Discover Oncology

REVIEW Open Access

Artificial intelligence-driven pathomics in hepatocellular carcinoma: current developments, challenges and perspectives

Wei Ding[1†], Jinxing Zhang[2†], Zhicheng Jin[3†], Hongjin Hua[4†], Qingquan Zu[2], Shudong Yang[5], Weidong Wang[1], Sheng Liu[2], Haifeng Zhou[2*] and Haibin Shi[2*]

[†]Wei Ding, Jinxing Zhang, Zhicheng Jin and Hongjin Hua are joint first authors.

*Correspondence:
Haifeng Zhou
hfzhou_ir@sina.com
Haibin Shi
shihb@njum.edu.cn

Full list of author information is available at the end of the article

Abstract

Hepatocellular carcinoma (HCC) is a highly malignant tumor with elevated incidence and mortality rates globally. Its complex etiology and pronounced heterogeneity present significant challenges in diagnosis and treatment. Recent advancements in artificial intelligence (AI) have demonstrated transformative potential to usher a new wave of precision oncology. Pathomics, an AI-based digital pathology technique, facilitates the extraction of extensive datasets from whole-slide histopathological images, enabling quantitative analyses to improve diagnosis, treatment, and prognostic prediction for HCC. Furthermore, emerging pathological foundation models are revolutionizing traditional paradigms and providing a robust framework for the development of specialized pathomics models tailored to specific clinical tasks in HCC. Despite its promise, pathomics research in HCC remains in its infancy, with clinical implementation hindered by challenges such as data heterogeneity, model interpretability, ethical concerns, regulatory issues, and the absence of standardized industry protocols. Future initiatives should prioritize the conduction of prospective multi-center studies, the integration of multi-modal data, the enhancement of regulatory frameworks, and the establishment of industry-wide standardized guidelines and compliant platform infrastructures to accelerate the clinical adoption of pathomics for personalized HCC treatment.

Keywords Hepatocellular carcinoma, Pathomics, Artificial intelligence, Pathological foundation model, Diagnosis, Prognosis, Challenge

1 Introduction

Hepatocellular carcinoma (HCC) ranks as the sixth most common cancer and the third leading cause of cancer-related deaths worldwide. In 2022, there were over 860,000 new cases and 750,000 deaths globally (GLOBOCAN) [1], with both incidence and mortality rates anticipated to rise further [2]. Surgical resection with curative intent remains the first-line treatment for early-stage HCC; however, the recurrence rate within five years post-surgery remains as high as 70% [3]. Tragically, approximately two-thirds of patients present with intermediate or advanced disease at diagnosis [4]. Current therapeutic

strategies for HCC are largely based on staging systems like the Barcelona Clinic Liver Cancer (BCLC) classification [5]; however, these methods lack the precision and personalization essential for modern oncology and fall short of meeting the demands of precision medicine. Therefore, there is an urgent clinical need for more precise prognostic biomarkers [6]. Additionally, the complex etiology and marked heterogeneity of HCC further complicate diagnosis and treatment [7, 8].

In recent years, artificial intelligence (AI) technologies, including machine learning (ML) and deep learning (DL), have seen rapid development in oncology, demonstrated by the exponential growth in AI-driven oncological research and applications [9–11]. Pathomics, an emerging technology for intelligent analysis of histopathological images, leverages AI algorithms to extract high-throughput biological features from whole-slide images (WSIs) that are otherwise imperceptible. This enables the visualization of tumor heterogeneity, assisting in tumor detection, histological subtyping, pathological grading, and clinical prognostication [12–14]. Compared to traditional variable-based models, pathomics models (PMs) exhibit superior predictive performance and cost-effectiveness. Their broad applicability and translational potential have been demonstrated across various cancers, including breast, colorectal, lung, renal, and prostate cancer [15–20]. Furthermore, pathomics can be integrated with radiomics, genomics, and clinical data to develop multi-modal AI models, facilitating a more comprehensive analysis of tumor biology, deeper insights into disease mechanisms, and more informed therapeutic decision-making.

Despite the increasing number of studies applying pathomics in HCC, both in mechanistic research and clinical translation, the field remains in its early stage, with several bottlenecks and challenges ahead. This narrative review aims to provide a clinically oriented overview of recent advancements in pathomics applications for HCC in the AI era, the current challenges, and future research directions.

2 The workflow of pathomics

The pathomics pipeline involves several critical steps [21], as illustrated in Fig. 1. The first step is the acquisition and preprocessing of WSIs, where researchers use high-resolution scanning devices to digitize tissue slides, followed by standardization to ensure the production of high-quality, uniform images. The second step is the segmentation and

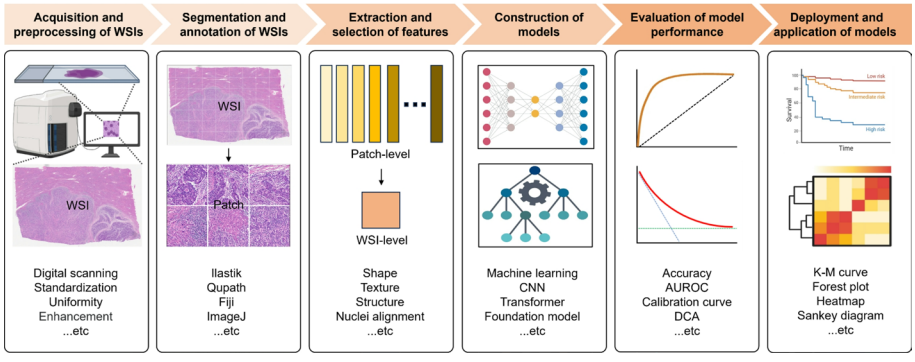

Fig. 1 Workflow of pathomics in hepatocellular carcinoma. This includes multiple stages such as the acquisition and preprocessing of WSIs, segmentation and annotation of WSIs, feature extraction and selection, model construction, model performance evaluation, and eventual model deployment. *WSI* whole-slide image; *CNN* convolutional neural network; *AUROC* area under the receiver operating characteristic curve; *DCA* decision curve analysis; *K-M* Kaplan-Meier

annotation of WSIs. Each WSI is subdivided into multiple patches, and regions of interest are meticulously annotated, which is crucial for the accurate extraction of pathomics features in subsequent step. The third step involves feature extraction and selection, where quantitative analysis of tissue structures such as cellular morphology, nuclei, and cytoplasm within target regions is performed. Features are extracted at the patch level using pretrained models and subsequently aggregated into slide-level features via attention-based weighted averaging mechanisms. The fourth step is model construction, which forms the core of the entire pipeline. Appropriate model architectures are selected to train on the extracted features, with pretrained PFMs typically yielding optimal performance. Following this, model performance is evaluated using various metrics, including accuracy, precision, F1 score, area under the receiver operating characteristic curve (AUROC), and decision curve analysis (DCA). These evaluations are vital for validating the feasibility and reliability of PMs, serving as essential prerequisites for the final step of clinical deployment and translation.

Instead of being limited to single-modality pathomics, Fig. 2 attempts to broaden this pipeline by integrating pathomics features with complementary modalities, including radiomics, genomics, transcriptomics and clinical records, into a cohesive multi-modal framework that enriches feature characterization and underpins the development of more accurate diagnostic, prognostic and therapeutic models.

3 Model architectures and key parameter choices

In AI-driven HCC pathomics studies, researchers rely on two main architectures, convolutional neural networks (CNNs) and vision transformers (ViTs), to tackle three core challenges: gigapixel-scale whole-slide images, diverse tumor patterns, and noisy clinical labels. CNN backbones (e.g., 50-layer ResNet or EfficientNet-B0, pretrained on ImageNet) remain the bedrock for patch-level feature extraction. Slides are divided into 224×224 to 512×512 px tiles, a size large enough to capture cellular detail and tissue context within typical graphics processing unit (GPU) limits [22, 23]. To model

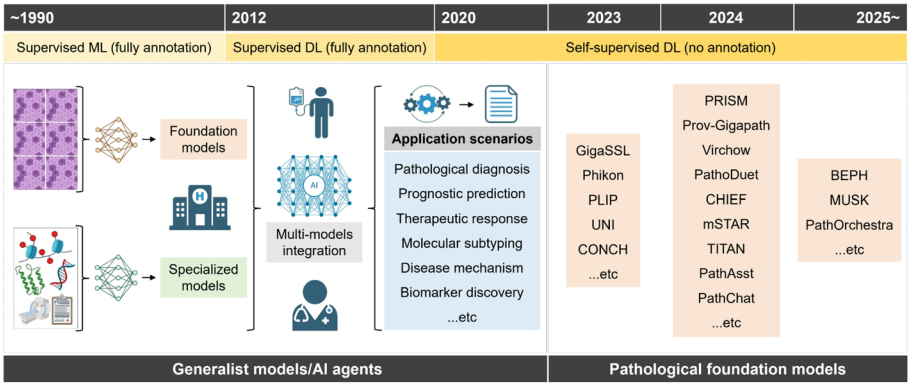

Fig. 2 Development of pathological foundation models and the emergence of future generalist models/AI agents. Since 2020, self-supervised learning has enabled training of pathological foundation models on large-scale, unlabeled datasets, reducing the need for task-specific annotations during fine-tuning downstream task models. Current research emphasizes the integration of multi-modal data, including pathology slides, radiological images, genomic and transcriptomic profiles, and clinical records, in order to develop specialized models for diverse professional tasks. In oncology, the ultimate goal for foundation models is to create generalist models that merge general and specialized capabilities, also referred to as AI agents, capable of assisting in diagnosis, treatment, and prognosis, as well as delivering analytical insights, medical interpretation, and interactive support. *ML* machine learning; *DL* deep learning; *AI* artificial intelligence

long-range tissue structure, newer pipelines introduce ViT-base transformers (12 layers) that tokenize slides into 16×16 or 32×32 px patches, learning relationships across distant regions [22]. Once extracted, patch features are aggregated via attention-based multi-instance learning (MIL) modules: a 128-256-dimensional attention head scores each patch, patches are grouped into 3–5 clusters, and the top 10–20 scores are pooled. This mechanism downweights uninformative areas (e.g., blank space or artifacts) and amplifies diagnostically or prognostically relevant regions, yielding a robust slide-level prediction.

To ensure stable training on small, heterogeneous HCC cohorts, common settings include the Adam optimizer with weight decay (1×10^{-5} to 1×10^{-4}), a learning rate of 1×10^{-4} to 1×10^{-5} (often cosine-annealed), and batch sizes of 16–32 tiles per GPU [22, 24]. Models typically train for 20–50 epochs, use ReLU activations throughout, and apply 20–50% dropout in final layers to prevent overfitting [25]. Each slide contributes 200–1000 "valid" tiles selected via Otsu thresholding and blur/artifact filtering to cover representative histology at a manageable compute cost [22]. Finally, a label-denoising step is introduced to reduce noise, for example by excluding patients who experienced ultra-early recurrences under 3 months from the training set [26].

These parameter choices and architectural designs are informed by prior literature and extensive empirical validation, aiming to maximize accuracy, stability, and clinical applicability. Understanding these technical details facilitates the reproducibility, benchmarking, and adaptation of AI models in diverse clinical scenarios.

4 Basic applications of pathomics in HCC: pathological diagnosis

Conventional histopathological assessment is inherently subjective, leading to inter-observer variability. AI-based pathomics has emerged as a transformative technology for quantifying tissue characteristics in pathology. The primary advantage of pathomics in basic applications lies in its ability to automate tasks that traditionally rely on manual interpretation, such as object detection, tumor segmentation, histological classification, pathological grading, anomaly detection, and quality control [27] (Table 1).

4.1 Tumor identification

While tumor detection does not directly yield a definitive diagnosis, it is a foundational task in the pathological diagnostic workflow and crucial for following judgment of the pathological nature. Notably, AI excels particularly in distinguishing tumor from non-tumor regions.

Liao et al. [28] employed a CNN-based DL method to analyze WSIs of HCC samples sourced from public databases. This pathomics technology accurately differentiated between normal and tumor tissues, enabling automated localization of HCC and achieving a perfect AUROC of 1. This allowed pathologists to complete their diagnoses based on AI-assisted identification. Similarly, other researchers utilized a global label DL framework combining transfer learning and multi-instance learning to classify histopathological images of HCC with high accuracy. In the test dataset, all performance metrics exceeded 0.95, with precision reaching 1. In the validation dataset, both recall and precision were 0.99, demonstrating strong potential to support precise HCC diagnosis [29]. More recently, a study by Troiano et al. [30] demonstrated that pathomics techniques based on nuclear fluorescent staining could automatically detect HCC cells,

Table 1 Summary of basic applications for pathological diagnosis via pathomics in hepatocellular carcinoma

Reference	Tasks	Study design	Number of patients/WSIs (tiles) and validation methods	Details and outcome measures
Liao et al. [28]	Tumor identification	Retrospective, single-center, and public database	Training (TCGA): 408 WSIs Internal test (TCGA): 73 WSIs Internal validation: 455 pts/719 dots (11 TMAs)	Performance of PM: AUROC of 0.998-1.000 (test) and 0.902–0.971 (validation)
Sun et al. [29]	Tumor identification	Retrospective, public database	Training (TCGA): 324 WSIs Internal test (TCGA): 69 WSIs Internal validation (TCGA): 69 WSIs	Performance of PM: all measures exceeded 0.95, especially precision of 1 (test); accuracy of 0.98 and precision of 0.99 (validation)
Troiano et al. [30]	Tumor nuclei identification	Retrospective, single-center	Training: 13,322 nuclei images (TMAs) Internal test: 3806 WSIs Internal validation: 1903 WSIs	Nuclear fluorescent staining-based PMs developed by ML and DL algorithms, respectively Performance of PMs: DL outperformed ML, accuracy of 88.10% vs. 62.12%, AUROC of 0.950 vs. 0.664, respectively
Cheng et al. [31]	Histological classification	Retrospective, multi-center	Training: 330 pts/454 WSIs Internal test: 132 pts/195 WSIs External validation: 234 pts/422 WSIs	Classify 7-category hepatocellular nodular lesions Performance of AI-based PM (HnAIM): AUROC of 0.9991 (internal test) and 0.935 (external validation)
Kiani et al. [32]	Histological classification	Retrospective, single-center, and public database	Training (TCGA): 20 WSIs Internal test (TCGA): 50 WSIs External validation: 80 WSIs	Differentiate between two subtypes of PLC (HCC and ICC) Performance of PM: accuracy of 0.885 (internal test) and 0.842 (external validation)
Jang et al. [33]	Histological classification	Retrospective, multi-center	Training: 480 WSIs Internal test: 320 WSIs External validation: 60 WSIs	Distinguish among HCC, ICC and metastatic colorectal cancer Performance of PM (AI diagnostic assistant): AUROC above 0.95 (both internal and entire dataset test) but of 0.745 (external validation)
Calderaro et al. [34]	Histological classification	Retrospective, multi-center, and public database	Training and internal test: 591 pts/1024 WSIs External validation (TCGA): 360 pts	Reclassify cHCC-CCA to HCC or ICC Performance of DL-based AI model: AUROC of 0.99 (internal test) and 0.94 (external validation)
Chen et al. [35]	Pathological grading	Retrospective, single-center	Training and internal test: 74 pts/444 WSIs	Classify multi-differentiated (poorly, moderately and well) HCC Performance of PM: SENet model proved with best accuracy of 95.27% among five DL models

Table 1 (continued)

Reference	Tasks	Study design	Number of patients/WSIs (tiles) and validation methods	Details and outcome measures
Chen et al. [36]	Pathological grading and tumor identification	Retrospective, single-center, and public database	Training and internal test (TCGA): 491 WSIs External validation: 101 WSIs	PM developed by DL algorithm (Inception V3) Performance of PM: accuracy of 96.0% accuracy for benign and malignant classification; that of 89.6% for poor, moderate and well tumor differentiation
Lin et al. [37]	Pathological grading	Retrospective, single-center	Training and internal test: 113 pts/217 WSIs (3472 tiles)	A multiphoton microscopy-based PM established by DL algorithm Performance of PM: accuracy over 90% and AUROC around 0.9 for classifying differentiation grades of HCC
Sun et al. [39]	MVI identification	Retrospective, single-center	Training: 56 pts/WSIs Internal test: 13 pts/WSIs	MVI regions identified by classification of MVI boundaries Performance of PM (PCformer): accuracy, recall rate and F1-score all of 80.06%
Chen et al. [40]	MVI identification	Retrospective, multi-center	Training: 270 pts/2284 WSIs Internal test: 80 pts/633 WSIs External validation: 120 pts/504 WSIs	Performance of PM: AUROC of 0.904 (internal test) and 0.871 (external validation); that of 0.875 (internal test) and 0.879 (external validation) in case of being only one WSI or biopsy samples
Zhang et al. [38]	MVI identification and quantification	Retrospective, single-center, and public database	Training: 530 pts/3866 WSIs Internal test: 223 pts/1,651 WSIs External validation (TCGA): 358 pts/WSIs	Performance of PM (MVI-AIDM): accuracy of 97.49% (internal test) and 94.25% (external validation); AUROC of 0.88 (internal test) and 0.82 (external validation); MVI positive rate of 70.85% for MVI-AIDM and 64.13% for pathologists ($P < 0.001$)

WSI whole-slide image, *pts* patients, *TCGA* the cancer genome atlas, *PM* pathomics model, *AUROC* area under the receiver operating characteristic curve, *TMA* tissue microarray, *ML* machine learning, *DL* deep learning, *PLC* primary liver cancer, *HCC* hepatocellular carcinoma, *ICC* intrahepatic cholangiocarcinoma, *cHCC-CCA* combined hepatocellular-cholangiocarcinoma, *AI* artificial intelligence, *MVI* microvascular invasion

achieving an average accuracy of 88% in distinguishing normal nuclei from tumor nuclei within HCC samples.

4.2 Histological classification

Histological subtyping of liver tumors is a labor-intensive task, often affected by significant inter-observer variability. The introduction of AI technologies is crucial to enhance the accuracy and efficiency of differential diagnoses.

A pioneering effort by Cheng et al. [31] introduced a DL model for the differential diagnosis of seven types of hepatocellular nodular lesions (HNLs). The PM, termed the hepatocellular-nodular AI model (HnAIM), achieved an AUROC of 0.935. In subgroup analyses of biopsy specimens, it demonstrated robust performance at both the patch- and slide-level, contributing to the early diagnosis of HCC and the risk stratification of

patients with HNLs. Several additional studies have focused on the differential diagnosis of typical HCC. For example, Kiani et al. [32] developed a PM to distinguish HCC from intrahepatic cholangiocarcinoma (ICC), achieving a diagnostic accuracy of 0.885 using a small training set of WSIs. Another research team designed a DL model that effectively differentiates among HCC, ICC, and colorectal liver metastases, with the AUROC value of the classifier surpassing 0.95 [33]. Furthermore, Calderaro et al. [34] conducted a DL-based phenotypic analysis in a rare cohort of patients with combined hepatocellular-cholangiocarcinoma (cHCC-CCA), demonstrating the model's capability to reclassify cHCC-CCA into either HCC or CCA.

4.3 Pathological grading

The degree of tumor differentiation is a critical factor in determining HCC prognosis. However, pathological grading (PG) remains time-consuming, resource-intensive, and labor-intensive. In this context, the role of AI-driven image classification becomes increasingly significant.

Chen et al. [35] utilized WSIs from The Cancer Genome Atlas (TCGA) database and trained a PM on the Easy DL platform to accurately classify the PG of HCC. The model performed comparably to pathologists with 5 years of experience, achieving an accuracy of 89.6% in classifying high-, intermediate-, and low-grade tumors. In a subsequent study, five AI models were evaluated for their potential in the intelligent classification scenario of histopathological images. Among them, a DL model called SENet demonstrated the highest performance with an accuracy of 95.27% in differentiating different PGs of HCC, and revealed good reliability and generalization ability [36]. Additionally, one team led by Zhuo developed an innovative approach integrating multiphoton microscopy with a DL algorithm, creating a label-free and automated system for PG classification in HCC, achieving classification accuracy exceeding 90% [37].

4.4 Identification of microvascular invasion

Microvascular invasion (MVI), a hallmark of HCC's aggressive behavior, is a well-established independent risk factor for early recurrence and poor prognosis. However, the MVI detection rate varies widely among pathologists, ranging from 7.8–74.4% [38]. Consequently, emerging pathomics technologies present a promising approach for the rapid and accurate identification of MVI.

Sun et al. [39] introduced a novel PM, named PCformer, which combines a CNN algorithm with a ViT architecture to identify MVI regions by classifying MVI boundaries in WSIs. However, the model achieved an F1 score of only 80.06% and lacked independent external validation. In the same year, Chen et al. [40] developed a PM based on WSIs from multi-center HCC cohorts using a weakly supervised DL approach for rapid MVI assessment. The model demonstrated excellent performance in both training and validation cohorts, with AUROC values of 0.904 and 0.871, respectively. More recently, another research group [38] developed an AI-based diagnostic model for MVI (MVI-AIDM) using WSIs. This model achieved a significantly higher MVI detection rate than pathologists (70.85% vs. 64.13%, $P < 0.001$), with an accuracy of 97.49% in internal validation and 94.25% in external validation, markedly improving the efficiency and accuracy of MVI diagnosis. Notably, it successfully identified subtle MVI that is often challenging for conventional pathology. Additionally, the model enabled automated quantification of

MVI areas and tumor cell counts, while providing spatial information on MVI, offering new insights for predicting postoperative recurrence in HCC.

5 Advanced applications of pathomics in HCC: prognostic prediction

Given the reliability and robustness of pathomics in the aforementioned basic applications, its expansion into advanced prognostic prediction tasks is progressing steadily and holds significant potential to directly impact clinical practice in HCC. These applications include the inference of genetic alterations and immune phenotypes, evaluation of high-risk pathological features, and prediction of survival outcomes and treatment responses [41] (Table 2).

5.1 Prediction of genetic markers and immune signatures

Although the detection of novel biomarkers, such as gene mutations or immune phenotypes, still largely depends on molecular testing technologies, advancements in AI have enabled pathomics to directly extract subtle morphological changes from WSIs for accurate prediction. This approach offers significant advantages in terms of cost-effectiveness and reproducibility.

Several studies have applied ML techniques to develop PMs that successfully predict the expression of high-risk genetic features. For example, Zhou et al. [42] demonstrated that their PM could accurately predict the expression of Enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2) and its associated prognosis in HCC. More recently, teams represented by Yan [43] and Huang [44] reported that emerging PM-based scoring systems improved the prediction of TNF receptor superfamily member 4 (TNFRSF4) and Angiopoietin-2 (ANGPT2) expression, respectively, also showing potential as prognostic biomarkers. Furthermore, bioinformatics analyses revealed that high- and low-pathomics scores (PSs) were associated with distinct HCC growth-related gene enrichment pathways, vascular endothelial growth factor (VEGF)-related gene expression, and levels of immune cell infiltration, suggesting the potential utility of PSs in guiding targeted and immunotherapies.

The tumor immune microenvironment (TIME) has garnered increasing attention in recent years for its pivotal role in the initiation, progression, and metastasis of HCC. Several research groups have utilized WSIs to develop DL-based PMs for predicting immune-related features, offering novel insights into HCC immunotherapy. Zeng et al. [45] developed a PM capable of predicting the activation of six immune gene signatures in HCC, effectively identifying those with high expression who may be more responsive to immunotherapy. Another team led by Song developed a PM that automatically estimates immune cell infiltration abundance in HCC, revealing significant differences in overall survival (OS) and immune checkpoint expression between high- and low-infiltration subgroups [46]. These innovative findings indicate that AI-powered pathomics signatures could represent a novel class of biological markers with promising potential for the clinical translation of HCC biological insights.

5.2 Prediction of high-risk pathological features

According to the latest HCC management guidelines from the EASL and AASLD [47, 48], MVI, poor PG, and satellite nodules are traditionally recognized as high-risk pathological features [49]. In recent years, new high-risk pathological phenotypes have been

Table 2 Summary of advanced applications for prognostic prediction via pathomics in hepatocellular carcinoma

Reference	Tasks	Study design	Number of patients/WSIs (tiles) and validation methods	Details and outcome measures
Zhou et al. [42]	Gene and survival prediction	Retrospective, public database	Training (TCGA): 187 pts/WSIs Internal test (TCGA): 80 pts/WSIs	Predict overexpression of EZH2 gene and OS Performance: for PM, AUROC of 0.815 (training) and 0.742 (test) for EZH2 prediction; for PS, high PS associated with worse OS
Yan et al. [43]	Gene and survival prediction	Retrospective, public database	Training and internal test (TCGA): 267 pts/WSIs (7: 3)	Predict overexpression of TNFRSF4 gene and OS Performance: for PM, AUROC of 0.787 (training) and 0.723 (test) for TNFRSF4 prediction; for PS, high PS related to unfavorable OS
Huang et al. [44]	Gene and survival prediction	Retrospective, single-center, and public database	Training (TCGA): 188 pts/WSIs Internal test (TCGA): 79 pts/WSIs External validation: 91 pts/WSIs	Predict ANGPT2 *g*ene expression and OS Performance: for PM, AUROC of 0.811 (training), 0.726 (test) and 0.719 (validation) for ANGPT2 prediction; for PS, independent risk factor of OS ($P < 0.001$)
Zeng et al. [45]	Immune prediction	Retrospective, single-center, and public database	Training and internal test (TCGA): 336 pts/349 WSIs (6:2:2) External validation: 139 pts	Predict activation of 6 immune gene signatures Performance of PM: AUROC of 0.780–0.914 (internal test) and 0.810–0.921 (external validation)
Jia et al. [46]	Immune prediction	Retrospective, single-center, and public database	Training and internal test (TCGA): 354 pts/100 WSIs (165293 tiles) Internal validation (TCGA): 7686 tiles External validation: 50 pts/1076 tiles	Evaluate immune infiltration of HCC tissue Performance of PM: Accuracy-Micro of 97.46% and Accuracy-Macro of 82.28%; AUROC over 0.95 (both internal and external validation sets)
Laurent-Bellue et al. [52]	High-risk pathological prediction	Retrospective, multi-center	Training: 60 pts/399 WSIs Internal test: 13 pts/77 WSIs Internal validation: 34 pts External validation: 29 pts/32 WSIs	Predict pejorative architectures group comprising MVI, poor pathological grading, MTM subtype and VETC pattern Performance of PM: accuracy of 0.864 at patch level and 0.823 at WSI level
Saillard et al. [23]	Survival prediction	Retrospective, single-center, and public database	Training and internal test: 194 pts/390 WSIs (7800000 tiles) External validation (TCGA): 328 pts/342 WSIs	Predict OS Performance of PM: in training set, C-index of 0.78 and 0.75 based on supervised and unsupervised DL algorithm, respectively; similarly in external validation set, C-index of 0.68 and 0.70
Shi et al. [54]	Survival prediction and histological classification	Retrospective, single-center, and public database	Training: 501 pts/1221 WSIs Internal validation: 528 pts/970 WSIs External validation (TCGA): 320 pts/WSIs	Predict OS and PFS, and histological classification Performance: for PM, accuracy of 0.982 (internal) and 0.921 (external) for histological classification; for PM-based tumor risk score, independent risk factors of OS (internal and external) and PFS (internal) ($P < 0.001$)

Table 2 (continued)

Reference	Tasks	Study design	Number of patients/WSIs (tiles) and validation methods	Details and outcome measures
Qu et al. [56]	Recurrence prediction and histological classification	Retrospective, single-center, and public database	Training: 402 pts/383 WSIs Internal validation: 174 pts/164 WSIs External validation (TCGA): 154 pts/147 WSIs	Predict RFS and histological subtyping Performance: for PM, overall accuracy of 94.17% for histological subtyping; for PS and combined score, AUROC of 0.837 and 0.857 for 1-year RFS, 0.857 and 0.852 for 3-year RFS, 0.826 and 0.845 for 5-year RFS, respectively; for PS, C-index of 0.804 (training), 0.739 (internal) and 0.708 (external)
Liu et al. [55]	Recurrence prediction	Retrospective, multi-center, and public database	Training (resection cohort): 552 pts/57,415 tiles Internal test (transplant cohort): 144 pts External validation (TCGA): 302 pts	Predict RFS Performance of PM (MobileNetV2_HCC_class): time-dependent accuracy of 0.631–0.731 (internal test) and 0.585–0.701 (external validation), time-dependent AUROC of 0.658–0.736 (internal test) and 0.641–0.712 (external validation)
Zeng et al. [60]	Therapeutic response and recurrence prediction	Retrospective, single-center, and public database	Training and internal test (TCGA): 336 pts/WSIs (6: 2: 2) External validation: 225 pts/WSIs (resection cohort), 157 pts/WSIs (biopsy cohort), 122 pts/WSIs (biopsy plus atezolizumab–bevacizumab cohort)	Predict response to atezolizumab-bevacizumab (ABRS) and PFS Performance of PM (ABRS-P): PCC of 0.60 (external resection cohort, $P < 0.0001$) and 0.53 (external biopsy cohort, $P < 0.0001$) between ABRS-P values and ABRS scores, respectively; in external atezolizumab–bevacizumab cohort, high ABRS-P value showed significantly longer median PFS (12 vs. 7 months, $P = 0.014$) than low value
Li et al. [26]	Therapeutic response and recurrence prediction	Retrospective, single-center, and public database	Training and internal test: 366 pts/724 WSIs (7: 3) External validation (TCGA): 78 pts	Predict postoperative recurrence and response to sorafenib Performance of PM: AUROC of 0.818 (internal) and 0.713 (external) for 1-year recurrence, that of 0.811 (internal) and 0.707 (external) for 2-year recurrence, respectively; effectively predict therapeutic response and survival benefit of sorafenib (both $P < 0.05$)

WSI whole-slide image; pts patients, *TCGA* the cancer genome atlas, *PM* pathomics model, *PS* pathomics score, *AUROC* area under the receiver operating characteristic curve, *C-index* concordance-index, *EZH2* Enhancer of zeste 2 polycomb repressive complex 2 subunit, *TNFRSF4* TNF receptor superfamily member 4, *ANGPT2* Angiopoietin-2, *HCC* hepatocellular carcinoma, *OS* overall survival, *PFS* progression-free survival, *RFS* relapse-free survival, *MVI* microvascular invasion, *MTM* macrotrabecular-massive, *VETC* vessels encapsulating tumor clusters, *DL* deep learning, *ABRS* atezolizumab–bevacizumab response signature, *ABRS-P* ABRS-prediction, *PCC* Pearson correlation coefficient

identified, including the vessels encapsulating tumor clusters (VETC) pattern and the macrotrabecular-massive (MTM) subtype [50, 51]. However, traditional assessment methods remain time-consuming and subjective, highlighting the potential for AI-based pathomics technologies to address these limitations.

While most previous studies on MVI prediction have focused on non-invasive radiomics methods, recent research [38–40] has explored pathomics-based predictive approaches, examining the association between pathological phenotypes and MVI. These studies provide histological insights into the invasive characteristics of the MVI microenvironment. Concurrently, the rapid identification and prognostic evaluation of new high-risk pathological features have become key areas of focus. Laurent-Bellue et al. [52] recently identified four histological phenotypes that most accurately predict

recurrence (MVI, poor PG, MTM subtype, and the VETC pattern) which were collectively categorized as non-pejorative architectures. Using a supervised DL algorithm, they developed a PM to identify and quantify these non-pejorative structures, achieving an accuracy of 0.864 at the patch level and 0.823 at the slide level. Correlation analysis suggested that identifying these structures could serve as an effective surrogate marker for MVI, assisting in recurrence risk prediction. This aligned with Chen et al. [40], who found that the MTM structure is associated with MVI positivity. More recently, Yu et al. [53] developed a DL-based radio-pathomics model using MRI and WSIs to predict the VETC pattern in HCC. Their nomogram scoring system enabled risk stratification and successfully assessed early recurrence and progression-free survival (PFS). Consequently, the quantitative characteristics of pathomics show promise as alternative biological markers for high-risk pathological phenotypes, with significant potential for clinical translation.

5.3 Prediction of survival and recurrence

Accurate survival prognosis prediction and risk stratification are crucial for treatment decision-making in HCC. A high recurrence risk supports the need for adjuvant therapy post-surgery, while a high mortality risk may require more aggressive treatment strategies. Notably, pathomics can directly extract both visible and sub-visible features from WSIs, offering a comprehensive reflection of survival outcomes.

Saillard et al. [23] developed two PMs based on DL architecture, achieving concordance indices of 0.78 and 0.75 for predicting OS after HCC surgery, both surpassing traditional variable-based models in terms of discriminatory ability. They identified high-risk pathological features, including the MTM subtype, vascular space, cellular atypia, and nuclear pleomorphism, while immune cell infiltration was identified as a low-risk feature. Following this, Shi et al. [54] introduced a tumor risk score (TRS) system, based on a weakly supervised DL framework, to assess post-operative survival prognosis in HCC. The predictive power of the TRS system outperformed that of clinical staging systems. Pathological review confirmed similar findings, identifying features such as sinusoidal capillarization, prominent nucleoli and nuclear membrane, nuclear/cytoplasm ratio, and inflammatory cell infiltration as key potential predictors for TRS. The consistency of these results suggests that tissue architecture, cellular morphology, and immune response are key pathological features influencing survival prognosis. Currently, combined anti-angiogenic therapy and immunotherapy represent first-line treatments for advanced HCC [5]. Therefore, TRS holds potential as a biomarker for predicting the efficacy of combined therapies.

Subsequent studies primarily employed DL algorithms to develop PMs for predicting recurrence risk following HCC resection or transplantation, enabling effective risk stratification [25, 55–57]. For example, Liu et al. [55] found that tumor areas with the highest recurrence risk exhibited pathological features such as stroma, high cellular atypia, deeply stained nuclei, and a lack of immune cell infiltration, consistent with the findings of afore-mentioned studies [23, 54]. Additionally, Qu et al. [56] observed that patients with high-risk recurrence scores had a higher proportion of poor PG and MVI, with immunohistochemical staining revealing their correlations with local immune cell infiltration. These findings underscore the importance of pathomics and the TIME in HCC prognosis assessment.

Table 3 Summary of pathomics-based multi-modal studies of prognostic prediction in hepatocellular carcinoma

Reference	Tasks	Study design	Number of patients/WSIs (tiles) and validation methods	Details and outcome measures
Yu et al. [53]	High-risk pathological and survival prediction	Retrospective, multi-center	Training: 317 pts Internal test: 137 pts External validation: 124 pts	Predict MTM subtype, early recurrence and PFS Performance: for PM, AUROC of 0.79 in predicting MTM subtype; for combined nomogram model (radio-pathomics), C-index of 0.60 (internal) and 0.67 (external) for early recurrence, AUROC of 0.81 (internal) and 0.78 (external) for 3-year PFS, respectively
Qu et al. [25]	Recurrence prediction	Retrospective, single-center	Training: 256 pts/302 WSIs Internal validation: 124 pts/193 WSIs	Predict TTR and RFS after transplantation Performance: for PS, C-index of 0.827 (training) and 0.794 (validation), AUROC of 0.861 and 0.795 for TTR and RFS, respectively; for combined score (PS plus clinical features/clinico-pathomics), C-index of 0.849 (entire cohort)
Schmauch et al. 2024 [57]	Recurrence and survival prediction	Retrospective, multi-center	Training and internal test: 194 pts External validation: 169 pts/385 WSIs (resection cohort), 300 pts/300 WSIs (transplant cohort)	Predict RFS and DSS Performance of combined model (clinico-pathomics) proved superior to PM: C-index of 0.64 and 0.77 for RFS and DSS in resection cohort, respectively, and that of 0.76 and 0.77 in transplant cohort
Feng et al. [58]	Survival prediction	Retrospective, single-center	Training and internal validation: 126 pts (7: 3)	Predict OS Performance of combined nomogram model (radio-pathomics) with most robustness: C-index of 0.840 (training) and 0.875 (validation); time-dependent AUROC of 0.899 (training) and 0.875 (validation)
Xie et al. [59]	Recurrence prediction	Retrospective, single-center	Training and internal test: 107 pts	Predict early recurrence Performance of multi-modal model (clinico-radio-pathomics) outperforming bi-modal models: AUROC of 0.743–0.863; and for best algorithm SVM, AUROC of 0.863, accuracy of 0.784, sensitivity of 0.731 and specificity of 0.826

WSI whole-slide image; pts patients, *PM* pathomics model, *PS* pathomics score, *AUROC* area under the receiver operating characteristic curve, *C-index* concordance-index, *MTM* macrotrabecular-massive, *OS* overall survival, *PFS* progression-free survival, *TTR* time to recurrence, *RFS* relapse-free survival, *DSS* disease-specific survival, *SVM* support vector machine

The integration of multi-modal data offers more comprehensive, multi-level, and cross-scale biological information, enhancing model prediction accuracy, which represents the current trend in HCC prediction research (Table 3). For instance, Schmauch et al. [57] demonstrated that a bi-modal model combining clinical and pathomics data significantly improved the performance of recurrence and survival prediction in HCC. Similarly, the dual-modal nomogram developed by Feng et al. [58], which fused MRI radiomics and pathomics features, demonstrated the most stable performance in predicting OS, achieving the highest net benefit. Furthermore, another research team developed a multi-modal model integrating clinical variables, radiomics, and pathomics features, validating its superior predictive performance over the bi-modal model [59]. Thus, the integration of additional data dimensions, such as genomics, transcriptomics, and proteomics, into AI-based predictive frameworks represents a promising direction.

5.4 Prediction of therapeutic response

Treatment options for unresectable HCC have become increasingly diversified, encompassing local-regional approaches and systemic therapies, particularly targeted therapy and immunotherapy [5]. However, these therapies are only effective for a subset of patients and may cause adverse effects in non-responders. Pathomics facilitates the identification of structural characteristics and transformation features in HCC tissues, enabling the exploration of predictive biomarkers for positive therapeutic responses. This approach aids in the precise identification of potential responders, thereby minimizing adverse effects in non-responders.

Zeng et al. [60] developed the PM, termed ABRS-P, which accurately evaluates the expression of the atezolizumab–bevacizumab response signature (ABRS) in HCC, further validating its potential as a novel biomarker for predicting PFS. Subsequent spatial transcriptomics analysis demonstrated that multiple immune-related factors were upregulated in patients with high ABRS-P scores. ABRS-P enhances immunotherapy precision in a simple and cost-effective manner, supporting personalized management of HCC. Li et al. [26] recently employed denoised recurrence labels at various thresholds to train the prognostic PM and named it CNN-SASM, which accurately stratified the risk of patients with HCC undergoing sorafenib therapy and effectively predicted survival benefits.

Currently, two primary methods are used to evaluate treatment responses based on WSIs. One involves utilizing pathomics to predict known treatment response-related molecular features, such as the ABRS [60]. Therefore, AI-assisted exploration of immune gene features associated with therapeutic responses in the TIME holds significant clinical translational value [45, 46]. The second method predicts treatment responses directly from WSIs. Although the latter is the preferred approach, such "end-to-end" workflows require training on large-scale cohorts with known treatment responses, as demonstrated in the sorafenib response study [26]. Due to limited access to such data, studies of this nature remain relatively rare, and the full potential of pathomics in predicting therapeutic responses has yet to be fully realized.

6 The emergence of pathological foundation models

Recent years have witnessed rapid advancements in large models, particularly in natural language processing and image/video analysis, often referred to as "foundation models". These neural network architectures are pre-trained on extensive general data and can be fine-tuned for various downstream tasks. With the increasing demand in clinical diagnosis and research, medical foundation models have emerged as a prominent topic in AI research. By pre-training on large-scale WSI data, pathological foundation models (PFMs) can accurately and efficiently assist in pathological diagnosis, feature analysis, and treatment decision-making. This not only reshapes the research landscape in pathology but also holds transformative potential for clinical translation [10]. However, the diversity of tasks and data heterogeneity in pathology renders a single PFM insufficient for all applications. Consequently, their primary value lies in accelerating the development of specialized vertical models for specific downstream tasks. Recent studies published in *J Clin Oncol* exemplify this trend, such as a specialized foundation model (abbreviated as DINOPath) for predicting survival prognosis and treatment benefits in patients with gastrointestinal cancers [61], and a multi-modal model (termed ArteraAI

Prostate Test) for accurately predicting long-term benefits of androgen deprivation therapy in high-risk patients with prostate cancer [62]. Similarly, there is an urgent need for vertical, task-specific foundation models in HCC.

The development of PFMs advanced significantly with the emergence of self-supervised learning. Numerous research outcomes have been published in leading journals such as *Nature* and *Nat Med*, including models like PLIP, Virchow, CONCH, UNI, Prov GigaPath, PRISM, PathChat, CHIEF, PathoDuet, and PathOrchestra [63]. Recently, models such as TITAN [64], mSTAR [65], BEPH [66], and MUSK [67] have been developed for additional downstream tasks. Interactive visual-language intelligent assistant models like PathChat [68] and PathAsst [69], known as "real-time pathology assistants," are also gaining prominence. PFMs are now evolving from "tool-based" to "cognitive-based" systems, with the future objective of creating universal generalist models (also described as AI agents) that combines both general and specialized capabilities. These AI agents, integrating various specially trained vertical models, will be capable of understanding user instructions, making autonomous decisions, and selecting the appropriate tools to perform specific tasks, ultimately automating the entire pathological pipeline (Fig. 2).

7 Challenges and future perspectives of pathomics in HCC

Despite the promising potential of AI-empowered pathomics in HCC diagnosis and prognosis, several challenges still persist across its entire workflow from data collection and model development to clinical translation. The following six sections will provide a critical analysis and future outlook, identify research gaps, and propose corresponding solutions.

7.1 Data heterogeneity and scarcity

High-quality, multi-center WSI datasets are essential for developing PMs with robust generalizability. However, significant variability in scanner hardware, staining protocols, image preprocessing pipelines, and annotation standards across institutions leads to severe domain shifts when models are applied across centers [70]. Although public repositories such as TCGA and the Clinical Proteomic Tumor Analysis Consortium (CPTAC) offer large collections of HCC WSIs, they underrepresent rare histological subtypes and lack standardized metadata, limiting their effectiveness in training on diverse samples [22, 70]. Weakly supervised and unsupervised annotation methods can reduce the costs of manual labeling but are typically validated only on small, single-center cohorts, while crowdsourced annotations introduce inter-observer variability that undermines data consistency [54, 71, 72].

To address these gaps, coordinated efforts are needed in annotation standards, collaborative training, and data sharing. First, an international WSI annotation consortium should be established to harmonize annotation workflows, implement multi-tiered quality audits, and define a unified metadata schema. Second, unified guidelines for staining, scanning resolution, color normalization, and artifact filtering must be developed and promoted, leveraging automated quality control tools such as GrandQC [73]to minimize source-level heterogeneity [74]. Third, privacy-preserving distributed training frameworks like federated learning and swarm learning should be adopted to enable collaborative model optimization without exchanging raw data [70]. Finally, collaboration between researchers, healthcare institutions, and industry partners is essential to

develop and publicly share multi-modal external validation cohorts integrating clinical, molecular, and imaging data. This will provide benchmarks for fair algorithm evaluation and ultimately enhance model generalization and clinical applicability.

7.2 Model transparency and interpretability

Although DL models have achieved continual breakthroughs in pathomics for HCC, their "black-box" nature remains a major barrier to clinical adoption. The lack of interpretable decision-making processes erodes the trust of clinicians and patients in these tools [75]. Common visualization techniques, such as Gradient-weighted Class Activation Mapping (Grad-CAM), SHapley Additive exPlanations (SHAP), and Concept Activation Vectors (CAVs), can highlight regions or cellular structures on which the model focuses. However, these techniques fall short in quantitatively linking these regions to underlying molecular pathways, gene expression, or patient outcomes, limiting their ability to fully restore confidence in AI-driven decisions [54, 60, 72, 75]. Moreover, relying solely on statistical metrics like AUROC and accuracy fails to capture a model's true clinical utility in real-world decision-making contexts.

To achieve genuinely useful explainability, several key advancements are required. First, we need to develop a multi-scale, quantitative interpretability framework that not only visualizes tumor regions of interest but also integrates molecular pathway and gene expression data. This would help quantify the relationship between image features and biological mechanisms [54]. Concurrently, real-world cohort studies should incorporate clinical utility metrics such as DCA, clinical impact curves (CIC), and net benefit, to assess how different explainability methods contribute to decision support [71]. Notably, a human-in-the-loop design is also indispensable; by iteratively incorporating pathologists' expertise into model refinement and visualization workflows, we can significantly enhance interpretability and trust [76]. Finally, as regulatory bodies like the Food and Drug Administration (FDA) and European Medicines Agency (EMA) increasingly mandate explainability for AI systems, it is imperative to establish a unified reporting standard. Developers need to provide specific types of interpretability evidence to meet clinical trial and approval requirements [77]. By implementing these measures, PMs of HCC can maintain high performance while becoming truly interpretable, verifiable, and deployable.

7.3 Ethical issues and regulatory frameworks

The clinical implementation of AI-driven pathomics presents significant ethical and regulatory challenges that require thorough consideration [24]. Patient privacy and data security are of paramount concerns, with stringent regulations such as Health Insurance Portability and Accountability Act (HIPAA) in the United States (US) and General Data Protection Regulation (GDPR) in the European Union (EU) mandating robust safeguards, including data encryption, access control, and de-identification [78]. However, obtaining meaningful informed consent is complicated by the often opaque, "black-box" nature of AI algorithms, which limits patients' understanding of how their data is used and how AI contributes to clinical decisions. Therefore, enhancing transparency and explainability in AI models is essential to uphold patient autonomy and build trust. Furthermore, ethical review is crucial in validating the legitimacy of medical research, and these issues must be addressed during the research design phase [79].

Algorithmic bias is another critical concern that can lead to inequitable healthcare outcomes. Biases may arise from unrepresentative training datasets or inherent flaws in AI model design, potentially disadvantaging certain patient groups [80]. To mitigate these risks, it is necessary to employ diverse and inclusive datasets, continuously monitor AI performance across different populations, implement fairness metrics during model evaluation, and adhere to the bias assessment frameworks outlined by TRIPOD-AI and PROBAST-AI during reporting [77]. These measures are vital to ensure that AI applications do not inadvertently perpetuate or exacerbate existing healthcare disparities.

From a regulatory perspective, agencies such as the US FDA, EMA in EU, and National Medical Products Administration (NMPA) in China have developed adaptive, risk-based frameworks to evaluate AI software as Software as a Medical Device (SaMD). These frameworks cover pre-market validation, risk classification, and post-market surveillance to ensure safety and efficacy [78]. However, regulatory heterogeneity across jurisdictions impedes the global deployment of AI tools, and there is still no comprehensive lifecycle management guidance addressing issues such as "post-approval model re-certification" and "continuous online monitoring". In response, the SPIRIT-AI Extension [81] and CONSORT-AI Extension [82] have supplemented clinical trial design and reporting guidelines for AI interventions, while the FUTURE-AI consensus has further underscored best practices for developing trustworthy and deployable AI [83]. Moving forward, it is imperative to establish full-process regulatory policies that encompass model development, clinical validation, deployment, and maintenance. These policies must clearly delineate the responsibilities and authorities of AI developers, healthcare providers, and regulatory bodies, striking a balance between innovation and patient safety.

7.4 Industry standards and platform construction

Current guidelines such as TRIPOD-AI and PROBAST-AI provide robust frameworks for reporting and bias assessment in AI predictive model research [77]; however, they do not address the multi-modal quality control of WSIs or the need for continuous online monitoring. To support end-to-end deployment, digital pathology platforms must implement a complete ML operations (ML-Ops) workflow encompassing data acquisition and preprocessing, model development and deployment, clinical integration, and ongoing maintenance [22, 84]. At the data level, it is essential to adopt unified WSI quality standards and leverage automated tools such as GrandQC [73] to verify scanning resolution, color consistency, and artifact presence [74, 85]. At the algorithm level, platforms should support continuous learning and real-time performance surveillance to detect data drift and automatically initiate model retraining or revalidation, thereby ensuring sustained safety and efficacy across evolving data environments [84].

Commercial implementation further requires clear policies on cost allocation and intellectual property. Platforms ought to include compliance-management modules that delineate each stakeholder's rights and responsibilities regarding data use, model updates, and revenue sharing. Furthermore, a dedicated certification standard for pathomics-based SaMD should be developed. This standard must cover software-lifecycle management, cybersecurity safeguards, user-access controls, and audit-logging requirements, in order to satisfy the specific evaluation criteria of regulators such as the FDA, EMA, and NMPA [78]. By building end-to-end, scalable, and compliant AI

platforms for digital pathology, institutions can more readily share pathomics innovations and accelerate their translation into real-world clinical practice.

7.5 Roadmap for external validation and clinical translation

Recognizing the persistent challenges in AI-driven pathomics for HCC, and building upon the research gaps identified above, we propose a six-step roadmap to address these unmet needs. First, an international WSI-annotation and data-sharing consortium should be established. This consortium will harmonize image-quality criteria, metadata standards, and annotation protocols. By adopting privacy-preserving learning (e.g., federated learning), it can facilitate multi-center model development without exchanging raw data. Second, standardized preprocessing and feature-extraction pipelines must be defined. These pipelines should cover tissue preparation, slide scanning parameters, color normalization, artifact filtering, and pathomics-feature computation, ensuring that extracted features are reproducible and comparable across centers. Third, a unified multi-modal interpretability framework should be developed. This framework will quantitatively link histopathological features with molecular pathways, gene-expression profiles, and complementary imaging modalities to produce transparent, explainable AI outputs that clinicians can trust.

Fourth, well-designed prospective, multi-center validation must be performed. Embedding models into clinical trials across diverse institutions will enable systematic collection of multi-dimensional endpoints including diagnostic accuracy, cost-effectiveness, and human-AI collaboration efficiency, thereby demonstrating real-world clinical utility. Fifth, robust ML-Ops infrastructure and SaMD regulatory pathways should be implemented. These end-to-end pipelines should cover data ingestion, model training, deployment, continuous performance monitoring, and automated retraining or revalidation, alongside compliance modules for cybersecurity, user-access control, and commercialization requirements. Finally, deployment-phase monitoring with formal recertification triggers is essential. Real-time surveillance to detect data drift and performance degradation should automatically initiate model updates or retraining, thereby ensuring ongoing compliance with SaMD lifecycle regulations.

By integrating these elements into a cohesive strategy that spans consortium formation, pipeline standardization, interpretability, prospective validation, ML-Ops compliance and continuous monitoring, AI-driven pathomics for HCC can accelerate translation from proof of concept to routine practice.

8 Conclusion

Looking ahead, AI-driven pathomics holds the promise to revolutionize HCC diagnosis, treatment planning, and prognostic evaluation. However, key barriers including data heterogeneity, the absence of standardized workflows, regulatory and ethical uncertainties, and the need for robust clinical validation remain to be addressed. By improving dataset quality and sharing, unifying metadata and annotation standards, establishing industry-wide guidelines and compliant platform infrastructures, conducting large-scale prospective multimodal studies, and continuously refining PFMs, we can accelerate the translation of pathomics into routine clinical practice. Ultimately, these concerted efforts will enable precise, personalized management of HCC and improve patient outcomes.

Acknowledgements
We are deeply indebted to all those who provided invaluable support and contributions to this work.

Author contributions
W Ding, JX Zhang and ZC Jin drafted the manuscript, collected the literature data and elaborated the figures and tables; HJ Hua, QQ Zu, and SD Yang conducted the literature collection, checking and revising the paper; HF Zhou, WD Wang, S Liu and HB Shi designed the review, revised the draft and provided the financial support. All authors have read and approved the final manuscript.

Funding
This work is supported by the Jiangsu Provincial Innovative and Entrepreneurial Doctor Program (303073543ER21, to Hai-Feng Zhou), the Young Scholars Fostering Fund of the First Affiliated Hospital of Nanjing Medical University (PY2023056, to Hai-Feng Zhou), the Natural Science Foundation of Jiangsu Province (Youth Fund Project) (BK20241121, to Hai-Feng Zhou), the Young Scholars Fostering Fund of the First Affiliated Hospital of Nanjing Medical University (PY2023051, to Jin-Xing Zhang), the Technological Achievements and Appropriate Technology Promotion Project of the Wuxi Health Commission (T202212, to Wei-Dong Wang), and the Jiangsu Province Capability Improvement Project through Science, Technology and Education (JSDW202243, to Hai-Bin Shi).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors consent to the publication of the manuscript.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Interventional Radiology, The Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi People's Hospital, Wuxi Medical Center, Nanjing Medical University, No. 299 Qingyang Road, Wuxi 214023, China
[2]Department of Interventional Radiology, The First Affiliated Hospital with Nanjing Medical University, No. 300 Guangzhou Road, Nanjing 210029, China
[3]Center of Interventional Radiology and Vascular Surgery, Department of Radiology, Zhongda Hospital, Medical School, Southeast University, No. 87 Dingjiaqiao Road, Nanjing 210009, China
[4]Department of Pathology, The First Affiliated Hospital with Nanjing Medical University, No. 300 Guangzhou Road, Nanjing 210029, China
[5]Department of Pathology, The Affiliated Wuxi People's Hospital of Nanjing Medical University, No. 299 Qingyang Road, Wuxi 214023, China

Received: 12 June 2025 / Accepted: 21 July 2025
Published online: 28 July 2025

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA-CANCER J CLIN. 2024;74(3):229–63. https://doi.org/10.3322/caac.21834.
2. Rumgay H, Arnold M, Ferlay J, Lesi O, Cabasag CJ, Vignat J, et al. Global burden of primary liver cancer in 2020 and predictions to 2040. J HEPATOL. 2022;77(6):1598–606. https://doi.org/10.1016/j.jhep.2022.08.021.
3. Llovet JM, Pinyol R, Yarchoan M, Singal AG, Marron TU, Schwartz M, et al. Adjuvant and neoadjuvant immunotherapies in hepatocellular carcinoma. NAT REV CLIN ONCOL. 2024;21(4):294–311. https://doi.org/10.1038/s41571-024-00868-0.
4. Wang Q, Li J, Zhang Z, Li Y, Liang Y, Chen X, et al. The effectiveness and safety of therapies for hepatocellular carcinoma with tumor thrombus in the hepatic vein, inferior Vena cave and/or right atrium: a systematic review and single-arm meta-analysis. EXPERT REV ANTICANC. 2025;25(5):561–70. https://doi.org/10.1080/14737140.2025.2489651.
5. Zhong B, Fan W, Guan JJ, Peng Z, Jia Z, Jin H, et al. Combination locoregional and systemic therapies in hepatocellular carcinoma. LANCET GASTROENTEROL. 2025;10(4):369–86. https://doi.org/10.1016/S2468-1253(24)00247-4.
6. Li J, Liang Y, Wang Q, Li Y, Chen X, Luo W, et al. Tumor-associated lymphatic vessel density is a postoperative prognostic biomarker of hepatobiliary cancers: a systematic review and meta-analysis. FRONT IMMUNOL. 2025;15:1519999. https://doi.org/10.3389/fimmu.2024.1519999.
7. Chan Y, Zhang C, Wu J, Lu P, Xu L, Yuan H, et al. Biomarkers for diagnosis and therapeutic options in hepatocellular carcinoma. MOL CANCER. 2024;23(1):189. https://doi.org/10.1186/s12943-024-02101-z.
8. Li Y, Wu S, Wu Y, Zhang W, Wang Y, Li Y, et al. Portal venous and hepatic arterial coefficients predict Post-Hepatectomy overall and Recurrence-Free survival in patients with hepatocellular carcinoma: A retrospective study. J HEPATOCELL CARCINO. 2024;11:1389–402. https://doi.org/10.2147/JHC.S462168.
9. Aggarwal A, Bharadwaj S, Corredor G, Pathak T, Badve S, Madabhushi A. Artificial intelligence in digital pathology - time for a reality check. NAT REV CLIN ONCOL. 2025. https://doi.org/10.1038/s41571-025-00991-6.

10. Ma Y, Jamdade S, Konduri L, Sailem H. AI in histopathology explorer for comprehensive analysis of the evolving AI landscape in histopathology. NPJ DIGIT MED. 2025;8(1):156. https://doi.org/10.1038/s41746-025-01524-2.
11. Wagner VM. Is histopathology deep learning artificial intelligence the future of precision oncology?? J CLIN ONCOL. 2024;42(30):3513–6. https://doi.org/10.1200/JCO-24-01271.
12. Gupta R, Kurc T, Sharma A, Almeida JS, Saltz J. The emergence of pathomics. CURR PATHOBIOL REP. 2019;7(3):73–84. https://doi.org/10.1007/s40139-019-00200-x.
13. Ochi M, Komura D, Ishikawa S. Pathology foundation models. JMA J. 2025;8(1):121–30. https://doi.org/10.31662/jmaj.2024-0206.
14. Song AH, Jaume G, Williamson DFK, Lu MY, Vaidya A, Miller TR, et al. Artificial intelligence for digital and computational pathology. Nat Reviews Bioeng. 2023. https://doi.org/10.1038/s44222-023-00096-8.
15. Bergstrom EN, Abbasi A, Diaz-Gay M, Galland L, Ladoire S, Lippman SM, et al. Deep learning artificial intelligence predicts homologous recombination deficiency and platinum response from histologic slides. J CLIN ONCOL. 2024;42(30):3550–60. https://doi.org/10.1200/JCO.23.02641.
16. Bjartell A, Krzyzanowska A, Liu VYT, Tierney M, Royce TJ, Sjostrom M, et al. Validation of a digital pathology-based multimodal artificial intelligence biomarker in a prospective, real-world prostate cancer cohort treated with prostatectomy. CLIN CANCER RES. 2025. https://doi.org/10.1158/1078-0432.CCR-24-3656.
17. Boehm KM, El Nahhas OSM, Marra A, Waters M, Jee J, Braunstein L, et al. Multimodal histopathologic models stratify hormone receptor-positive early breast cancer. NAT COMMUN. 2025;16(1):2106. https://doi.org/10.1038/s41467-025-57283-x.
18. Chen T, Wen J, Shen X, Shen J, Deng J, Zhao M, et al. Whole slide image based deep learning refines prognosis and therapeutic response evaluation in lung adenocarcinoma. NPJ DIGIT MED. 2025;8(1):69. https://doi.org/10.1038/s41746-025-01470-z.
19. Jasti J, Zhong H, Panwar V, Jarmale V, Miyata J, Carrillo D, et al. Histopathology based AI model predicts anti-angiogenic therapy response in renal cancer clinical trial. NAT COMMUN. 2025;16(1):2610. https://doi.org/10.1038/s41467-025-57717-6.
20. Zheng J, Wang T, Wang H, Yan B, Lai J, Qiu K, et al. Use of a pathomics nomogram to predict postoperative liver metastasis in patients with stage III colorectal Cancer. ANN SURG ONCOL. 2025;32(2):771–82. https://doi.org/10.1245/s10434-024-16519-8.
21. Hosseini MS, Bejnordi BE, Trinh VQ, Chan L, Hasan D, Li X, et al. Computational pathology: A survey review and the way forward. J Pathol Inf. 2024;15:100357. https://doi.org/10.1016/j.jpi.2023.100357.
22. Komura D, Ochi M, Ishikawa S. Machine learning methods for histopathological image analysis: updates in 2024. COMPUT STRUCT BIOTEC. 2025;27:383–400. https://doi.org/10.1016/j.csbj.2024.12.033.
23. Saillard C, Schmauch B, Laifa O, Moarii M, Toldo S, Zaslavskiy M, et al. Predicting survival after hepatocellular carcinoma resection using deep learning on histological slides. Hepatology. 2020;72(6):2000–13. https://doi.org/10.1002/hep.31207.
24. Marra A, Morganti S, Pareja F, Campanella G, Bibeau F, Fuchs T, et al. Artificial intelligence entering the pathology arena in oncology: current applications and future perspectives. ANN ONCOL. 2025;36(7):712–25. https://doi.org/10.1016/j.annonc.2025.03.006.
25. Qu W, Tian M, Lu H, Zhou Y, Liu W, Tang Z, et al. Development of a deep pathomics score for predicting hepatocellular carcinoma recurrence after liver transplantation. HEPATOL INT. 2023;17(4):927–41. https://doi.org/10.1007/s12072-023-10511-2.
26. Li Y, Xiong J, Hu Z, Chang Q, Ren N, Zhong F, et al. Denoised recurrence label-based deep learning for prediction of postoperative recurrence risk and Sorafenib response in HCC. BMC MED. 2025;23(1):162. https://doi.org/10.1186/s12916-025-03977-4.
27. Pantanowitz L, Pearce T, Abukhiran I, Hanna M, Wheeler S, Soong TR, et al. Nongenerative artificial intelligence in medicine: advancements and applications in supervised and unsupervised machine learning. Mod PATHOL. 2025;38(3):100680. https://doi.org/10.1016/j.modpat.2024.100680.
28. Liao H, Long Y, Han R, Wang W, Xu L, Liao M, et al. Deep learning-based classification and mutation prediction from histopathological images of hepatocellular carcinoma. CLIN TRANSL MED. 2020;10(2):e102. https://doi.org/10.1002/ctm2.102.
29. Sun C, Xu A, Liu D, Xiong Z, Zhao F, Ding W. Deep Learning-Based classification of liver Cancer histopathology images using only global labels. IEEE J BIOMED HEALTH. 2020;24(6):1643–51. https://doi.org/10.1109/JBHI.2019.2949837.
30. Troiano M, Grignaffini F, De Stefanis C, D'Oria V, Bianchi M, Mangini F, et al. Comparison between two artificial intelligence models to discriminate cancerous cell nuclei based on confocal fluorescence imaging in hepatocellular carcinoma. DIGEST LIVER DIS. 2024. https://doi.org/10.1016/j.dld.2024.11.026.
31. Cheng N, Ren Y, Zhou J, Zhang Y, Wang D, Zhang X, et al. Deep Learning-Based classification of hepatocellular nodular lesions on Whole-Slide histopathologic images. Gastroenterology. 2022;162(7):1948–61. https://doi.org/10.1053/j.gastro.2022.02.025.
32. Kiani A, Uyumazturk B, Rajpurkar P, Wang A, Gao R, Jones E, et al. Impact of a deep learning assistant on the histopathologic classification of liver cancer. NPJ DIGIT MED. 2020;3:23. https://doi.org/10.1038/s41746-020-0232-8.
33. Jang H, Go J, Kim Y, Lee SH. Deep learning for the pathologic diagnosis of hepatocellular carcinoma, cholangiocarcinoma, and metastatic colorectal cancer. CANCERS. 2023. https://doi.org/10.3390/cancers15225389.
34. Calderaro J, Ghaffari Laleh N, Zeng Q, Maille P, Favre L, Pujals A, et al. Deep learning-based phenotyping reclassifies combined hepatocellular-cholangiocarcinoma. NAT COMMUN. 2023;14(1):8290. https://doi.org/10.1038/s41467-023-43749-3.
35. Chen M, Zhang B, Topatana W, Cao J, Zhu H, Juengpanich S, et al. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. NPJ PRECIS ONCOL. 2020;4:14. https://doi.org/10.1038/s41698-020-0120-3.
36. Chen C, Chen C, Ma M, Ma X, Lv X, Dong X, et al. Classification of multi-differentiated liver cancer pathological images based on deep learning attention mechanism. BMC MED Inf DECIS. 2022;22(1):176. https://doi.org/10.1186/s12911-022-01919-1.
37. Lin H, Wei C, Wang G, Chen H, Lin L, Ni M, et al. Automated classification of hepatocellular carcinoma differentiation using multiphoton microscopy and deep learning. J BIOPHOTONICS. 2019;12(7):e201800435. https://doi.org/10.1002/jbio.201800435.

38. Zhang X, Yu X, Liang W, Zhang Z, Zhang S, Xu L, et al. Deep learning-based accurate diagnosis and quantitative evaluation of microvascular invasion in hepatocellular carcinoma on whole-slide histopathology images. CANCER MED-US. 2024;13(5):e7104. https://doi.org/10.1002/cam4.7104.
39. Sun L, Sun Z, Wang C, Cheng S, Wang K, Huang M. PCformer: an MVI recognition method via classification of the MVI boundary according to histopathological images of liver cancer. J OPT SOC AM A. 2022;39(9):1673–81. https://doi.org/10.1364/JOSAA.463439.
40. Chen Q, Xiao H, Gu Y, Weng Z, Wei L, Li B, et al. Deep learning for evaluation of microvascular invasion in hepatocellular carcinoma from tumor areas of histology images. HEPATOL INT. 2022;16(3):590–602. https://doi.org/10.1007/s12072-022-10323-w.
41. Echle A, Rindtorff NT, Brinker TJ, Luedde T, Pearson AT, Kather JN. Deep learning in cancer pathology: a new generation of clinical biomarkers. BRIT J CANCER. 2021;124(4):686–96. https://doi.org/10.1038/s41416-020-01122-x.
42. Zhou X, Man M, Cui M, Zhou X, Hu Y, Liu Q, et al. Relationship between EZH2 expression and prognosis of patients with hepatocellular carcinoma using a pathomics predictive model. HELIYON. 2024;10(20):e38562. https://doi.org/10.1016/j.heliyon.2024.e38562.
43. Yan Z, Li X, Li Z, Liu S, Chang H. Prognostic significance of TNFRSF4 expression and development of a pathomics model to predict expression in hepatocellular carcinoma. HELIYON. 2024;10(11):e31882. https://doi.org/10.1016/j.heliyon.2024.e31882.
44. Huang X, Zheng S, Li S, Huang Y, Zhang W, Liu F, et al. Machine Learning-Based pathomics model predicts Angiopoietin-2 expression and prognosis in hepatocellular carcinoma. AM J PATHOL. 2025;195(3):561–74. https://doi.org/10.1016/j.ajpath.2024.12.005.
45. Zeng Q, Klein C, Caruso S, Maille P, Laleh NG, Sommacale D, et al. Artificial intelligence predicts immune and inflammatory gene signatures directly from hepatocellular carcinoma histology. J HEPATOL. 2022;77(1):116–27. https://doi.org/10.1016/j.jhep.2022.01.018.
46. Jia W, Shi W, Yao Q, Mao Z, Chen C, Fan A, et al. Identifying immune infiltration by deep learning to assess the prognosis of patients with hepatocellular carcinoma. J CANCER RES CLIN. 2023;149(14):12621–35. https://doi.org/10.1007/s00432-023-05097-z.
47. EASL Clinical Practice Guidelines on the management of hepatocellular carcinoma. J HEPATOL. 2025;82(2):315–74. https://doi.org/10.1016/j.jhep.2024.08.028.
48. Singal AG, Llovet JM, Yarchoan M, Mehta N, Heimbach JK, Dawson LA, et al. AASLD practice guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. Hepatology. 2023;78(6):1922–65. https://doi.org/10.1097/HEP.0000000000000466.
49. Zhang K, He K, Zhang L, Li W, Xie S, Cui Y, et al. Gadoxetic Acid-enhanced MRI scoring model to predict pathologic features of hepatocellular carcinoma. Radiology. 2025;314(2):e233229. https://doi.org/10.1148/radiol.233229.
50. Feng Z, Li H, Liu Q, Duan J, Zhou W, Yu X, et al. CT radiomics to predict Macrotrabecular-Massive subtype and immune status in hepatocellular carcinoma. Radiology. 2023;307(1):e221291. https://doi.org/10.1148/radiol.221291.
51. Pan J, Huang H, Zhang S, Zhu Y, Zhang Y, Wang M, et al. Intraindividual comparison of CT and MRI for predicting vessels encapsulating tumor clusters in hepatocellular carcinoma. EUR RADIOL. 2025;35(1):61–72. https://doi.org/10.1007/s00330-024-10944-9.
52. Laurent-Bellue A, Sadraoui A, Claude L, Calderaro J, Posseme K, Vibert E, et al. Deep learning classification and quantification of pejorative and nonpejorative architectures in resected hepatocellular carcinoma from digital histopathologic images. AM J PATHOL. 2024;194(9):1684–700. https://doi.org/10.1016/j.ajpath.2024.05.007.
53. Yu Y, Cao L, Shen B, Du M, Gu W, Gu C, et al. Deep learning radiopathomics models based on Contrast-enhanced MRI and pathologic imaging for predicting vessels encapsulating tumor clusters and prognosis in hepatocellular carcinoma. Radiol Imaging Cancer. 2025;7(2):e240213. https://doi.org/10.1148/rycan.240213.
54. Shi J, Wang X, Ding G, Dong Z, Han J, Guan Z, et al. Exploring prognostic indicators in the pathological images of hepatocellular carcinoma based on deep learning. Gut. 2021;70(5):951–61. https://doi.org/10.1136/gutjnl-2020-320930.
55. Liu Z, Liu Y, Zhang W, Hong Y, Meng J, Wang J, et al. Deep learning for prediction of hepatocellular carcinoma recurrence after resection or liver transplantation: a discovery and validation study. HEPATOL INT. 2022;16(3):577–89. https://doi.org/10.1007/s12072-022-10321-y.
56. Qu W, Tian M, Qiu J, Guo Y, Tao C, Liu W, et al. Exploring pathological signatures for predicting the recurrence of early-stage hepatocellular carcinoma based on deep learning. FRONT ONCOL. 2022;12:968202. https://doi.org/10.3389/fonc.2022.968202.
57. Schmauch B, Elsoukkary SS, Moro A, Raj R, Wehrle CJ, Sasaki K, et al. Combining a deep learning model with clinical data better predicts hepatocellular carcinoma behavior following surgery. J Pathol Inf. 2024;15:100360. https://doi.org/10.1016/j.jpi.2023.100360.
58. Feng L, Huang W, Pan X, Ruan F, Li X, Tan S, et al. Predicting overall survival in hepatocellular carcinoma patients via a combined MRI radiomics and pathomics signature. TRANSL ONCOL. 2025;51:102174. https://doi.org/10.1016/j.tranon.2024.102174.
59. Xie Q, Zhao Z, Yang Y, Wang X, Wu W, Jiang H, et al. A clinical-radiomic-pathomic model for prognosis prediction in patients with hepatocellular carcinoma after radical resection. CANCER MED-US. 2024;13(11):e7374. https://doi.org/10.1002/cam4.7374.
60. Zeng Q, Klein C, Caruso S, Maille P, Allende DS, Minguez B, et al. Artificial intelligence-based pathology as a biomarker of sensitivity to atezolizumab-bevacizumab in patients with hepatocellular carcinoma: a multicentre retrospective study. LANCET ONCOL. 2023;24(12):1411–22. https://doi.org/10.1016/S1470-2045(23)00468-0.
61. Wang X, Jiang Y, Yang S, Wang F, Zhang X, Wang W, et al. Foundation model for predicting prognosis and adjuvant therapy benefit from digital pathology in GI cancers. J Clin Oncol. 2025. https://doi.org/10.1200/JCO-24-01501.
62. Armstrong AJ, Liu VYT, Selvaraju RR, Chen E, Simko JP, DeVries S, et al. Development and validation of an artificial intelligence digital pathology biomarker to predict benefit of long-term hormonal therapy and radiotherapy in men with high-risk prostate cancer across multiple phase III trials. J Clin Oncol. 2025. https://doi.org/10.1200/JCO.24.00365.
63. Da Q, Wang SH, Wang WW, Yang CX, Wang B, Ruan M, et al. [Progress and challenges of pathological artificial intelligence in the era of large models]. Zhonghua Bing Li Xue Za Zhi. 2025;54(3):305–9. https://doi.org/10.3760/cma.j.cn112151-20240926-00640.

64. Ding T, Wagner SJ, Song AH, Chen RJ, Lu MY, Zhang A et al. Multimodal whole slide foundation model for pathology. 2024. Preprint. Available from: arXiv:2411.19666v1.
65. Xu Y, Wang Y, Zhou F, Ma J, Jin C, Yang S et al. A Multimodal knowledge-enhanced whole-slide pathology foundation model. 2024. Preprint. Available from: arXiv:2411.19666v1.
66. Yang Z, Wei T, Liang Y, Yuan X, Gao R, Xia Y, et al. A foundation model for generalizable cancer diagnosis and survival prediction from histopathological images. NAT COMMUN. 2025;16(1):2366. https://doi.org/10.1038/s41467-025-57587-y.
67. Xiang J, Wang X, Zhang X, Xi Y, Eweje F, Chen Y, et al. A vision-language foundation model for precision oncology. Nature. 2025;638(8051):769–78. https://doi.org/10.1038/s41586-024-08378-w.
68. Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, Chow AK, et al. A multimodal generative AI copilot for human pathology. Nature. 2024;634(8033):466–73. https://doi.org/10.1038/s41586-024-07618-3.
69. Sun Y, Zhu C, Zheng S, Zhang K, Sun L, Shui Z et al. PathAsst: a generative foundation AI assistant towards artificial general intelligence of pathology. 2023. Preprint. Available from: arXiv:2305.15072v2. Cited 19 Feb 2024.
70. Wu X, Li W, Tu H. Big data and artificial intelligence in cancer research. TRENDS CANCER. 2024;10(2):147–60. https://doi.org/10.1016/j.trecan.2023.10.006.
71. Ke J, Shen Y, Lu Y, Guo Y, Shen D. Mine local homogeneous representation by interaction information clustering with unsupervised learning in histopathology images. COMPUT METH PROG BIO. 2023;235:107520. https://doi.org/10.1016/j.cmpb.2023.107520.
72. Jiang B, Bao L, He S, Chen X, Jin Z, Ye Y. Deep learning applications in breast cancer histopathological imaging: diagnosis, treatment, and prognosis. BREAST CANCER RES. 2024;26(1):137. https://doi.org/10.1186/s13058-024-01895-6.
73. Weng Z, Seper A, Pryalukhin A, Mairinger F, Wickenhauser C, Bauer M, et al. GrandQC: A comprehensive solution to quality control problem in digital pathology. NAT COMMUN. 2024;15(1):10685. https://doi.org/10.1038/s41467-024-54769-y.
74. Deng Y, Feng M, Jiang Y, Zhou Y, Qin H, Xiang F, et al. Development of pathological reconstructed high-resolution images using artificial intelligence based on whole slide image. MEDCOMM. 2020;1(3):410–7. https://doi.org/10.1002/mco2.39.
75. Gao P, Xiao Q, Tan H, Song J, Fu Y, Xu J, et al. Interpretable multi-modal artificial intelligence model for predicting gastric cancer response to neoadjuvant chemotherapy. CELL REP MED. 2024;5(12):101848. https://doi.org/10.1016/j.xcrm.2024.101848.
76. Perez-Lopez R, Ghaffari Laleh N, Mahmood F, Kather JN. A guide to artificial intelligence for cancer researchers. NAT REV CANCER. 2024;24(6):427–41. https://doi.org/10.1038/s41568-024-00694-7.
77. Collins GS, Dhiman P, Andaur Navarro CL, Ma J, Hooft L, Reitsma JB, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ OPEN. 2021;11(7):e048008. https://doi.org/10.1136/bmjopen-2020-048008.
78. Pantanowitz L, Hanna M, Pantanowitz J, Lennerz J, Henricks WH, Shen P, et al. Regulatory aspects of artificial intelligence and machine learning. Mod PATHOL. 2024;37(12):100609. https://doi.org/10.1016/j.modpat.2024.100609.
79. Arora A, Alderman JE, Palmer J, Ganapathi S, Laws E, McCradden MD, et al. The value of standards for health datasets in artificial intelligence-based applications. NAT MED. 2023;29(11):2929–38. https://doi.org/10.1038/s41591-023-02608-w.
80. Hanna MG, Pantanowitz L, Jackson B, Palmer O, Visweswaran S, Pantanowitz J, et al. Ethical and Bias considerations in artificial intelligence/machine learning. Mod PATHOL. 2025;38(3):100686. https://doi.org/10.1016/j.modpat.2024.100686.
81. Rivera SC, Liu X, Chan A, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. BMJ-BRIT MED J. 2020;370:m3210. https://doi.org/10.1136/bmj.m3210.
82. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. BMJ-BRIT MED J. 2020;370:m3164. https://doi.org/10.1136/bmj.m3164.
83. Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ-BRIT MED J. 2025;388:e081554. https://doi.org/10.1136/bmj-2024-081554.
84. Hanna MG, Pantanowitz L, Dash R, Harrison JH, Deebajah M, Pantanowitz J, et al. Future of artificial Intelligence-Machine learning trends in pathology and medicine. Mod PATHOL. 2025;38(4):100705. https://doi.org/10.1016/j.modpat.2025.100705.
85. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. LANCET ONCOL. 2019;20(5):e253–61. https://doi.org/10.1016/S1470-2045(19)30154-8.

Publisher's note