ANNALS OF MEDICINE
2025, VOL. 57, NO. 1, 2568118
https://doi.org/10.1080/07853890.2025.2568118

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Artificial intelligence for the prediction of posthepatectomy recurrence in hepatocellular carcinoma: a systematic review and meta-analysis

Zhiqiang Xiang[a†], Jing Deng[b†], Hao Liang[c†], Mengliang Jiang[c†], Yuhan Liang[d], Zhaohai Liu[c], Yachen Wu[c], Leyuan Peng[a], Xiaoming Dai[c] and Zhu Zhu[c]

[a]Department of Hepatobiliary Surgery, Hunan University of Medicine General Hospital, Huaihua, China; [b]Department of Thoracic Surgery, The Second Xiangya Hospital of Central South University, Changsha, China; [c]The First Affiliated Hospital, Department of Hepatobiliary Surgery, Hengyang Medical School, University of South China, Hengyang, China; [d]College of Medical Imaging Laboratory and Rehabilitation, Xiangnan University, Chenzhou, China

ABSTRACT

Objective: Posthepatectomy recurrence of hepatocellular carcinoma (HCC) is a major cause of poor prognosis. Accurate prediction is essential for reducing the burden of advanced disease and improving outcomes.

Methods: A systematic search of the PubMed, Embase, and Cochrane Library databases was conducted from their inception to December 31, 2024. The standard quality assessment of diagnostic accuracy studies (QUADAS-2) tool was utilized to analyse the methodological quality of the included studies. Bivariate linear mixed models were used to pool diagnostic estimates, including sensitivity (Se), specificity (Sp), positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR). Additionally, the area under the receiver operating characteristic curves (AUC) of the included studies was utilized to evaluate the diagnostic value.

Results: A total of 6665 HCC patients included in 20 studies were enrolled. The pooled Se, Sp, PLR, NLR, DOR and AUC for the overall AI-assisted diagnostic performance for postoperative HCC recurrence were 0.87 (95% CI: 0.72–0.83), 0.85 (95% CI: 0.80–0.90), 5.39 (95% CI: 3.85–7.55), 0.25 (95% CI: 0.20–0.33), 21 (95% CI: 13–35), and 0.89 (95% CI: 0.86–0.91), respectively.

Conclusion: AI showed high accuracy in predicting the posthepatectomy recurrence of HCC and would shed light on screening and monitoring high-risk patients following liver resection for further treatment.

ARTICLE HISTORY
Received 17 May 2025
Revised 18 August 2025
Accepted 24 September 2025

KEYWORDS
Hepatocellular carcinoma; recurrence; prediction; artificial intelligence; meta-analysis

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common malignancy worldwide and the third leading cause of death [1]. Although hepatectomy is one of the most effective radical treatments for HCC, the postoperative prognosis is unsatisfactory, with a 5-year recurrence rate as high as 70% and intrahepatic recurrence accounting for 60%–80% of these cases [2–5]. The accurate prediction of posthepatectomy recurrence of HCC remains a challenging clinical issue. Several biomarkers are currently employed to predict HCC recurrence after surgery, including Alpha-fetoprotein (AFP) and protein induced by vitamin K absence or antagonist II (PIVKA-II), with a discouraging prediction accuracy of 60%–80% [6,7]. Therefore, more precise prediction methods are needed to address the issue.

In recent years, the rapid development of artificial intelligence (AI) technology has facilitated significant progress in auxiliary diagnosis, decision-making and disease prediction. In terms of data processing capacity, feature extraction, and learning mechanisms, conventional methodologies predominantly rely

CONTACT Zhu Zhu zhuzhu027@gmail.com The First Affiliated Hospital, University of South China, Hengyang, Hunan, 421001, China; Leyuan Peng ply1013@163.com Department of Hepatobiliary Surgery, Hunan University of Medicine General Hospital, Huaihua, Hunan 418000, China.
†These authors contributed equally to this work, and they are co-first authors.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/07853890.2025.2568118.
© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

on manual data collection, organization, and analysis, which inherently constrains dataset scalability. Notably, AI-driven approaches demonstrate superior capability in handling massive datasets, enabling the discovery of more comprehensive and nuanced feature patterns that enhance predictive accuracy [8]. Recently, several studies have applied AI models to predict the risk of posthepatectomy recurrence of HCC [9–28]. However, their overall accuracy of prediction cannot be effectively evaluated due to the various evaluation methods and model algorithms used in the studies.

In this study, we conducted a meta-analysis to assess the accuracy of AI technology in predicting the posthepatectomy recurrence of HCC and systematically evaluated the diagnostic performance of AI technology in predicting the risk of HCC recurrence after surgery.

Materials and methods

The work has been reported in line with the guidelines of PRISMA [25] (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and AMSTAR [26] (Assessing the methodological quality of systematic reviews).

Protocol

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [29] were followed in this systematic review and meta-analysis.

Study selection and literature search

We conducted a systematic search of the PubMed, Embase, and Cochrane Library databases from their inception to December 31, 2024, using subject and free terms including 'hepatocellular carcinoma', 'artificial intelligence', 'deep learning', 'machine learning', 'recurrence', and their variations (see search strategy in Supplementary Table S1). Moreover, we carefully examined the references of each identified primary study and past systematic review to identify additional relevant articles.

All publications were independently screened by two authors. After eliminating duplicate papers, all remaining publications' titles and abstracts were reviewed. The full text was then evaluated by two authors to assess whether the article fulfilled the inclusion criteria or whether the outcome variables met the requirements. If issues could not be addressed *via* conversation, a consensus was reached through the judgment of a third examiner.

The inclusion criteria were as follows: patients pathologically diagnosed with primary HCC following hepatectomy with a risk of recurrence predicted using AI methods. The exclusion criteria were as follows: incomplete or invalid outcome data; inclusion of other regimens in the study cohort; and nondiagnostic tests, case reports, reviews, expert opinions, or conference abstracts.

Data extraction and quality assessment

We extracted data from each eligible study, including the first author, publication date, country, research design, AI learning model, true-positive (TP), false-positive (FP), false-negative (FN), and true-negative (TN) rates, verification method, area under the curve (AUC), training dataset, and main content of the model. If a model's dataset was validated multiple times, the model with the highest sensitivity (Se) and specificity (Sp) was selected. This work was finalized independently by two reviewers, with the assistance of a third party wherever deemed necessary.

To analyse the methodological quality of the included research, we utilized the standard quality assessment of diagnostic accuracy studies (QUADAS-2) tool [30]. STATA version 15 (Stata Corp LP, College Station, USA) was used to analyse the risk of bias and applicability using the scoring rules for each item on the checklist. The four areas of evaluation included patient selection, indicator testing, reference standard, processes, and timing. To highlight the possibility of bias, each question was answered as 'Yes', 'No', or 'Unclear', and 'High Risk', 'Low Risk' , or 'Unclear Risk' were considered to indicate an applicability concern.

Statistical analysis

To analyse the raw data, we utilized the MIDAS module of STATA 15 (StataCorp LP, College Station, USA, Texas). The results were used to plot a summary receiver operating characteristic (SROC) curve, with the AUC demonstrating the diagnostic ability of the included studies [31]. For stratified analyses comparing models using only preoperative variables (Pre) with those incorporating both preoperative and postoperative variables (Pre + Post), the reported AUCs and their 95% confidence intervals (CIs) were extracted from each study. When not directly reported, standard errors (SEs) were calculated from the CIs. To stabilize variance and account for the bounded distribution of AUC values, all AUCs were transformed to the logit scale prior to pooling. Pooled estimates were obtained using a random-effects model with restricted maximum likelihood (REML) estimation and subsequently back-transformed to the original AUC scale to facilitate interpretation. AUC values of 0.5–0.7, 0.7–0.9, and >0.9 were considered to indicate low, medium, and high performance, respectively.

To demonstrate variations between studies, we created a forest plot and detected heterogeneity in the Se and Sp of the pooled data. The MIDAS module in STATA was applied to assess heterogeneity due to the threshold effects, whereas I^2 and Q tests were employed to quantify heterogeneity generated by the non-threshold effects. I^2 ≥50% was considered to indicate moderate to high heterogeneity among studies, and p <.05 was considered to indicate a significant difference [30]. Moreover, we utilized Deeks's asymmetry regression test [32] to examine potential publication bias. A slope coefficient with p <.10 was considered to indicate a significant publication bias.

Results

Literature selection

Figure 1 depicts the literature search and selection process [29]. A total of 20 studies were included and published from 2019 to 2024, with 5 studies based on clinical characteristics, 13 studies based on

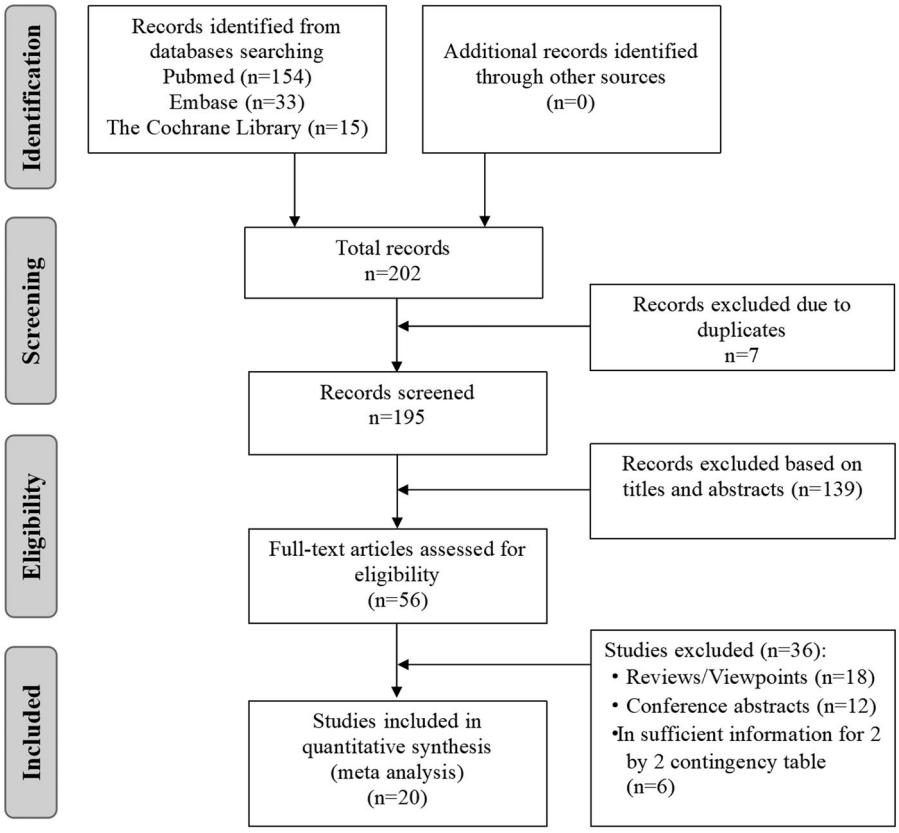


Figure 1. PRISMA flow diagram of the study process.

radiomics, and 2 studies based on pathology. A total of 6665 HCC patients were enrolled in the study, including 2976 patients (44.65%) with recurrence and 3689 patients (55.35%) without recurrence.

Data extraction and quality assessment

We extracted data and evaluated the quality of the 20 included studies; the relevant characteristics are illustrated in Table 1, with detailed modeling architectures, input features, and methodological characteristics provided in Supplementary Table S2. The distribution of methodological quality assessments for the included studies based on the QUADAS-2 scale is shown in Figure 2. According to the quality assessment, no studies were excluded from the analysis. The Deeks funnel plot asymmetry test slope coefficient ($p = 0.20$) revealed no publication bias (Figure 3).

Data analysis

The forest plots and comprehensive results of all studies included in this diagnostic meta-analysis are shown in Figure 4. The pooled Se and Sp of the overall AI-assisted diagnostic performance for postoperative HCC recurrence was 78% (95% CI: 72–83%) and 85% (95% CI: 80–90%), respectively, while the PLR, NLR, and combined DOR was 5.39 (95% CI: 3.85–7.55), 0.25 (95% CI: 0.20–0.33), and 21 (95% CI: 13–35), respectively. On the SROC curves for all datasets ($n = 20$), circles indicate the observed data, while solid squares indicate the total Se and total Sp aggregation points. The upper and lower boundaries indicate Se confidence, whereas 1-Sp confidence indicate the left and right boundaries (Figure 5). The AUC was 0.89, with a 95% CI of 0.86–0.91, suggesting high diagnostic value (Figure 5). Among the included studies, two employed time-to-event survival models, reporting time-dependent AUCs of 0.83 at 2 years and 0.683 at 5 years. The positive prior probability (Prior Prob) for HCC recurrence was 40% in this research, and a PLR of 5 boosted the postprobability positive (Post Prob Pos) to 78%. Similarly, when the NLR was 0.25, the postprobability negative (Post Prob Neg) decreased to 14% (Figure 6). Nevertheless, I^2 values exhibited significant heterogeneity among the studies.

In stratified analyses, models using only preoperative variables (Pre, $n = 11$) yielded a pooled AUC of 0.84 (95% CI: 0.81–0.88), which was comparable to that of models incorporating both preoperative and postoperative variables (Pre + Post, $n = 7$; AUC = 0.88, 95% CI: 0.85–0.91), with no statistically significant difference between the two approaches ($p = 0.37$).

Subgroup analysis

There was significant heterogeneity among studies (I^2 = 99%, 95% CI: 98%–99%, $p < .0001$). Similarly, Se and Sp displayed high heterogeneity in the forest plots (I^2 = 93.09%, 95% CI: 91.01%–95.17%, $p < .001$;

Table 1. Baseline characteristics of included studies.

Study	Country	Recurrence Rate	TP	FN	FP	TN	AUC
Lee (2021)	China	46.2%	8	3	2	25	0.741
Liu (2021)	China	Not mentioned	44	2	10	25	0.917
Sun (2021)	China	85.7%	541	46	130	462	0.934
Mai (2021)	China	35.9%	231	109	93	246	0.753
Yamshita (2021)	United States	34.7%	63	10	6	119	0.956
Cui (2022)	China	43.3%	61	26	12	79	0.864
Gao (2022)	China	28.2%	99	76	8	195	0.911
Huang (2022)	China	52.4%	84	41	34	130	0.83
Ren (2022)	China	17.8%	16	8	1	29	0.864
Shi (2022)	China	50.8%	53	7	10	39	0.93
Wu (2022)	China	Not mentioned	17	3	3	18	0.948
Wang (2022)	China	49.52%	43	10	10	42	0.834
Zeng (2022)	China	43.4%	281	175	97	453	0.785
Zhang (2022)	China	39.5%	48	16	0	57	0.942
Li (2023)	China	33.5%	8	3	2	25	0.857
Lv (2023)	China	51.3%	107	8	7	102	0.98
Kinoshita (2023)	Japan	41.0%	10	5	13	26	0.73
Altaf (2024)	United States	49.7%	219	72	364	755	0.858
Yan (2024)	China	Not mentioned	24	13	12	57	0.733
Zandavi (2024)	Australia	49.7%	118	25	18	99	0.835

Notes: TP: true-positive; FP: false-positive; FN: false-negative; TN: true-negative; AUC: area under the curve.

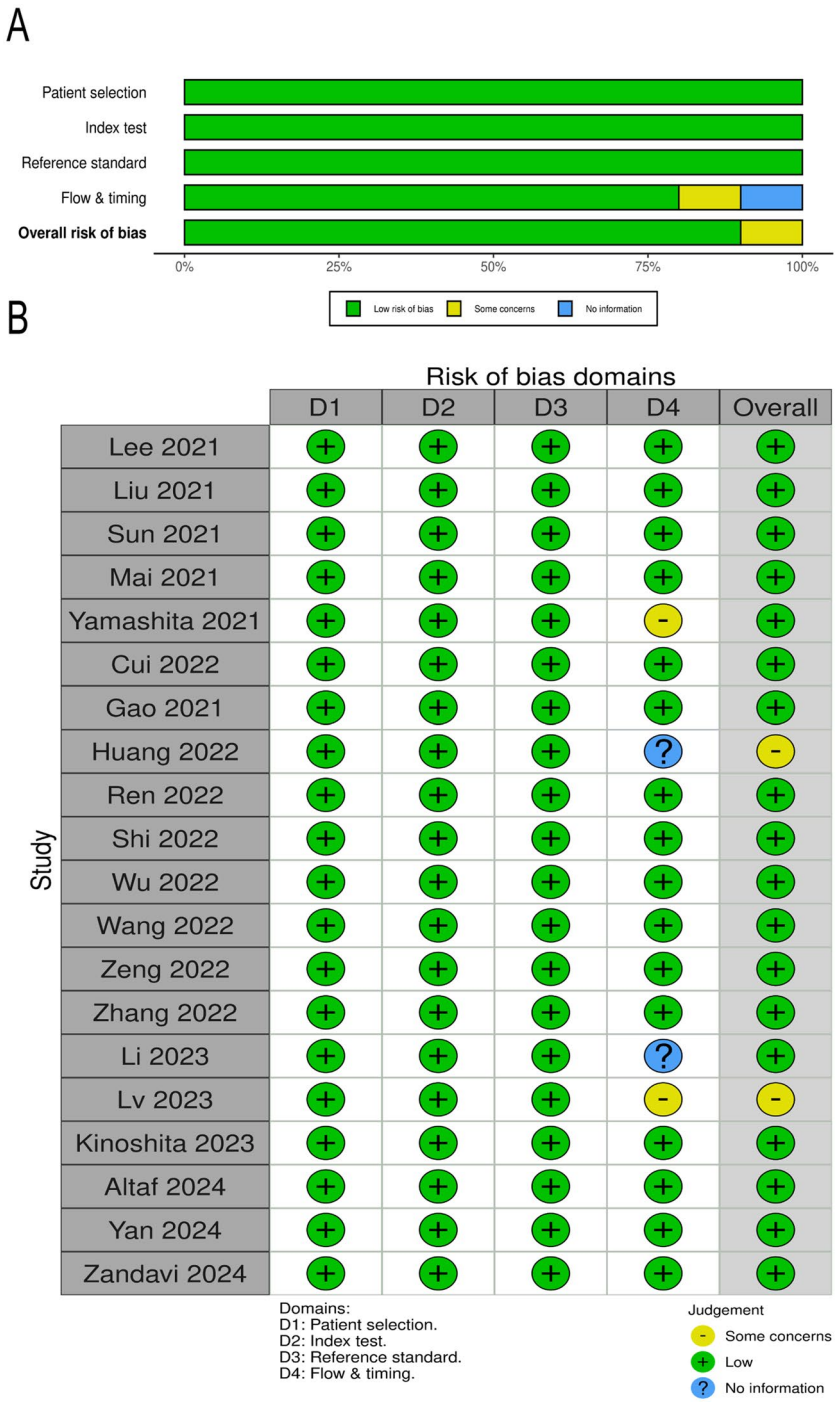


Figure 2. Stacked bar charts of the QUADAS-2 scores for methodological quality. (A) Summary across all studies; (B) individual studies. For each domain, the proportions indicating low, high, or unclear risks of bias, and applicability concerns, are depicted in green, red, and yellow, respectively.

$I^2 = 93.78\%$, 95% CI: 91.96%–95.60%, $p < .001$). To further investigate the causes of heterogeneity, we conducted a subgroup analysis of various variables, including the incorporation of pathology or MVI, the timing of tumour recurrence, and the type of medical imaging. Studies using pathological data demonstrated higher Se compared to those without pathological data (91%, 95% CI: 82%–100% vs.75%, 95% CI: 68%–81%; see Figure 7A), Interestingly, studies including MVI exhibited slightly lower Se and Sp compared to those without MVI (Se, 71%, 95% CI: 62%–79% vs. 78%, 95% CI: 72%–85%; Sp, 77%, 95% CI: 69%–86% vs. 85%, 95% CI: 79%–90%; see Figure 7B). The time pattern of recurrence for HCC was divided into early (within two years of operation) and late recurrence (>2 years after resection). Studies defined

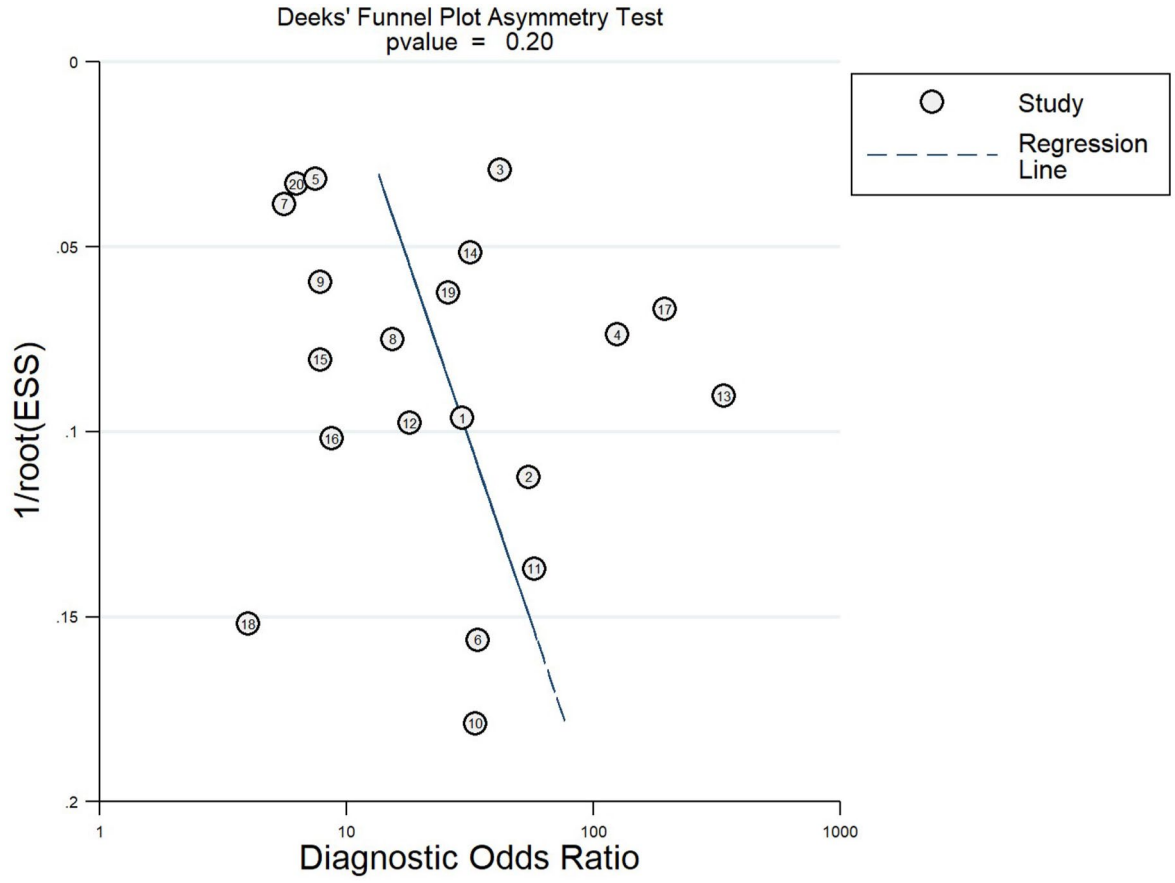


Figure 3. Deeks' funnel plot for publication bias assessment. The number within each circle indicates the study ID. ESS, effective sample size.

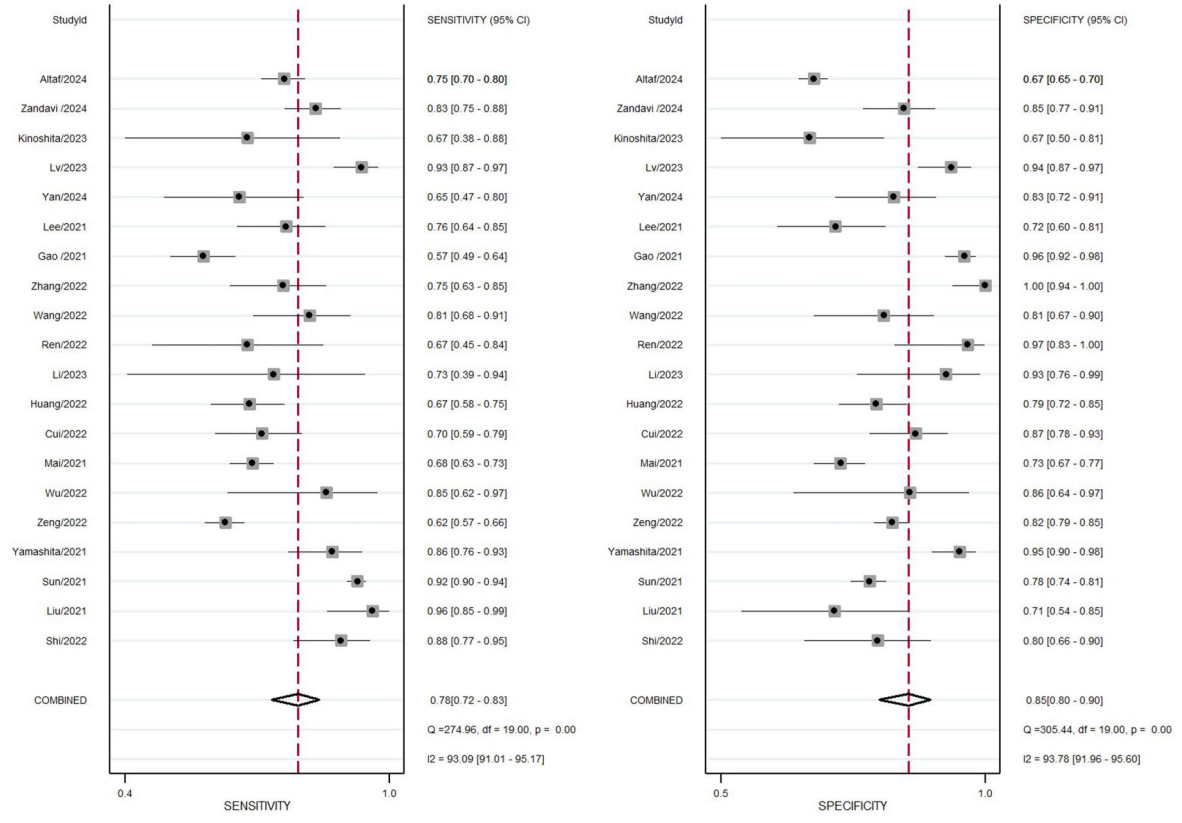


Figure 4. Forest plot showing sensitivity and specificity estimates of the included studies. Vertical lines indicate pooled estimates. $I^2 > 50\%$ indicates significant heterogeneity across studies.

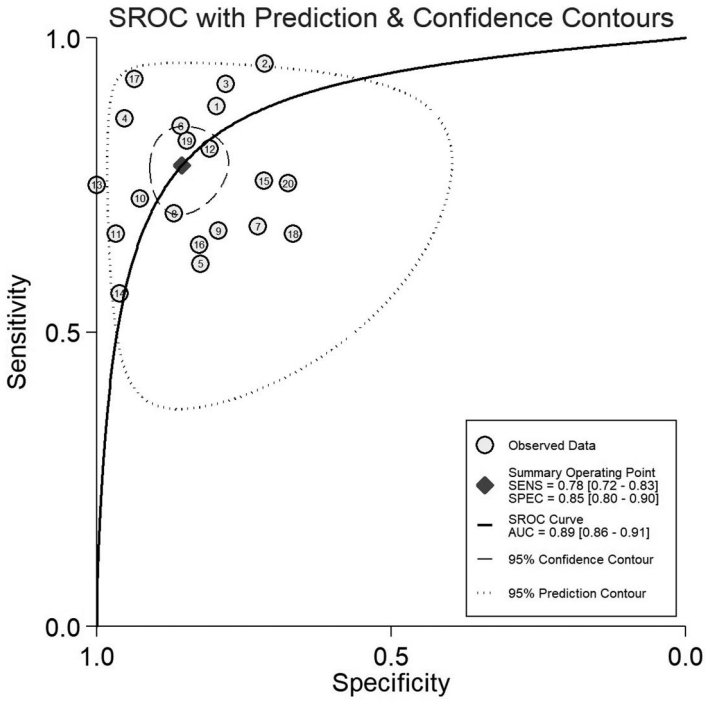


Figure 5. Summary receiver operating characteristic (SROC) curve. Each circle represents an included study, with the 95% confidence interval (CI) shown in parentheses. AUC, area under the curve.

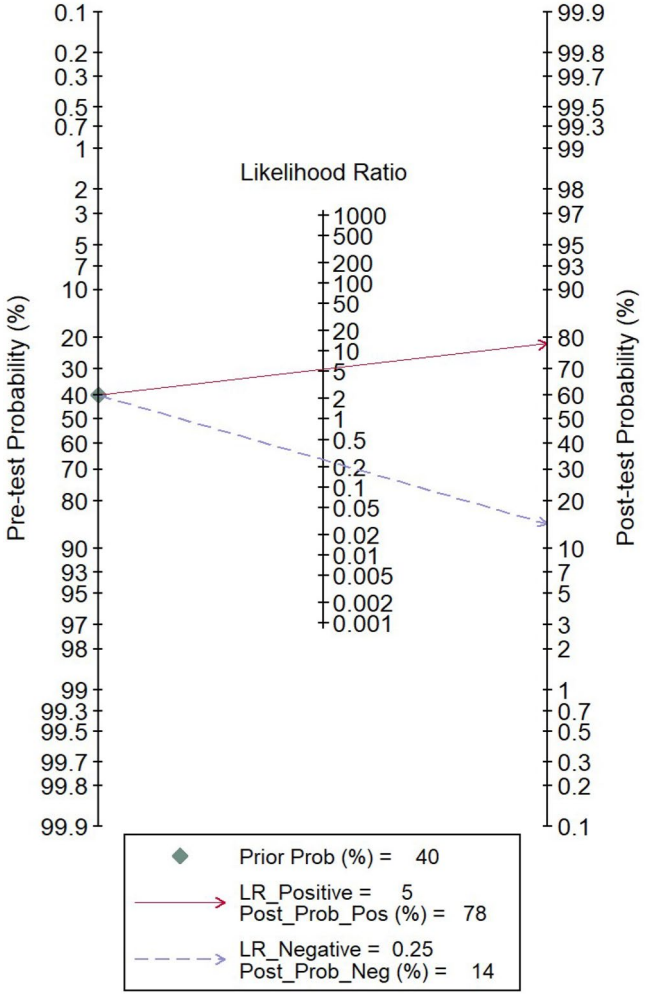


Figure 6. Fagan nomogram. LR, likelihood ratio; prob, probability; pos, positive; neg, negative.

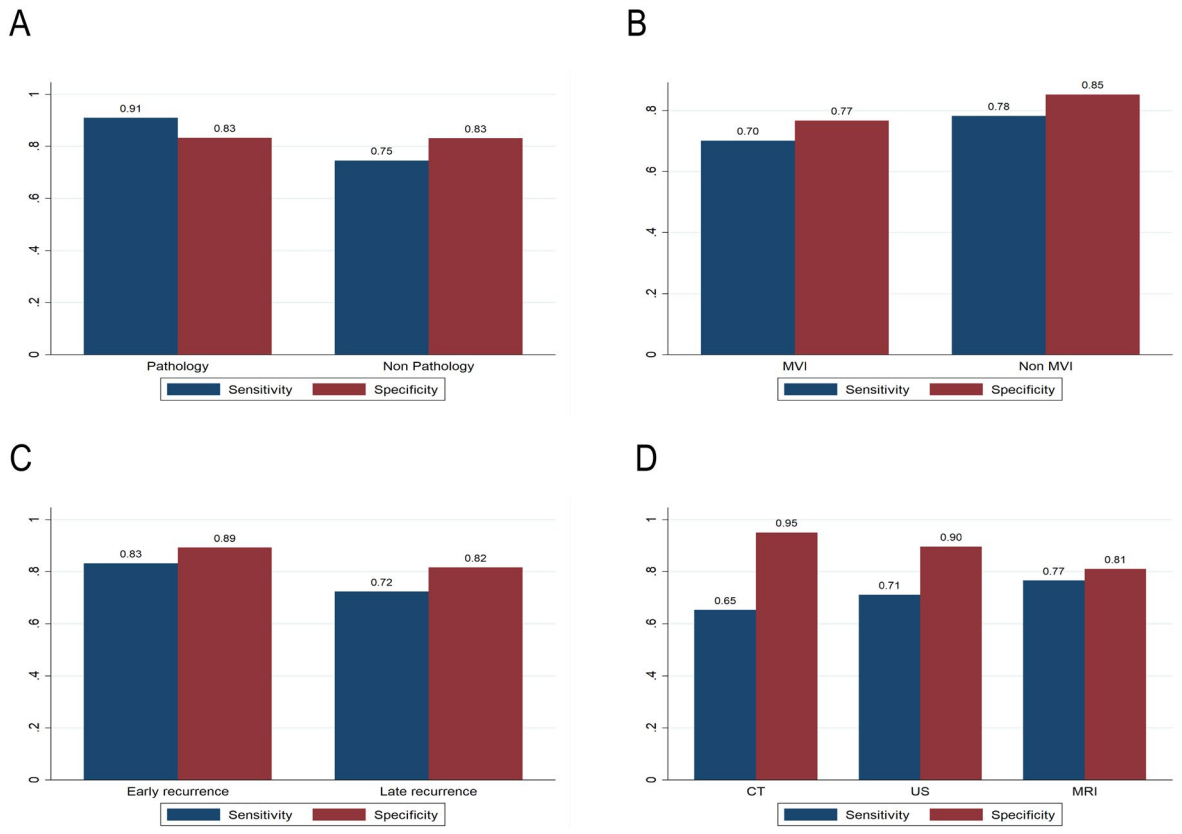


Figure 7. Bar charts comparing sensitivity and specificity across subgroups: (A) pathology versus non-pathology; (B) MVI versus non-MVI; (C) early recurrence versus late recurrence; (D) different imaging modalities (CT, US, and MRI). MVI, microvascular invasion; CT, computed tomography; US, ultrasound; MRI, magnetic resonance imaging.

as late recurrence demonstrated higher Se and Sp than those characterized as early recurrence (Se, 87%, 95% CI: 82%–91% vs. 72%, 95% CI: 65%–79%; Sp, 89%, 95% CI: 82%–97% vs. 80%, 95% CI: 74%–86%; see Figure 7C). Additionally, studies using MRI-based radiomics achieved superior Sp compared to those with CT or ultrasound (96%, 95% CI: 94%–98% vs. 82%, 95% CI: 75%–89% vs. 79%, 95% CI: 75%–84%), though showing the lowest Se (64%, 95% CI: 53%–74%, see Figure 7D) The Forest Plots of their Sp and Se are shown in the Supplementary Figure S1.

Discussion

Our meta-analysis indicated that the pooled sensitivity and specificity were 78% and 85%, respectively, with an AUC of 0.89. To the best of our knowledge, this is the first meta-analysis on AI-assisted prediction of the risk of posthepatectomy recurrence of HCC patients, demonstrating the significant accuracy of prediction using AI in this field. Based on this study, AI was four times more likely to identify recurrence than to mistakenly identify recurrence correctly, and the posttesting probability of a positive test result was 78%. Similarly, the likelihood of misjudging nonrecurrence was 0.25 times higher than that of correctly judging nonrecurrence, and the posttesting probability of a negative test result was 14%. Our findings show that AI has the potential to predict the posthepatectomy recurrence risk of HCC, which is beneficial for the identification of high-risk patients following surgery and the development of optimal prevention and treatment strategies for them.

Compared to conventional methods, AI-based models exhibit significantly higher accuracy in predicting postoperative recurrence of HCC, with their advantages in the integration of multimodal data and algorithmic innovation. AI-based models not only incorporate clinical and pathological parameters but also include imaging features extracted quantitatively through radiomics approaches, which reveal the characteristics of cancer comprehensively. A scoping review showed that multi-modal machine learning

was more robust in its estimations over unimodal methods, with an average increase of 6.4% in predictive accuracy [33]. In contrast to linear models that struggle to identify complex links between variables, AI-based models leverage various nonlinear algorithms to detect their hidden correlations and provide adaptive, high-precision, and real-time risk prediction, such as gradient-boosted decision trees and convolutional neural networks [34,35]. Notably, AI-based models have demonstrated remarkable performance in postoperative recurrence prediction across multiple malignancies. Xiang et al. developed a multimodal transformer with Unified maSKed modeling based on clinical data and pathological images and achieved the highest AUC of 0.833 in predicting five-year recurrence risk for melanoma [36]. In addition, an artificial neural network model gained an AUC value of 0.979 for the prediction of recurrence risk in patients with T1 rectal cancer following local excision [37]. Therefore, AI-based models probably have a general potential in the risk prediction for the recurrence of various malignancies after surgery. Furthermore, our analysis demonstrates that models relying exclusively on preoperative variables can achieve predictive accuracy comparable to those incorporating postoperative data. Such models provide robust prognostic estimates prior to surgery, enabling timely risk stratification, personalized perioperative planning, and informed patient counseling. Restricting inputs to preoperative data also enhances generalizability by eliminating variability introduced through postoperative pathological assessments and institutional practices, and offers particular advantages in settings where histopathology may be delayed or inaccessible. Harnessing advanced machine learning to derive prognostic insights from preoperative clinical, laboratory, and imaging features supports a paradigm shift toward more timely, proactive, and individualized decision-making in surgical oncology.

Comparative evaluations of AI-generated predictions versus traditional physician assessments have emerged as a critical topic amid the ongoing digital transformation in healthcare. By integrating multidimensional data and capturing complex nonlinear relationships, AI systems offer notable advantages in standardizing structured tasks while mitigating subjective bias. A recent meta-analysis covering diverse AI models across multiple specialties demonstrated that AI performance [38], in certain diagnostic contexts, approaches that of general physicians. However, in complex or ambiguous cases, AI remains outperformed by expert clinicians, whose expertise in interpretability, adaptive decision-making, and patient-specific contextualization remains indispensable. Nonetheless, AI holds significant promise for improving healthcare access in resource-constrained settings and advancing medical education. As AI becomes more deeply integrated into clinical workflows, its core value is to augment—rather than replace—clinical judgment. Future developments should prioritize multimodal data integration and specialty-specific model refinement to enhance robustness, interpretability, and clinical generalizability, thereby enabling AI to evolve from an 'auxiliary tool' into a reliable 'clinical decision support partner'.

Despite promising in predicting postoperative recurrence of HCC, the positive predictive value and negative predictive value of AI-based models are still not satisfactory, which may be attributed to unclear pathogenesis of HCC, insufficient model generalizability, and lack of dynamic capabilities. The molecular mechanisms underlying the recurrence of HCC are still obscure, which leads to debates regarding the prognostic value of certain clinical, pathological, and imaging biomarkers. These uncertainties may compromise the decision-making accuracy of AI-based models that predominantly rely on statistical correlations. Additionally, data heterogeneity, especially demographic characteristics such as the mobility of chronic liver diseases, would undermine the model's generalizability across various populations. Moreover, predominantly based on static preoperative data, current models underestimate dynamic postoperative variables, such as the adherence to longitudinal follow-up and changes in the immune status of patients. Therefore, to enhance the predictive performance of AI-based models, a series of improvements is required, including further studies to better understand the pathogenesis of HCC recurrence, the establishment of global multimodal databases to strengthen the model generalizability, and dynamic models to enable real-time risk assessment.

Our study showed that the included studies exhibited heterogeneity, which may be related to the incorporation of pathology or MVI, timing of tumour recurrence, and type of medical imaging. Pathological parameters (e.g. histological differentiation and MVI) directly reflect tumor biology, enabling higher sensitivity in recurrence prediction models. Detection of late recurrence benefits from distinct features (e.g. extrahepatic spread and multifocal lesions), while examination of early recurrence remains challenging as a result of difficulties in imaging diagnosis, such as identification of micro-metastases [39,40]. Imaging

modality selection requires sensitivity-specificity tradeoffs: MRI gains an advantage over lesion characterizations through superior soft-tissue resolution, whereas contrast-enhanced ultrasound (CEUS) demonstrates sensitivity advantages in the detection of micro-lesions *via* dynamic perfusion monitoring [41,42]. Algorithmically, the integration of the multi-algorithm approach achieves balanced performance through feature complementarity, error compensation, and adaptive mechanisms. Hence, future developments should prioritize lightweight ensemble architectures combined with explainability enhancement to facilitate the clinical adoption of AI-based models.

This study still has several limitations. First, the analysis only included retrospective studies and lacked high-quality prospective studies such as randomized controlled clinical trials. Second, only two studies applied time-to-event survival models, which are conceptually distinct from fixed-time classification models and therefore were summarized narratively rather than pooled quantitatively. Third, only a small proportion of studies reported decision thresholds, and the reported values varied considerably. Given the influence of threshold selection on sensitivity and specificity, such variability and incomplete reporting likely increased heterogeneity and reduced comparability. Finally, since AI-based models are typically developed through black-box methods, AI is primarily in charge of the process. Although humans are aware of the specific operations of each step in a model's decision-making process, the operations cannot be expressed in human-readable semantics, which might prevent us from specifically analyzing the impact of certain factors. Therefore, large-scale, prospective, multicenter, randomized controlled studies using transparent and standardized modeling approaches are needed to improve both performance and interpretability in predicting posthepatectomy HCC recurrence.

Conclusions

In conclusion, our meta-analysis demonstrates the high accuracy of AI in predicting the posthepatectomy recurrence of HCC. AI is a promising noninvasive method for predicting posthepatectomy HCC recurrence, with substantial implications for screening high-risk patients and developing adjuvant therapies throughout postoperative monitoring.

Author contributions

CRediT: **Zhiqiang Xiang**: Conceptualization, Data curation, Formal analysis, Writing – original draft; **Jing Deng**: Data curation, Investigation; **Hao Liang**: Data curation, Methodology, Software, Visualization; **Mengliang Jiang**: Data curation, Investigation, Software; **Yuhan Liang**: Data curation, Formal analysis; **Zhaohai Liu**: Validation; **Yachen Wu**: Validation; **Leyuan Peng**: Supervision, Writing – review & editing; **Xiaoming Dai**: Funding acquisition, Supervision, Writing – review & editing; **Zhu Zhu**: Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The study was funded by the Clinical Medicine Technological Innovation Leading Project of Hunan province (2020SK51818) and the Natural Science Foundation of Hunan province (2022JJ70119).

ORCID

Zhu Zhu http://orcid.org/0000-0003-0274-8007

Data availability statement

All data generated or analysed in this study are contained within the published article and its supplementary information files. The underlying raw data used for the statistical analyses are available from the corresponding author (Zhu Zhu) on reasonable request.

References

[1] Siegel RL, Miller KD, Fuchs HE, et al. Cancer statistics, 2022. CA Cancer J Clin. 2022;72(1):7–33. doi:10.3322/caac.21708.

[2] Dhir M, Melin AA, Douaiher J, et al. A review and update of treatment options and controversies in the management of hepatocellular carcinoma. Ann Surg. 2016;263(6):1112–1125. doi:10.1097/SLA.0000000000001556.

[3] Xiang ZQ, Wu YC, Qu XL, et al. Laparoscopic hepatectomy for hepatocellular carcinoma in patients with clinically significant portal hypertension: a systematic review and meta-analysis. World J Surg Oncol. 2024;22(1):3. doi:10.1186/s12957-023-03264-7.

[4] Dai XM, Xiang ZQ, Wang Q, et al. Oncological outcomes of anatomic versus non-anatomic resections for small hepatocellular carcinoma: systematic review and meta-analysis of propensity-score matched studies. World J Surg Oncol. 2022;20(1):299. doi:10.1186/s12957-022-02770-4.

[5] Xiang ZQ, Zhu FF, Zhao SQ, et al. Laparoscopic versus open repeat hepatectomy for recurrent hepatocellular carcinoma: a systematic review and meta-analysis of propensity score-matched cohort studies. Int J Surg. 2023;109(4):963–971. doi:10.1097/JS9.0000000000000305.

[6] Suh SW, Lee KW, Lee JM, et al. Prediction of aggressiveness in early-stage hepatocellular carcinoma for selection of surgical resection. J Hepatol. 2014;60(6):1219–1224. doi:10.1016/j.jhep.2014.01.027.

[7] Suehiro T, Sugimachi K, Matsumata T, et al. Protein induced by vitamin K absence or antagonist II as a prognostic marker in hepatocellular carcinoma. Comparison with alpha-fetoprotein. Cancer. 1994;73(10):2464–2471. doi:10.1002/1097-0142(19940515)73:10<2464::AID-CNCR2820731004>3.0.CO;2-9.

[8] Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018;2(10):719–731. doi:10.1038/s41551-018-0305-z.

[9] Altaf A, Endo Y, Munir MM, et al. Impact of an artificial intelligence based model to predict non-transplantable recurrence among patients with hepatocellular carcinoma. HPB (Oxford). 2024;26(8):1040–1050. doi:10.1016/j.hpb.2024.05.006.

[10] Cui H, Wang KY, Li WY, et al. CT images-based 3D convolutional neural network to predict early recurrence of solitary hepatocellular carcinoma after radical hepatectomy. Diagn Interv Radiol. 2022;28(6):524–531. doi:10.5152/dir.2022.201097.

[11] Gao W, Wang W, Song D, et al. A predictive model integrating deep and radiomics features based on gadobenate dimeglumine-enhanced MRI for postoperative early recurrence of hepatocellular carcinoma. Radiol Med. 2022;127(3):259–271. doi:10.1007/s11547-021-01445-6.

[12] Huang Z, Shu Z, Zhu RH, et al. Deep learning-based radiomics based on contrast-enhanced ultrasound predicts early recurrence and survival outcome in hepatocellular carcinoma. World J Gastrointest Oncol. 2022;14(12):2380–2392. doi:10.4251/wjgo.v14.i12.2380.

[13] Kinoshita M, Ueda D, Matsumoto T, et al. Deep learning model based on contrast-enhanced computed tomography imaging to predict postoperative early recurrence after the curative resection of a solitary hepatocellular carcinoma. Cancers (Basel). 2023;15(7):2140. doi:10.3390/cancers15072140.

[14] Lee IC, Huang JY, Chen TC, et al. Evolutionary learning-derived clinical-radiomic models for predicting early recurrence of hepatocellular carcinoma after resection. Liver Cancer. 2021;10(6):572–582. doi:10.1159/000518728.

[15] Li Q, Wei Y, Zhang T, et al. Predictive models and early postoperative recurrence evaluation for hepatocellular carcinoma based on gadoxetic acid-enhanced MR imaging. Insights Imaging. 2023;14(1):4. doi:10.1186/s13244-022-01359-5.

[16] Liu IT, Yen CS, Wang WL, et al. Predict early recurrence of resectable hepatocellular carcinoma using multi-dimensional artificial intelligence analysis of liver fibrosis. Cancers (Basel). 2021;13(21):5323. doi:10.3390/cancers13215323.

[17] Lv C, He N, Yang JJ, et al. Prediction of 3-year recurrence rate of hepatocellular carcinoma after resection based on contrast-enhanced CT: a single-centre study. Br J Radiol. 2023;96(1145):20220702. doi:10.1259/bjr.20220702.

[18] Mai RY, Zeng J, Meng WD, et al. Artificial neural network model to predict post-hepatectomy early recurrence of hepatocellular carcinoma without macroscopic vascular invasion. BMC Cancer. 2021;21(1):283. doi:10.1186/s12885-021-07969-4.

[19] Ren Y, Bo L, Shen B, et al. Development and validation of a clinical-radiomics model to predict recurrence for patients with hepatocellular carcinoma after curative resection. Med Phys. 2023;50(2):778–790. doi:10.1002/mp.16061.

[20] Shi HY, Lee KT, Chiu CC, et al. 5-year recurrence prediction after hepatocellular carcinoma resection: deep learning vs. Cox regression models. Am J Cancer Res. 2022;12(6):2876–2890.

[21] Sun LY, Ouyang Q, Cen WJ, et al. A model based on artificial intelligence algorithm for monitoring recurrence of hcc after hepatectomy. Am Surg. 2023;89(5):1468–1478. doi:10.1177/00031348211063549.

[22] Wang L, Wu M, Zhu C, et al. Ensemble learning based on efficient features combination can predict the outcome of recurrence-free survival in patients with hepatocellular carcinoma within three years after surgery. Front Oncol. 2022;12:1019009. doi:10.3389/fonc.2022.1019009.

[23] Wu C, Yu S, Zhang Y, et al. CT-based radiomics nomogram improves risk stratification and prediction of early recurrence in hepatocellular carcinoma after partial hepatectomy. Front Oncol. 2022;12:896002. doi:10.3389/fonc.2022.896002.

[24] Yamashita R, Long J, Saleem A, et al. Deep learning predicts postsurgical recurrence of hepatocellular carcinoma from digital histopathologic images. Sci Rep. 2021;11(1):2047. doi:10.1038/s41598-021-81506-y.
[25] Yan X, Li Y, Qin W, et al. Radiomics model based on contrast-enhanced computed tomography imaging for early recurrence monitoring after radical resection of AFP-negative hepatocellular carcinoma. BMC Cancer. 2024;24(1):700. doi:10.1186/s12885-024-12436-x.
[26] Zandavi SM, Kim C, Goodwin T, et al. AI-powered prediction of HCC recurrence after surgical resection: personalised intervention opportunities using patient-specific risk factors. Liver Int. 2024;44(10):2724–2737. doi:10.1111/liv.16050.
[27] Zeng J, Zeng J, Lin K, et al. Development of a machine learning model to predict early recurrence for hepatocellular carcinoma after curative resection. Hepatobiliary Surg Nutr. 2022;11(2):176–187. doi:10.21037/hbsn-20-466.
[28] Zhang H, Huo F. Prediction of early recurrence of HCC after hepatectomy by contrast-enhanced ultrasound-based deep learning radiomics. Front Oncol. 2022;12:930458. doi:10.3389/fonc.2022.930458.
[29] McInnes MDF, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA. 2018;319(4):388–396. doi:10.1001/jama.2017.19163.
[30] Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557–560. doi:10.1136/bmj.327.7414.557.
[31] Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med. 1993;12(14):1293–1316. doi:10.1002/sim.4780121403.
[32] Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005;58(9):882–893. doi:10.1016/j.jclinepi.2005.01.016.
[33] Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: a scoping review. NPJ Digit Med. 2022;5(1):171. doi:10.1038/s41746-022-00712-8.
[34] Shin H, Hur MH, Song BG, et al. AI model using CT-based imaging biomarkers to predict hepatocellular carcinoma in patients with chronic hepatitis B. J Hepatol. 2025;82(6):1080–1088.
[35] Moons KGM, Damen JAA, Kaul T, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ. 2025;388:e082505. doi:10.1136/bmj-2024-082505.
[36] Xiang J, Wang X, Zhang X, et al. A vision-language foundation model for precision oncology. Nature. 2025;638(8051):769–778. doi:10.1038/s41586-024-08378-w.
[37] Su J, Liu Z, Li H, et al. Artificial intelligence-based model to predict recurrence after local excision in T1 rectal cancer. Eur J Surg Oncol. 2025;51(6):109717. doi:10.1016/j.ejso.2025.109717.
[38] Takita H, Kabata D, Walston SL, et al. A systematic review and meta-analysis of diagnostic performance comparison between generative AI and physicians. NPJ Digit Med. 2025;8(1):175. doi:10.1038/s41746-025-01543-z.
[39] Chen M, Cao J, Hu J, et al. Clinical-radiomic analysis for pretreatment prediction of objective response to first transarterial chemoembolization in hepatocellular carcinoma. Liver Cancer. 2021;10(1):38–51. doi:10.1159/000512028.
[40] Zhang L, Chen J, Jiang H, et al. MR elastography as a biomarker for prediction of early and late recurrence in HBV-related hepatocellular carcinoma patients before hepatectomy. Eur J Radiol. 2022;152:110340. doi:10.1016/j.ejrad.2022.110340.
[41] Mao Y, Wang J, Zhu Y, et al. Gd-EOB-DTPA-enhanced MRI radiomic features for predicting histological grade of hepatocellular carcinoma. Hepatobiliary Surg Nutr. 2022;11(1):13–24. doi:10.21037/hbsn-19-870.
[42] Huang Z, Zhou P, Li S, et al. MR versus CEUS LI-RADS for distinguishing hepatocellular carcinoma from other hepatic malignancies in high-risk patients. Ultrasound Med Biol. 2021;47(5):1244–1252. doi:10.1016/j.ultrasmedbio.2021.01.020.