeed, the RSF model did not consider aetiology and liver cirrhosis as important predictors of recurrence. Further studies will need to be conducted with the RSF model using patients of different aetiologies of liver disease to validate its use across different demographics in predicting HCC recurrence and reduce selection bias.

The user-friendly aspect of the web-tool developed, encompasses multiple complex aspects of the predictive model to increase its application in clinical practice.

In conclusion, the RSF model emerges as a beacon in the quest for precision postoperative care in HCC. Its demonstrated superiority over traditional models, coupled with its ability to stratify risk, ushers in a new era of individualized treatment strategies. The future role of artificial intelligence (AI) in evaluating hepatic diseases holds tremendous promise for revolutionizing diagnostic and treatment approaches. AI technologies,

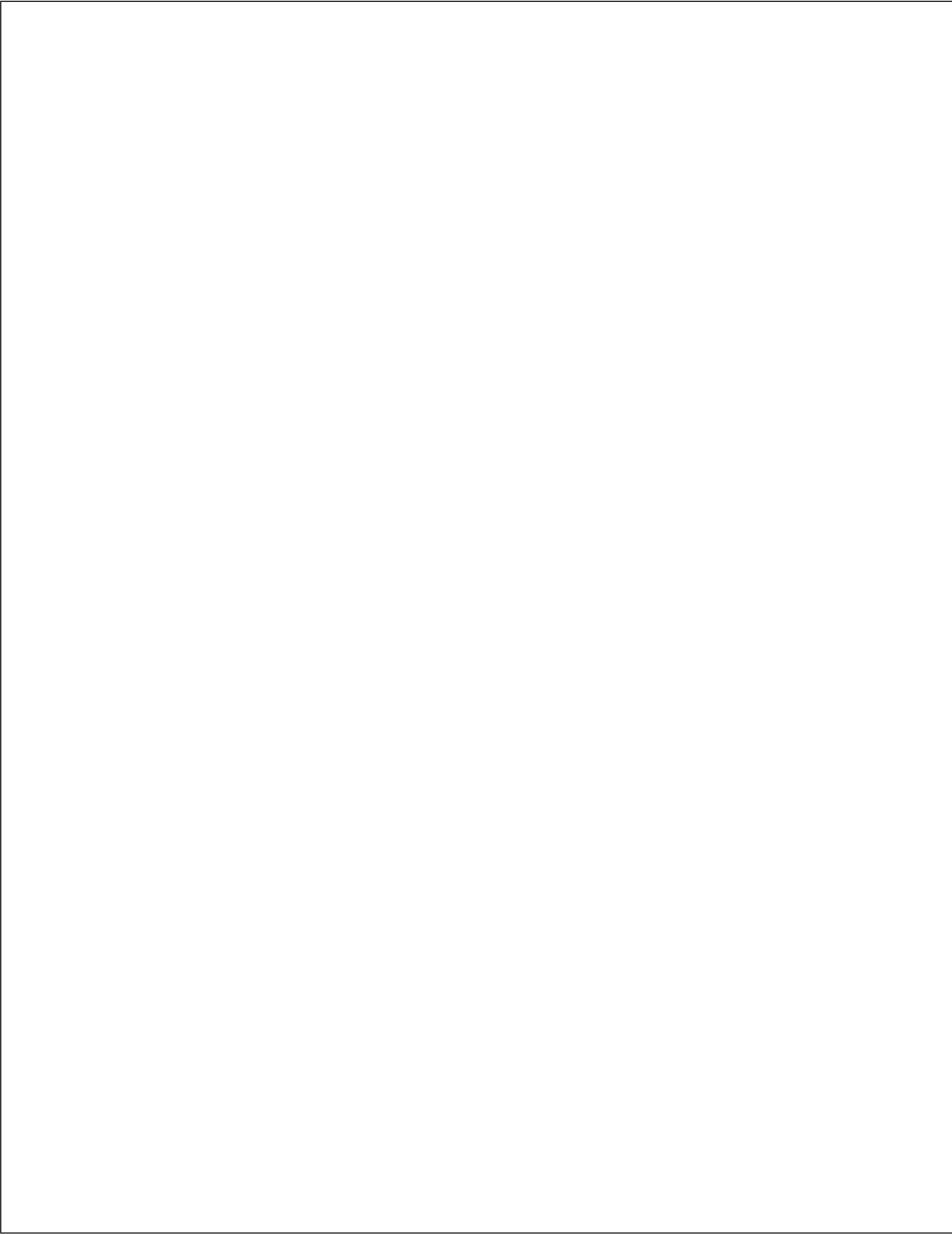
particularly machine learning algorithms, can analyse vast amounts of medical data, including imaging studies, laboratory results, and patient histories, to identify patterns and subtle anomalies that may escape the human eye. In hepatic diseases, AI can play a crucial role in early detection, risk assessment, and personalized treatment planning. Advanced imaging techniques, such as MRI and CT scans, can be enhanced by AI algorithms to provide more accurate and timely diagnoses of liver conditions. Machine learning models can also predict disease progression, helping healthcare professionals tailor interventions based on individual patient profiles.

As we navigate this machine learning odyssey, the RSF model stands poised to redefine the landscape of HCC prognostication and guide clinicians toward more informed and personalized patient care.

## **CONCLUSION**

In conclusion, the RSF model emerges as a beacon in the quest for precision postoperative care in HCC. Its demonstrated superiority over traditional models, coupled with its ability to stratify risk, ushers in a new era of individualized treatment strategies. The future role of artificial intelligence (AI) in evaluating hepatic diseases holds tremendous promise for revolutionizing diagnostic and treatment approaches. AI technologies, particularly machine learning algorithms, can analyse vast amounts of medical data, including imaging studies, laboratory results, and patient histories, to identify patterns and subtle anomalies that may escape the human eye. In hepatic diseases, AI can play a crucial role in early detection, risk assessment, and personalized treatment planning. Advanced imaging techniques, such as MRI and CT scans, can be enhanced by AI algorithms to provide more accurate and timely diagnoses of liver conditions. Machine learning models can also predict disease progression, helping healthcare professionals tailor interventions based on individual patient profiles.

As we navigate this machine learning odyssey, the RSF model stands poised to redefine the landscape of HCC prognostication and guide clinicians toward more informed and personalized patient care.



5%

SIMILARITY INDEX

PRIMARY SOURCES

1

Jianxing Zeng, Jinhua Zeng, Kongying Lin, Haitao Lin, Qionglan Wu, Pengfei Guo, Weiping Zhou, Jingfeng Liu. "Development of a machine learning model to predict early recurrence for hepatocellular carcinoma after curative resection", Hepatobiliary Surgery and Nutrition, 2021

Crossref

59 words — 4%

2

discovery.researcher.life

Internet

27 words — 2%