

89790_Auto_Edited.docx

Name of Journal: *World Journal of Gastroenterology*

Manuscript NO: 89790

Manuscript Type: EDITORIAL

From prediction to prevention: Machine learning revolutionizes hepatocellular carcinoma recurrence monitoring

Ramírez-Mejía MM *et al.* From Prediction to Prevention: Machine Learning for HCC

Mariana Michelle Ramírez-Mejía, Nahum Mendez-Sanchez

Abstract

In this editorial we comment on the article by Zhang YB *et al* entitled *Development of a machine learning-based model for predicting risk of early postoperative recurrence of hepatocellular carcinoma*. Hepatocellular carcinoma (HCC) remains a major global health challenge, with high incidence and mortality rates, primarily due to the critical issue of postoperative recurrence. Early recurrence within the initial 2 years post-treatment, linked to the hidden spread of the primary tumor, significantly impacts patient survival. Traditional predictive factors, encompassing patient and treatment-related aspects, have limitations in accurately prognosticating HCC recurrence. The integration of machine learning algorithms, fueled by the exponential growth of computational power, has revolutionized HCC research. This study, conducted by Zhang YB *et al.* demonstrates a groundbreaking preoperative prediction model for early postoperative HCC recurrence.

Challenges persist, including sample size constraints, data handling issues, and the need for extensive validation and interpretability. The study emphasizes collaborative efforts, urging multicenter studies and comparative analyses to validate and refine the model. Overcoming these challenges and exploring innovative approaches, such as multi-omics integration, will enhance personalized oncology care. This study marks a

significant stride toward precise, efficient, and personalized oncology practices, offering hope for improved patient outcomes in the realm of HCC treatment.

Key Words: Hepatocellular carcinoma; early recurrence; machine learning; XGBoost model; predictive precision medicine; clinical utility; personalized interventions.

Ramírez-Mejía MM, Mendez-Sanchez N. From Prediction to Prevention: Machine Learning Revolutionizes Hepatocellular Carcinoma Recurrence Monitoring. *World J Gastroenterol* 2023; In press

Core Tip: Machine learning provides a pivotal moment in personalized oncology care, paving the way for precise and individualized postoperative strategies, ultimately enhancing patient outcomes in the field of hepatocellular carcinoma treatment. Ongoing collaboration, expanded sample sizes, and multicenter studies are crucial for validating and refining this innovative predictive model, ensuring its applicability and reliability in diverse clinical settings.

1 **INTRODUCTION**

Hepatocellular carcinoma (HCC) is the most common form of primary liver cancer and is considered a major global health challenge due to its high incidence and mortality rates [1, 2]. Despite the advances in medical and surgical interventions, recurrence remains a critical problem affecting the long-term survival rates of patients with HCC [3, 4]. Recurrence of HCC within the initial 2 years post-treatment is categorized as early, typically arising from the hidden spread of the primary tumor within the liver and correlating with the tumor's size and extent. On the other hand, a recurrence occurring beyond the 2-year mark is termed late and is associated with de novo HCC, indicating the development of new cancerous growth independent of the original tumor [5]. Several predictive factors associated with recurrence have been recognized. Factors contributing to early recurrence encompass patient-related aspects such as age,

presence of underlying health conditions, liver function, viral load, presence and activity of hepatitis, metabolic dysfunction-associated fatty liver disease (MAFLD), alcoholic liver disease (ALD) and others etiologies, and the existence and activity of liver cirrhosis. Additionally, treatment-related factors, including the type of treatment employed, surgical margins, and the specifics of the resection, also play a crucial role in predicting early recurrence [6]. The intricate nature of liver cancer, coupled with the diverse factors influencing recurrence, makes accurate prognosis a significant challenge [7, 8]. In the constantly evolving landscape of HCC research, the quest for predicting early recurrence has undergone a remarkable transformation in recent decades. Initially, researchers focused on deciphering the morphological characteristics of tumors as a basis for predictions [9]. Factors such as vascular invasion, tumor multiplicity and large tumor size emerged as fundamental, although somewhat rudimentary, indicators that established the basis for understanding the complexities of HCC recurrence [10, 11]. Nevertheless, the paradigm shifted with the arrival of molecular analysis. Elevated alpha-fetoprotein (AFP) levels emerged as one of the first markers used for HCC prediction, offering insight into the intricate molecular landscape of this aggressive cancer [12, 13]. Despite these advances, the multifaceted nature of HCC recurrence demanded more nuanced and sophisticated approaches. Researchers and clinicians alike recognized the limitations of relying solely on morphologic and molecular analyses [14]. The quest for predictive accuracy urged the scientific community to explore uncharted territory, harnessing the transformative power of technology, especially in the fields of imaging, genetics and computational sciences [15]. The evolution of technology emerged as a pivotal catalyst, propelling HCC research into an era of unprecedented possibilities. Cutting-edge imaging techniques, coupled with advancements in genetic profiling, provided researchers with a comprehensive understanding of the tumor microenvironment [16-19]. These insights, combined with the computational prowess of modern data analysis, paved the way for a new generation of predictive models. These models transcended the limitations of traditional analyses, offering a more nuanced and accurate glimpse into the future course of HCC [20-23].

MACHINE LEARNING IN HCC RESEARCH

The exponential growth of computational power has heralded a new era in HCC research, where machine learning algorithms have emerged as valuable tools in handling vast datasets and deciphering complex patterns [24-26]. This convergence of computational capabilities and healthcare needs represents a significant paradigm shift, transforming the landscape of HCC research. The integration of machine learning into the study of HCC offers a multitude of benefits and promises to address long-standing challenges in this field[27, 28].

In a retrospective study, Zhang YB *et al.*, harnessed the potential of supervised machine learning to develop a state-of-the-art preoperative prediction model for early postoperative HCC recurrence. Leveraging readily available clinical and imaging data, the team built six different risk prediction models, using ensemble learning, linear and neural network models, each meticulously designed to identify patients at high risk of recurrence. The study methodology consisted of analyzing the demographic and clinical data of 371 patients with HCC, excluding cases with incomplete data or previous neoadjuvant treatments. Using machine learning algorithms, the researchers identified eight key variables to predict early HCC recurrence: age, intratumoral arteries, AFP, blood glucose, number of tumors, glucose-to-lymphocyte ratio (GLR), liver cirrhosis, and platelet count (PLT). These variables formed the basis for the construction of six different prediction models, of which the XGBoost model proved to be the most robust (Figure 1). The XGBoost model, outperformed its peers, showing unmatched performance on several metrics. In the training dataset, the model achieved an impressive area under the receiver operating characteristic curve (AUROC) of 0.993, proof of its accuracy. Even in the validation and test data sets, the XGBoost model maintained its excellence, with AUROC values of 0.734 and 0.706, respectively. Calibration curve analysis underscored the reliability of the model, confirming its alignment with real-world results. Furthermore, decision curve analysis highlighted the clinical utility of the XGBoost model, highlighting its potential to guide surgical

strategies and usher in an era of individualized postoperative medicine. By employing the SHAP (SHapley Additive exPlanations) package, the study provided a detailed interpretation of the model results, unraveling the intricate relationships between variables. Preoperative glycemia emerged as a key factor, in line with previous research highlighting its role in HCC progression. The predictive power of the model was further demonstrated using an online calculator, designed to assist physicians in their daily practice. This user-friendly tool represents a major breakthrough, as it ensures the seamless integration of predictive analytics into clinical decision making.

The relevance of this study goes far beyond conventional medical research. Machine learning algorithms, used to decipher the intricate web of preoperative variables, have ushered in an era where predictive precision medicine reigns supreme. The identification of these eight key variables is a pivotal moment, providing physicians with unprecedented insight into the intricate factors that determine early postoperative recurrence.

Several research projects have been conducted to explore the use of machine learning in predicting HCC recurrence. Kucukkaya AS *et al.* developed a predictive model based on the analysis of pre-treatment magnetic resonance imaging (MRI) using the VGG16 and XGBoost machine learning models. This model aimed to predict recurrence in six different time intervals, ranging from 1 to 6 years, and demonstrated performance with AUROC values between 0.71 and 0.85^[26]. In another study, Zeng J *et al.* compared the performance of random survival forest (RSF) models with Cox proportional hazard (CPH) models in predicting early recurrence using clinical features of the participants. In training and internal and external validation cohorts, the C-index of the RSF model was 0.725, 0.762, and 0.747, respectively ^[24]. Although both studies highlighted the utility of machine learning, neither included the analysis of clinical and imaging variables, assuming a linear interaction of predictors for HCC recurrence. In this context, Zhang YB *et al.* propose a solution in their study, addressing the need to

include clinical and imaging variables in the analysis. Their approach seeks to overcome the limitation of assuming linear interactions among predictors of HCC recurrence.

CHALLENGES AND FUTURE DIRECTIONS

This represents a significant advancement in early postoperative HCC recurrence prediction. Future directions should focus on overcoming challenges related to sample size, data handling, validation, and interpretability. The authors acknowledge the limitations, emphasizing the need for future research endeavors to validate and refine the model further. A call to action resounds within these findings, urging the scientific community to collaborate, expand sample sizes, and conduct multicenter studies. Comparative analyses with existing prediction models stand as a crucial step, ensuring the reliability and applicability of this innovative approach [29, 30].

CONCLUSION

The development of accurate, interpretable, and widely applicable prediction models for early postoperative HCC recurrence represents a significant advance in personalized medicine. Addressing the challenges associated with data quality, model interpretability, while exploring innovative approaches, such as multi-omics integration and continuous model refinement, will pave the way for improved patient outcomes and healthcare practices in the field of HCC treatment. Through collaborative efforts, continued research and a patient-centered approach, the field of oncology can continue its journey toward more precise, efficient and personalized oncology care.

3%

SIMILARITY INDEX

PRIMARY SOURCES

- 1

pharmaceutical-journal.com
Internet

18 words — 1%
- 2

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

14 words — 1%
- 3

Yu-Bo Zhang, Gang Yang, Yang Bu, Peng Lei, Wei Zhang, Dan-Yang Zhang. "Development of a machine learning-based model for predicting risk of early postoperative recurrence of hepatocellular carcinoma", World Journal of Gastroenterology, 2023
Crossref

14 words — 1%
- 4

www.mdpi.com
Internet

13 words — 1%

EXCLUDE QUOTES ON

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES

EXCLUDE MATCHES

< 12 WORDS

< 12 WORDS