

Review Article



Artificial Intelligence in the Pathology of Gastric Cancer

Sangjoon Choi ¹, Seokhwi Kim ^{2,3}

¹Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

²Department of Pathology, Ajou University School of Medicine, Suwon, Korea

³Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Korea

OPEN ACCESS

Received: May 29, 2023

Revised: Jul 9, 2023

Accepted: Jul 14, 2023

Published online: Jul 26, 2023

Correspondence to

Seokhwi Kim

Department of Pathology, Ajou University School of Medicine, Department of Biomedical Sciences, Ajou University Graduate School of Medicine, 164 Worldcup-ro, Yeongtong-gu, Suwon 16499, Korea.
Email: seokhwikim@ajou.ac.kr

Copyright © 2023. Korean Gastric Cancer Association

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs

Sangjoon Choi

<https://orcid.org/0000-0003-2108-0575>

Seokhwi Kim

<https://orcid.org/0000-0001-7646-5064>

Funding

This work was supported by the new faculty research fund of Ajou University School of Medicine.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

ABSTRACT

Recent advances in artificial intelligence (AI) have provided novel tools for rapid and precise pathologic diagnosis. The introduction of digital pathology has enabled the acquisition of scanned slide images that are essential for the application of AI. The application of AI for improved pathologic diagnosis includes the error-free detection of potentially negligible lesions, such as a minute focus of metastatic tumor cells in lymph nodes, the accurate diagnosis of potentially controversial histologic findings, such as very well-differentiated carcinomas mimicking normal epithelial tissues, and the pathological subtyping of the cancers. Additionally, the utilization of AI algorithms enables the precise decision of the score of immunohistochemical markers for targeted therapies, such as human epidermal growth factor receptor 2 and programmed death-ligand 1. Studies have revealed that AI assistance can reduce the discordance of interpretation between pathologists and more accurately predict clinical outcomes. Several approaches have been employed to develop novel biomarkers from histologic images using AI. Moreover, AI-assisted analysis of the cancer microenvironment showed that the distribution of tumor-infiltrating lymphocytes was related to the response to the immune checkpoint inhibitor therapy, emphasizing its value as a biomarker. As numerous studies have demonstrated the significance of AI-assisted interpretation and biomarker development, the AI-based approach will advance diagnostic pathology.

Keywords: Artificial intelligence; Gastric cancer; Computer-assisted diagnosis; Immunohistochemistry; Biomarker

INTRODUCTION

Gastric cancer (GC) is one of the most common malignancies and remains the sixth-leading cause of cancer-related deaths in Korea [1]. The heterogeneity of GC is well-known, with varied histological morphology, tumor differentiation, and molecular alterations that are significantly associated with patient prognosis [2,3]. In the current era of precision medicine and cancer therapeutics, the accurate diagnosis and staging as well as the precise molecular classification and identification of novel biomarkers are critical for optimal treatment.

Author Contributions

Conceptualization: K.S.; data curation: C.S., K.S.; funding acquisition: K.S.; supervision: K.S.; writing - original draft: C.S., K.S.; writing - review & editing: C.S., K.S.

Microscopic evaluation of the histomorphology from hematoxylin and eosin (H&E)-stained slides is the gold standard in GC diagnosis [4]. However, histopathological diagnosis faces several challenges, including a scarcity of skilled pathologists, inter-observer discordance, and the inherent human limits of erroneous diagnosis [5-7]. Consequently, the demand for supporting systems that can contribute to more precise and consistent diagnostic practices continues to grow. In recent years, digital pathology using whole-slide images (WSIs) has emerged as a powerful tool for the analysis of pathologic specimens [8,9]. The use of slide scanners has enabled the digitalization of large numbers of slides and the construction of numerous datasets for diagnosis and research [10].

The analysis of such large pathologic image datasets is time-consuming and requires significant expertise. The application of artificial intelligence (AI) to pathologic image analysis is a promising solution to the challenges. Deep learning (DL) is one of the several machine learning AI algorithms that can learn hierarchical features from large-scale data. In contrast to conventional machine learning algorithms, DL includes data-based feature extraction in the training process and comprises multi-layered neural network architecture that can learn the hierarchical features [10-14]. The convolutional neural network (CNN) is the most common DL strategy used for image analysis, wherein the predictive features are not defined but the model learns the concepts and features that are useful for explaining the relationships between inputs and outputs. According to the degree of labeling of the input data, the DL algorithm can be classified as supervised, weakly-supervised, and unsupervised learning [15]. Currently, DL algorithms are applied for precise pathologic diagnosis [16,17] and to determine predictive and prognostic biomarkers from routine histology images [18-20]. The integration of AI with pathological diagnosis is particularly important in the diagnosis of GC, considering its histopathological and molecular heterogeneity [2]. The application of AI to pathology enables precise diagnoses, optimal staging, and clear understanding of the molecular characteristics of GC. This advancement enables more effective and personalized medical decision-making, ultimately leading to improved outcomes for GC patients.

In this review, we aimed to introduce the up-to-date utilization of AI in GC pathology. In particular, we summarized the representative data on the following 3 aspects: 1) detection of basic histologic components and diagnosis; 2) assessment of clinically important immunohistochemical staining; and 3) prediction of clinical outcomes and biomarker detection (**Table 1**).

DETECTING BASIC HISTOLOGIC COMPONENTS AND DIAGNOSIS

Histopathologic diagnosis

The pathologic diagnosis of biopsy and resection specimens of GC is primarily conducted based on the histomorphology of H&E slides. Recently developed DL algorithms enabled the accurate detection of cancer and non-cancer cell types as well as the cancer area and stroma in various types of cancers (**Fig. 1**).

Among the various studies regarding AI-based histomorphologic diagnosis of GC, a recent study described the development of a GC detection algorithm in biopsy specimens of daily pathologic practice [21]. A DL algorithm to classify gastric biopsy images into 3 categories

Table 1. Pivotal studies on the application of artificial intelligence to gastric cancer pathology

Reference	Purpose	Datasets	Algorithm	Results summary
Cancer detection and histological diagnosis				
Park et al. [21]	Histopathologic diagnosis of NFD, TA, and GC	9,866 WSIs	Representation aggregation CNN	AUROC 0.9790 for NFD vs. non-NFD (sensitivity 1.000, specificity 0.9749); 47% reduction of review time per image with DL assistance
Song et al. [22]	Histopathologic diagnosis of GC	4,794 WSIs	DeepLab v3, ResNet-50	AUROC 0.986 (accuracy 87.3%, sensitivity 99.6%, specificity 84.3%); Improved diagnostic accuracy of DL-assisted group compared with microscope or WSI-alone groups
Kanavati et al. [23]	Histopathologic diagnosis of GC	4,969 WSIs	EfficientNetB1, Inception v3	AUROC 0.95–0.99 (average accuracy 91.7%, sensitivity 93.6%, specificity 91.7%)
Ba et al. [24]	Histopathologic diagnosis of GC	110 WSIs	DeepLab v3	Higher sensitivity with DL-assistance (90.63% vs. 82.75%); Reduced review time per slide with DL-assistance (22.68 vs. 26.37 seconds)
Matsushima et al. [25]	LN metastasis detection	40 patients	ResNet-152	AUROC 0.9994, FROC of 1.00 sensitivity with only 3 false positives
Wang et al. [26]	LN metastasis detection T/MLN validation	9,366 WSIs	ResNet-50	LN metastasis detection accuracy 96.9%, sensitivity 98.5%, specificity 96.1%; High T/MLN was independent prognostic factor (univariate analysis, HR, 2.05, 95% CI, 1.66–2.54; multivariate analysis, HR, 1.39, 95% CI, 1.10–1.75)
Hu et al. [27]	LN metastasis detection	921 WSIs	Xception, DenseNet-121	LN quantification accuracy 97.13%; LN metastasis detection PPV 93.53%, NPV 97.99%
Clinically important immunohistochemistry interpretation				
Kim et al. [28]	PD-L1 CPS scoring Pembrolizumab response prediction	39 WSIs	Aperio IHC image analysis	84.6% concordance with pathologists; No significant difference for pembrolizumab response prediction
Han et al. [29]	HER2 scoring	183 WSIs	RepVGG	Accuracy 94.0%
Prediction of clinical outcome and biomarker detection				
Huang et al. [30]	Prognostic prediction	2,508 pictures	GastroMIL (RegNet, RNN), MIL-GC (MLP)	OS prediction C-index (0.657–0.728); Risk score by DL-model as an independent prognostic predictor (univariate analysis, HR, 2.414; multivariate analysis, HR, 1.803)
Wei et al. [31]	Prognostic prediction	382 WSIs	DeepCox-SC, MultiDeepCox-SC (InceptionResNetV2)	OS prediction C-index based on histopathology image only (0.660) and histopathology + clinical + gene expression (0.744)
Muti et al. [32]	EBV and MSI prediction	2,617 WSIs, 366 TMAs	ShuffleNet	AUROC 0.723–0.863 for MSI; AUROC 0.672–0.859 for EBV
Kather et al. [33]	MSI prediction	500 WSIs	ResNet-18	AUROC 0.81 (95% CI, 0.69–0.90) in TCGA; AUROC 0.69 (95% CI, 0.52–0.82) in Japanese cohort
Jeong et al. [34]	EBV prediction	487 WSIs	EBVNet (ResNet50 and Inception v3)	Better performance than pathologists (AUROC 0.88 vs. 0.75; AUPRC 0.65 vs. 0.41)
Lee et al. [35]	MSI prediction	714 WSIs	Inception v3	AUROC 0.893 (TCGA frozen), 902 (TCGA FFPE), 0.874 (external validation)
Kather et al. [20]	Molecular alteration prediction	321 WSIs	ResNet-18, AlexNet, Inception v3, DenseNet-201, ShuffleNet	AUROC for the top 8 mutations ranged from 0.66 to 0.78

DL = deep learning; NFD = negative for dysplasia; TA = tubular adenoma; GC = gastric carcinoma; WSI = whole-slide image; CNN = convolutional neural network; AUROC = area under receiver operating characteristic curve; LN = lymph node, FROC = free receiver operating characteristic curve; T/MLN = tumor-area-to-metastatic lymph node-area ratio; HR = hazard ratio; CI = confidence interval; PPV = positive predictive value; NPV = negative predictive value; PD-L1 = programmed death-ligand 1; CPS = combined positive score; IHC = immunohistochemistry; HER2 = human epidermal growth factor receptor 2; OS = overall survival; EBV = Epstein-Barr virus; MSI = microsatellite instability; TMA = tissue microarray; TCGA = The Cancer Genome Atlas; AURPC = area under recall-precision curve; FFPE = formalin-fixed paraffin-embedded.

(negative for dysplasia, tubular adenoma, or carcinoma) was derived from 2,434 WSIs as training and validation sets, and the performance of the DL model was evaluated by 7,432 test set WSIs. The diagnostic performance of the DL model achieved an area under the receiver operating curve (AUROC) value of 0.9790 for 2-tier classification (negative for dysplasia vs. positive). The sensitivity and specificity for diagnosis were 100% and 97.49%, respectively, when only epithelial tumors were included for analysis. Moreover, the average review time per

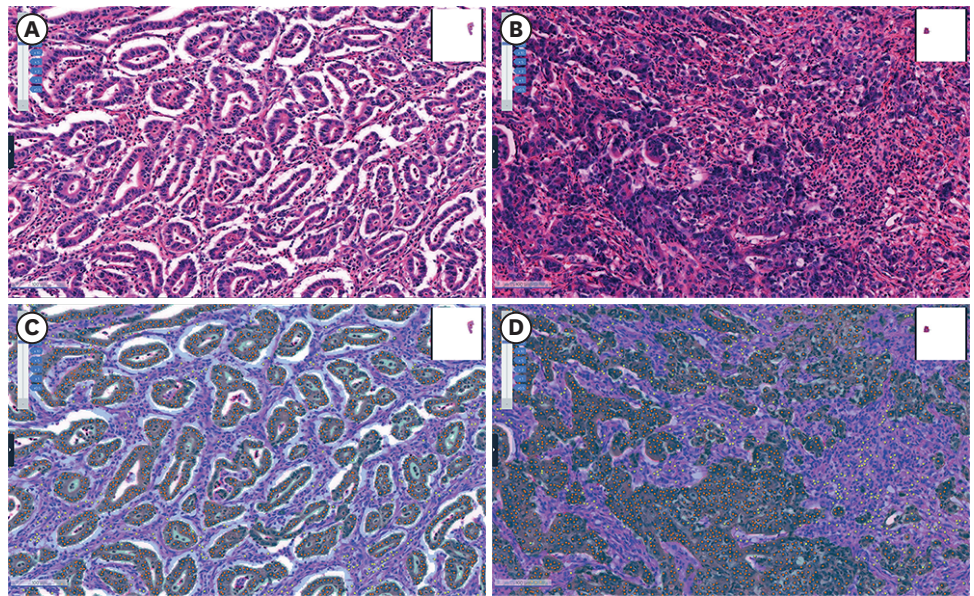


Fig. 1. Representative image of cancer detection on hematoxylin and eosin slides by the DL model. Well-to-moderately differentiated (A) and poorly differentiated (B) gastric tubular adenocarcinoma and adjacent reactive stroma with inflammatory infiltrates. The DL model can not only recognize cancer epithelium (green area) and peritumoral stroma (blue area) but also differentiate the tumor cells (orange dot) and lymphoplasmacytic infiltration (yellow dot) in (B) well-to-moderately differentiated gastric cancer, (C) poorly differentiated gastric cancer, and (D) poorly differentiated gastric cancer. DL = deep learning.

slide was significantly shorter in the DL-assisted digital image viewer group (18.90 seconds; 95% confidence interval [CI], 17.44–20.36) than in the digital image viewer (35.70 seconds; 95% CI, 33.24–38.15) and conventional microscope (44.97 seconds; 95% CI, 41.43–48.52) groups. The common cause of misinterpretation of tubular adenoma was regenerative glandular atypia and intestinal metaplasia. Inflammatory tissues, including ulcer detritus and granulation tissue, were the most common cause of misinterpretation as carcinoma.

Another study by Song et al. [22] also suggested that the standalone performance of AI is sufficiently high to detect and diagnose GC in the H&E WSI images. They trained the deep CNN with 2,123 pixel-level annotated H&E-stained WSI and validated the algorithm with 3,212 WSIs of gastric tissue slides. The DL model revealed a stable performance with an average AUROC of 0.986 (accuracy, sensitivity, and specificity of 87.3%, 99.6%, and 84.3%, respectively), showing reproducible sensitivity and specificity across independent institutions. Furthermore, the diagnostic accuracy of pathologists who used DL models improved compared with that of pathologists who did not use a DL-assisted diagnosis system for both microscopy and WSI. The DL algorithm also identified 2 initially misdiagnosed cases as negative for malignancy, which were overlooked by pathologists. Conversely, the DL model misinterpreted some cases as false positive malignancies, such as those involving mucus retention extravasation, foamy histiocytes, inflammatory exudates, myofibroblast proliferation, overstained images, or folded sections.

Poorly differentiated adenocarcinoma and signet ring cell carcinoma are often inconspicuous and occasionally misdiagnosed as gastritis by pathologists. Kanavati et al. [23] trained the DL model to classify diffuse type GC from H&E slides using 2 different approaches. First, the model was trained to directly classify diffuse type cancer. Alternatively, the 2-stage approach

involved an initial step where the model was used to detect adenocarcinoma, followed by a second stage of further classifying the cases into either diffuse or intestinal type. Both models obtained high AUROCs in the range of 0.95–0.99 (average accuracy, sensitivity, specificity of 91.7%, 93.6%, and 91.7%, respectively) in 5 independent test datasets. The rare false-positive and false-negative predictions stemmed from the misinterpretation of lymphoplasmacytic infiltration in lamina propria as cancer cells and missing degenerative cancer cells in necrotic tissue, respectively.

A recent study by Ba et al. [24] highlighted that DL-assisted interpretation of pathologists could enhance the accuracy and efficiency of GC diagnosis. They conducted a multi-reader study of 16 pathologists using 110 WSIs with or without DL assistance. The AUROC in the DL-assisted pathologist group was significantly higher (0.911 vs. 0.863, $P=0.003$) than that of pathologists without DL assistance. Additionally, pathologists with DL assistance showed higher sensitivity in GC detection than those without assistance (90.63% vs. 82.75%, $P=0.010$), whereas no significant difference was observed in specificity between the 2 groups (78.34% vs. 79.90%, $P=0.468$). Additionally, the use of DL assistance reduced the average review time per WSI in the DL-assisted group compared with that in the group without assistance (22.68 vs. 26.37 seconds, $P=0.033$).

These findings demonstrated that DL algorithms can detect and differentiate GC cells in H&E slides with sufficiently high performance. Nevertheless, nearly all DL algorithms occasionally revealed misinterpretations that could easily be noticed and corrected by pathologists. Therefore, using DL algorithms as a screening tool or assistance modality for pathologists in the diagnosis of GC specimens can be a plausible approach.

Detection of lymph node metastasis

The regional lymph node metastasis is one of the most important prognostic factors in GC as the tumor, node, and metastasis stage is determined by the number of metastatic lymph nodes [36]. The evaluation of lymph node metastasis under the microscope is laborious for pathologists, especially in the case of micrometastasis (size ranging between 0.2 and 2 mm) [37]. Several studies have demonstrated that DL-based cancer detection on lymph nodes can aid pathologists in their daily practice.

Matsushima et al. [25] trained a DL algorithm with H&E slides, including a total of 51 metastasis-positive nodes and 776 metastasis-negative nodes from 20 patients. The developed DL model achieved an excellent performance with an AUROC of 0.9994. The free-response ROC showed perfect sensitivity with only 3 false positives, wherein histiocytes were detected as cancer cells. Furthermore, the model's performance was validated on an independent dataset, which yielded similar results with an AUROC of 0.9914.

Another study by Wang et al. [26] that used 6,418 WSIs, including 15,362 lymph nodes with 4,974 metastases, also showed that the DL model's tumor detection performance in metastatic lymph nodes was comparable with that of experienced pathologists. The accuracy of the DL algorithm was 96.9%, with sensitivity and specificity values of 98.5% and 96.1%, respectively. The sensitivity was influenced by the histological tumor subtype, and misdiagnosis by the DL model mainly occurred in cases of mucinous adenocarcinoma and signet ring cell carcinoma. Moreover, the diagnosis of DL-assisted pathologists showed that 4.6% of cases were initially under-staged by pathologists owing to the neglect of micrometastases. The high tumor-area-to-metastatic lymph node-area ratio (≥ 0.45) was

associated with worse prognosis in GC patients of the same N stage, revealing the value as an independent prognostic factor in both univariate (hazard ratio [HR], 2.05; 95% CI, 1.66–2.54; $P < 0.001$) and multivariate (HR, 1.39; 95% CI, 1.10–1.75; $P = 0.007$) analyses. The results suggest that the DL model can not only provide an accurate histopathologic diagnosis but also a novel prognostic factor of lymph node metastasis in H&E slides, which could be difficult to manually quantify by pathologists.

Hu et al. [27] also utilized 921 WSIs of lymph nodes to develop DL algorithms for both lymph node identification and metastatic cancer detection. The model was highly efficient and accurate in detecting metastatic lymph nodes; the accuracy of lymph node quantification was 97.13%, and the positive and negative predictive values were 93.53% and 97.99%, respectively. As revealed by the results of Wang et al. [26], the false negative rate of the DL model in poorly differentiated adenocarcinoma and signet ring cell carcinoma was higher compared with that of well to moderately differentiated adenocarcinoma. An example of a false positive cancer area was the germinal center of the lymph node, which includes histomorphologically bizarre lymphoid cells owing to a high level of proliferation.

Mitosis counting

Mitosis detection is one of the earliest applications of AI because the mitotic count is included in the histopathologic grading of neoplasms, such as breast cancer and soft tissue sarcoma [38,39]. The mitotic index is also one of the histological prognostic factors of gastrointestinal stromal tumor (GIST) [40] and neuroendocrine tumors [41]. A higher mitotic index is associated with aggressive tumor behavior, including recurrence and distant metastasis, leading to a poorer prognosis. Studies have shown that AI algorithms could successfully count the mitosis of tumor cells for neuroendocrine tumor, invasive breast carcinoma, and leiomyosarcoma [41–43]; however, its effectiveness in GIST has not yet been demonstrated. As the morphology of mitotic figures is similar across different tumor types, similar AI models could be applied to GIST. The AI-based mitosis detection algorithm could help pathologists in the accurate counting of mitosis with reduced time and labor.

IMMUNOHISTOCHEMISTRY (IHC) INTERPRETATION

Programmed death-ligand 1 (PD-L1) scoring

In addition to the image analysis of H&E-stained WSI, AI algorithms have been used for IHC interpretation to reduce possible inter-observer variability of reading [44–47]. Among the various IHC markers, AI-based interpretation of PD-L1 is important in the histopathologic diagnosis of GC (**Fig. 2**). PD-1 is an inhibitory checkpoint receptor protein found on immune cells, including cytotoxic T-cells [48]. Some tumor cells express elevated PD-L1 expression on their surface, leading to the inactivation of cytotoxic T-cells and the downregulation of immune responses [49]. PD-L1 protein expression is assessed by IHC for selecting patients eligible for immune checkpoint inhibitor (ICI) treatment [50,51]. Several studies have demonstrated that digital image analysis of the tumor proportion score (TPS) of PD-L1 is concordant with manual assessment by pathologists in various tumors, although AI-aided PD-L1 TPS assessment in GC has not been attempted yet [46,52,53]. A DL algorithm for PD-L1 TPS interpretation in non-small cell lung cancer demonstrated that the DL-assisted evaluation of PD-L1 IHC by pathologists could increase the concordance of readings, especially around the clinically relevant TPS value of 1, leading to a more accurate prediction of the ICI response [44].

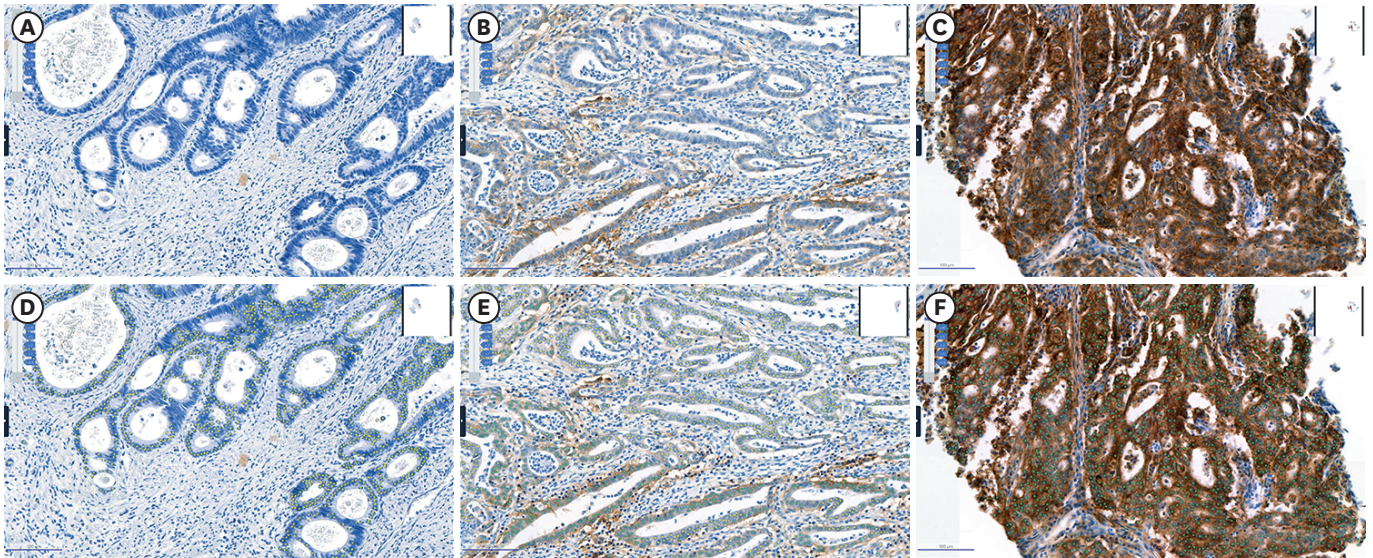


Fig. 2. Representative image of DL-based PD-L1 CPS quantification from PD-L1 IHC slides of gastric tubular adenocarcinoma. PD-L1 IHC exhibits heterogeneous immunoreactivity by tumor and mononuclear cells, with (A) CPS <1, (B) CPS 1–49, and (C) CPS ≥50. The DL-model can distinguish PD-L1 positive (purple dot) vs. negative (sky-blue dot) tumor cells and PD-L1 positive (yellow dot) vs. negative (red dot) lymphoplasmic cells (D–F). DL = deep learning; PD-L1 = programmed death-ligand 1; CPS = combined positive score; IHC = immunohistochemistry.

Interpretation of the combined positive score (CPS), which is utilized in GC, is more complex than that of TPS because the numerator not only includes the PD-L1 positive tumor cells but also the mononuclear inflammatory cells in the tumoral stroma. A few AI-based PD-L1 CPS interpretation models in urothelial carcinoma and head and neck squamous cell carcinoma have been reported to show good performance [54,55]. For GC, an AI-based PD-L1 CPS interpretation model has not yet been developed. However, Kim et al. [28] utilized a digital image analysis model, namely, the Aperio imagescope IHC membrane image analysis algorithm, to obtain the PD-L1 CPS, and compared the results with the interpretation of pathologists using 39 cases of GC with a CPS cutoff of 1. The results of the image analysis model and scoring of pathologists were also compared with the clinical responses of patients to pembrolizumab. Although the model was not AI-based, it showed 84.6% concordance with the interpretation of pathologists. Moreover, in predicting pembrolizumab responses, the PD-L1 CPS obtained by the model and the score interpreted by pathologists did not reveal a statistically significant difference ($P=0.1856$). The development and validation of AI-based CPS interpretation algorithms in GC will be invaluable for the stratification of patients who will benefit from ICI treatment.

Human epidermal growth factor receptor 2 (HER2) scoring

Another application of AI to the IHC of GC is the quantification of HER2. HER2 is known to be associated with tumor cell proliferation, adhesion, migration, and differentiation [56]. The overexpression or amplification of HER2 is identified in a subset of GC cases, and this is particularly important because of the well-established targeted therapeutics, namely, trastuzumab [57]. Current guidelines recommend IHC and in-situ hybridization for detecting the overexpression or amplification of HER2 [58]. HER2 IHC is scored based on the staining intensity (0, 1+, 2+, and 3+) and the proportion of tumor cells with membranous staining. The HER2 score evaluation on IHC images by AI models has been thoroughly studied for breast cancer [59–62] but has rarely been investigated in GC. Han et al. [29] developed a DL-based HER2 scoring algorithm for GC using a total of 183 HER2-stained WSIs for training,

validation, and testing. The DL model successfully distinguished the staining intensity and proportion of tumor cells and showed effectiveness with an accuracy of 94% for HER2 scoring prediction on WSIs.

PREDICTION OF THE CLINICAL OUTCOME AND BIOMARKER DETECTION

Prognostic prediction from H&E images

AI algorithms could not only detect and diagnose cancer cells but also predict the clinical outcome of patients based on the histopathology of various types of cancers [63-66]. Huang et al. [30] recently developed a DL algorithm for the prognostic prediction of GC using 2,333 H&E images from real-world datasets and The Cancer Genome Atlas (TCGA) program data. The algorithm calculated the risk score based on the most suspicious tiles of WSIs and grouped the GC patients into low- and high-risk subgroups based on the H&E morphology. Notably, the risk-predictive tiles represented features of high-grade malignancies, such as necrosis, vascular or neural invasion, and single-cell invasive morphology. The DL-based prediction of overall survival (OS) in GC patients showed C-index values of 0.728 and 0.671 in the training and internal validation sets, respectively, and the performance was also sufficiently high in the independent external validation set (C-index = 0.657). Moreover, the risk score generated by the DL model was a strong prognostic indicator of GC patients both in univariate (HR, 2.414; $P < 0.0001$) and multivariate (HR, 1.803; $P = 0.043$) analyses. This study demonstrated that the DL model could effectively stratify GC patients based on their risk level, thereby enabling the selection of patients who are more likely to benefit from intensive adjuvant therapy and close follow-ups.

Another study by Wei et al. [31] introduced a DL-based model, called MultiDeepCox-SC, that predicts OS in GC patients by integrating histopathological images with clinical and gene expression data. They trained the model using 10-fold cross-validation with 382 WSIs from the TCGA dataset, and the model calculated the risk score using H&E images, age, gender, tumor grade, tumor stage, and gene expression profiles to predict the patient's risk of death. The C-index of the model had a prognostic accuracy of 0.660 when relying solely on histopathological images; however, the value significantly increased to 0.744 when clinical and genetic risk factors associated with prognosis were included to calculate the risk score. Subgroup analysis revealed a significant difference in OS between the high- and low-risk score groups among GC patients of the same stage. Moreover, the risk score was an independent predictor of OS in the multivariate analysis on the TCGA dataset (HR, 1.555; $P = 3.53 \times 10^{-8}$) and the external test dataset (HR, 2.912; $P = 9.42 \times 10^{-4}$).

Prediction of molecular subtypes of GC

The 4 molecular subtypes of gastric adenocarcinoma proposed by TCGA project are as follows: 1) Epstein-Barr (EBV)-infected, 2) microsatellite instability (MSI), 3) genomically stable (GS), and 4) chromosomally unstable tumors (chromosomal instability; CIN) [2]. Subsequent studies have revealed that these molecular subgroups were associated with distinct clinical outcomes and therapeutic responses [67,68]. The prediction of molecular subtypes of GC from H&E images can provide a cost-effective method for patient stratification for clinicians.

Two DL-based models for predicting molecular subgroups of GC have been introduced to date. Wang et al. [69] developed a DL model using 332 H&E-stained WSIs from the TCGA

database that were segmented into 372,505 tiles with molecular subgroup and clinical outcome information. The authors split the entire dataset into 10 subsets and each was trained by the individual DL model; subsequently, an ensemble model was constructed according to the average values of the outputs from the 10 models. The ensemble model showed an enhanced prediction of the 4 molecular subtypes compared with the individual models; it achieved AUROC values of 0.785, 0.668, 0.762, and 0.811 for the prediction of GS, EBV, CIN, and MSI subtypes in tile-level images, and AUROC values of 0.897, 0.764, 0.890, and 0.898 for each subtype at the patient-level. Flinner et al. [70] conducted an independent study to develop a DL model to predict the molecular subtypes of GC using 133 WSIs from the TCGA-stomach adenocarcinoma (STAD) dataset for training and 65 United Kingdom Collaborative (UKC) images for validation. They also utilized an ensemble approach based on majority voting on multiple CNNs to maximize predictability. Compared with laboratory-based molecular testing to determine the molecular subtypes of their own GC cohort collected at UKC, the DL model showed the variable levels of prediction of each molecular subtype. The authors presumed that the intratumoral heterogeneity in GC might have resulted in the low predictability in some cases.

Among the molecular subtypes, the response of GC to ICIs, such as pembrolizumab, is known to be associated with EBV positivity or MSI-high status [71]. Therefore, studies have rigorously been conducted to develop AI algorithms to predict the MSI status and EBV infection directly from H&E histology for their use as beneficial predictive biomarkers.

In a retrospective multicenter cohort study, Muti et al. [32] successfully developed and validated a DL model to detect the EBV and MSI status from histology slides of GC. They obtained tissue samples from ten clinical cohorts of more than 2,500 GC patients from 7 countries (South Korea, Switzerland, Japan, Italy, Germany, the UK, and the USA), and trained the DL model to detect MSI and EBV positivity from digitized H&E slides. First, they trained and validated the DL model within each cohort in a 3-fold cross-validation design. Second, they selected the 5 largest cohorts as pooled training datasets and tested them on the 5 remaining cohorts. In the within-cohort cross-validation, the AUROC values ranged from 0.597 to 0.836 for the detection of MSI, and from 0.819 to 0.897 for the detection of EBV. When trained on a pooled cohort, the AUROC values ranged from 0.723 to 0.863 for the detection of MSI, and from 0.672 to 0.859 for the detection of EBV.

Kather et al. [33] conducted an independent study showing that their DL algorithm can directly predict the MSI status in GC from H&E histology. They utilized histology images of 315 patients from the TCGA-STAD dataset for training, whereas a randomly selected group of 99 patients was used as a test set. As GC in Asian patients has different clinicopathological features compared with that in non-Asian patients, and TCGA-STAD mostly comprises non-Asian patients, an additional cohort of Japanese patients (n=185) was used for external validation. The DL algorithm achieved an AUROC value of 0.81 and 0 in the TCGA-STAD and Japanese cohorts, respectively.

In a study based on the Korean population, Jeong et al. [34] developed a DL-based EBV prediction method from H&E-stained WSIs, using 319 slides from the TCGA dataset and 108 slides from an independent institution. An additional 60 WSIs from other institutions were used for external validation. The EBV prediction model achieved an AUROC value of 0.88 and an area under the precision-recall curve (AUPRC) value of 0.65, both of which were higher than the mean AUROC and AUPRC values of 0.75 and 0.41, respectively, achieved by

pathologists. Moreover, the model achieved negative predictive value, sensitivity, specificity, precision, and F1-score of 0.98, 0.86, 0.92, 0.60, and 0.71, respectively.

Another study by Lee et al. [35] demonstrated the automated classification of the MSI status in GC tissue slides using the DL model. They utilized a total of 331 WSIs from the TCGA database, including both frozen and formalin-fixed paraffin-embedded (FFPE) tissues, and conducted 10-fold cross-validation. External validation was conducted using 383 WSIs of gastrectomy specimens of Korean patients. The AUROC values of the model were 0.893 and 0.902 for the frozen and FFPE tissues from TCGA, respectively. The model also achieved good performance with an AUROC value of 0.874 on the external validation cohort.

Prediction of individual genetic alterations

The AI algorithm can serve as a screening tool for the molecular profiling of cancer [20,72-74]. Currently, targetable genetic alterations of cancers are mainly identified through next-generation sequencing (NGS); however, insufficient or inadequate sampling of the tumor and the relatively high cost of NGS are practical limitations. Therefore, studies have attempted to develop DL-based models to predict genetic alterations directly from H&E images.

Kather et al. [20] reported that genetic mutations, molecular tumor subtypes, gene expression signatures, and standard pathology biomarkers could be predicted from H&E images by DL algorithms in various types of cancers. For GC, 321 TCGA-STAD WSIs were used for the training and evaluation of DL networks by 3-fold cross-validation. Among all tested tumor types, GC showed the highest absolute number of detectable mutations, and the mean cross-validated AUROC value for the top 8 mutations ranged from 0.66 to 0.78. Key driver mutations of GC, including *AMER1*, *TP53*, *MTOR*, *FBXW*, and *PIK3CA*, were successfully detected ($P < 0.05$ after false discovery rate correction). Moreover, gene expression signatures of proliferation, hypermutation, stem cell properties, macrophages, homologous recombination repair deficiency, and wound healing could be detected from histology by the DL algorithm. These results suggest that AI can bridge the gap between genomics and phenomics in GC.

Classifying the tumor immune microenvironment as a biomarker

The distribution of tumor-infiltrating lymphocytes (TILs) in the tumor microenvironment (TME) is a key biomarker in predicting the ICI response in various cancers [75]. Previous studies have demonstrated that the TIL distribution patterns in tumor areas and peritumoral stroma could be classified into 3 distinct immune phenotypes: inflamed (TILs located inside the tumor), excluded (TILs confined to the peritumoral stroma), and desert (few TILs in both tumor nests and peritumoral stroma) [76-78]. Each phenotype encompassed distinct biological traits, and the inflamed phenotype showed the increased expression of PD-L1 by tumor and immune cells, leading to a superior response to immunotherapy compared with other phenotypes [77,78]. Park et al. [79] developed an AI-powered WSI analyzer of TIL in the TME to classify immune phenotypes and demonstrated that the spatial analysis of TIL was significantly correlated with tumor response and patient prognosis in advanced non-small cell lung cancer patients. As GC is known to have a heterogenous tumor immune microenvironment, namely, considering the amount of TIL, immune cell subtypes, and their distribution, a similar AI-based approach could be applicable to GC to select patients who are likely to benefit from ICI treatment.

EMERGING AREAS AND FUTURE CONSIDERATIONS

Recent studies have shown that diverse components of the TME other than TILs, including peritumoral fibrosis, cancer-associated fibroblasts (CAFs), and tertiary lymphoid structures (TLSs), are significantly associated with the biological behavior and prognosis of GC [80-83]. The fibrosis pattern in GC has not yet been analyzed by AI, but several AI-based diagnostic algorithms have successfully been developed to evaluate liver fibrosis [84,85], raising the possibility to analyze the fibrosis pattern or CAF distribution in GC by AI algorithms. A recent study revealed that an AI-based classification of TLS in gastrointestinal cancers, including GC, was significantly associated with the patient's prognosis [86]. The H&E and IHC images showing the distribution of tumor cells, lymphoid cells, and other cell types with the expression of relevant proteins, as well as the molecular profiles predicted by the AI algorithm, can provide a comprehensive spatial understanding of TME [87]. These approaches will help expand our understanding of tumor biology, provide precise prognostication, and identify novel therapeutic targets of GC.

No commercialized AI algorithms are currently available for GC pathology. Paige Prostate, an AI tool used to assist pathologists in the diagnosis of prostate cancer, is the only Food and Drug Administration (FDA)-approved, commercialized, clinical-grade product to date [88,89]. In the endoscopic field, an AI-based endoscopic imaging analysis tool is used for early GC detection but has only acquired FDA's Breakthrough Device designation [90]. The delayed development and productization of AI pathology tools for GC largely stem from practical challenges and limitations. The main challenge/limitation for the commercialization and clinical integration of AI algorithms is the financial issue. The development of algorithms is a time-consuming task, and the financial cost is high but the reimbursement for the cost is largely unknown [16]. From this perspective, most AI software companies recognize that the development of AI pathology algorithms is not profitable, which profoundly hampers the progress of this field [91]. Other hurdles include the generation of high-quality annotated data, integration of AI tools into existing workflows, and ensuring regulatory compliance [92-94]. Ethical issues, including data ownership and patient confidentiality, also need to be addressed [95].

Another important consideration in AI pathology is whether the AI-standalone interpretation can be utilized as a confirmative diagnostic tool without a pathologist's reading. Despite the sophistication of the algorithm, the mainstay of DL algorithms remains recognized as a "black box" and can be affected by the Clever Hans effect [96-98]; this hinders the analysis of errors even by software developers and cannot account for the legal issues. Therefore, supervision by pathologists is required for AI-based pathological analysis systems [99]. The role of pathologists is pivotal in various aspects of the development and utilization of AI pathology algorithms. Histologic slides may have different image features owing to the staining methods or slide scanners. Therefore, instead of training using a specific dataset, data from multiple institutions should be pooled to reduce bias [100]. The contributions of pathologists are inevitable during the annotation and labeling of the training of multiple datasets. Additionally, pathologists can conduct secondary validation, review outputs of AI models, and offer their expert opinions for quality assurance. When validating the developed model, both the internal and external sets must be validated to evaluate the performance of the actual AI model [11]. Based on their knowledge of tissue sample analysis, pathologists can collaborate with clinicians, data scientists, bioinformaticians, and AI experts to achieve significant progress in cancer research and clinical practice.

CONCLUSION

We reviewed the pivotal studies utilizing AI models to enhance histopathologic diagnosis and biomarker/clinical outcome predictions in the field of GC pathology. The current AI algorithms can successfully detect GC on H&E images, score IHC slides, and predict key biomarkers for clinical and therapeutic outcomes. The algorithms have showed sufficiently high performance, indicating their potential as standalone tools to detect well-established and novel biomarkers for GC from massive, diverse datasets in a more efficient manner. However, in most studies, AI-assisted interpretation by pathologists showed the best results among the AI-standalone and AI-unaided pathologist's interpretations. Numerous challenges also need to be overcome, such as the lack of standardization among AI models and the potential misinterpretation by AI models. Nevertheless, further investigation of AI-assisted pathology in GC will undoubtedly provide benefits in the era of precision medicine.

ACKNOWLEDGMENTS

We thank Dr. Soo Ick Cho for information on the artificial intelligence (AI) algorithmic specifications and commercialization of the products.

REFERENCES

1. Jung KW, Kang MJ, Park EH, Yun EH, Kim HJ, Kong HJ, et al. Prediction of cancer incidence and mortality in Korea, 2023. *Cancer Res Treat* 2023;55:400-407.
[PUBMED](#) | [CROSSREF](#)
2. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 2014;513:202-209.
[PUBMED](#) | [CROSSREF](#)
3. Nakamura Y, Kawazoe A, Lordick F, Janjigian YY, Shitara K. Biomarker-targeted therapies for advanced-stage gastric and gastro-oesophageal junction cancers: an emerging paradigm. *Nat Rev Clin Oncol* 2021;18:473-487.
[PUBMED](#) | [CROSSREF](#)
4. Guideline Committee of the Korean Gastric Cancer Association (KGCA), Development Working Group & Review Panel. Korean practice guideline for gastric cancer 2018: an evidence-based, multi-disciplinary approach. *J Gastric Cancer* 2019;19:1-48.
[PUBMED](#) | [CROSSREF](#)
5. Kim JM, Sohn JH, Cho MY, Kim WH, Chang HK, Jung ES, et al. Inter-observer reproducibility in the pathologic diagnosis of gastric intraepithelial neoplasia and early carcinoma in endoscopic submucosal dissection specimens: a multi-center study. *Cancer Res Treat* 2019;51:1568-1577.
[PUBMED](#) | [CROSSREF](#)
6. Renshaw AA, Gould EW. Measuring errors in surgical pathology in real-life practice: defining what does and does not matter. *Am J Clin Pathol* 2007;127:144-152.
[PUBMED](#) | [CROSSREF](#)
7. Robboy SJ, Weintraub S, Horvath AE, Jensen BW, Alexander CB, Fody EP, et al. Pathologist workforce in the United States: I. Development of a predictive model to examine factors influencing supply. *Arch Pathol Lab Med* 2013;137:1723-1732.
[PUBMED](#) | [CROSSREF](#)
8. Nam S, Chong Y, Jung CK, Kwak TY, Lee JY, Park J, et al. Introduction to digital pathology and computer-aided pathology. *J Pathol Transl Med* 2020;54:125-134.
[PUBMED](#) | [CROSSREF](#)
9. Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW, et al. Artificial intelligence in pathology. *J Pathol Transl Med* 2019;53:1-12.
[PUBMED](#) | [CROSSREF](#)

10. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 2019;16:703-715.
[PUBMED](#) | [CROSSREF](#)
11. van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat Med* 2021;27:775-784.
[PUBMED](#) | [CROSSREF](#)
12. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44-56.
[PUBMED](#) | [CROSSREF](#)
13. Niazi MK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol* 2019;20:e253-e261.
[PUBMED](#) | [CROSSREF](#)
14. Colling R, Pitman H, Oien K, Rajpoot N, Macklin P, Snead D, et al. Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. *J Pathol* 2019;249:143-150.
[PUBMED](#) | [CROSSREF](#)
15. Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, et al. Artificial intelligence for multimodal data integration in oncology. *Cancer Cell* 2022;40:1095-1110.
[PUBMED](#) | [CROSSREF](#)
16. Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. *Mod Pathol* 2022;35:23-32.
[PUBMED](#) | [CROSSREF](#)
17. Cao R, Tang L, Fang M, Zhong L, Wang S, Gong L, et al. Artificial intelligence in gastric cancer: applications and challenges. *Gastroenterol Rep (Oxf)* 2022;10:goac064.
[PUBMED](#) | [CROSSREF](#)
18. Coudray N, Tsirigos A. Deep learning links histology, molecular signatures and prognosis in cancer. *Nat Can* 2020;1:755-757.
[PUBMED](#) | [CROSSREF](#)
19. Fu Y, Jung AW, Torne RV, Gonzalez S, Vöhringer H, Shmatko A, et al. Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nat Can* 2020;1:800-810.
[PUBMED](#) | [CROSSREF](#)
20. Kather JN, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, et al. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Can* 2020;1:789-799.
[PUBMED](#) | [CROSSREF](#)
21. Park J, Jang BG, Kim YW, Park H, Kim BH, Kim MJ, et al. A prospective validation and observer performance study of a deep learning algorithm for pathologic diagnosis of gastric tumors in endoscopic biopsies. *Clin Cancer Res* 2021;27:719-728.
[PUBMED](#) | [CROSSREF](#)
22. Song Z, Zou S, Zhou W, Huang Y, Shao L, Yuan J, et al. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun* 2020;11:4294.
[PUBMED](#) | [CROSSREF](#)
23. Kanavati F, Tsuneki M. A deep learning model for gastric diffuse-type adenocarcinoma classification in whole slide images. *Sci Rep* 2021;11:20486.
[PUBMED](#) | [CROSSREF](#)
24. Ba W, Wang S, Shang M, Zhang Z, Wu H, Yu C, et al. Assessment of deep learning assistance for the pathological diagnosis of gastric cancer. *Mod Pathol* 2022;35:1262-1268.
[PUBMED](#) | [CROSSREF](#)
25. Matsushima J, Sato T, Ohnishi T, Yoshimura Y, Mizutani H, Koto S, et al. The use of deep learning-based computer diagnostic algorithm for detection of lymph node metastases of gastric adenocarcinoma. *Int J Surg Pathol* 2022;10668969221113475.
[PUBMED](#) | [CROSSREF](#)
26. Wang X, Chen Y, Gao Y, Zhang H, Guan Z, Dong Z, et al. Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nat Commun* 2021;12:1637.
[PUBMED](#) | [CROSSREF](#)
27. Hu Y, Su F, Dong K, Wang X, Zhao X, Jiang Y, et al. Deep learning system for lymph node quantification and metastatic cancer identification from whole-slide pathology images. *Gastric Cancer* 2021;24:868-877.
[PUBMED](#) | [CROSSREF](#)
28. Kim HN, Jang J, Heo YJ, Kim B, Jung H, Jang Y, et al. PD-L1 expression in gastric cancer determined by digital image analyses: pitfalls and correlation with pathologist interpretation. *Virchows Arch* 2020;476:243-250.
[PUBMED](#) | [CROSSREF](#)

29. Han Z, Lan J, Wang T, Hu Z, Huang Y, Deng Y, et al. A deep learning quantification algorithm for HER2 scoring of gastric cancer. *Front Neurosci* 2022;16:877229.
[PUBMED](#) | [CROSSREF](#)
30. Huang B, Tian S, Zhan N, Ma J, Huang Z, Zhang C, et al. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: a retrospective multicentre study. *EBioMedicine* 2021;73:103631.
[PUBMED](#) | [CROSSREF](#)
31. Wei T, Yuan X, Gao R, Johnston L, Zhou J, Wang Y, et al. Survival prediction of stomach cancer using expression data and deep learning models with histopathological images. *Cancer Sci* 2023;114:690-701.
[PUBMED](#) | [CROSSREF](#)
32. Muti HS, Heij LR, Keller G, Kohlruss M, Langer R, Dislich B, et al. Development and validation of deep learning classifiers to detect Epstein-Barr virus and microsatellite instability status in gastric cancer: a retrospective multicentre cohort study. *Lancet Digit Health* 2021;3:e654-e664.
[PUBMED](#) | [CROSSREF](#)
33. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019;25:1054-1056.
[PUBMED](#) | [CROSSREF](#)
34. Jeong Y, Cho CE, Kim JE, Lee J, Kim N, Jung WY, et al. Deep learning model to predict Epstein-Barr virus associated gastric cancer in histology. *Sci Rep* 2022;12:18466.
[PUBMED](#) | [CROSSREF](#)
35. Lee SH, Lee Y, Jang HJ. Deep learning captures selective features for discrimination of microsatellite instability from pathologic tissue slides of gastric cancer. *Int J Cancer* 2023;152:298-307.
[PUBMED](#) | [CROSSREF](#)
36. Amin MB, Edge SB, Greene FL, Byrd DR, Brookland RK, Washington MK, et al., *AJCC Cancer Staging Manual*. Berlin: Springer, 2017.
37. Steiner DF, MacDonald R, Liu Y, Truszkowski P, Hipp JD, Gammage C, et al. Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. *Am J Surg Pathol* 2018;42:1636-1646.
[PUBMED](#) | [CROSSREF](#)
38. Trojani M, Contesso G, Coindre JM, Rouesse J, Bui NB, de Mascarel A, et al. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. *Int J Cancer* 1984;33:37-42.
[PUBMED](#) | [CROSSREF](#)
39. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991;19:403-410.
[PUBMED](#) | [CROSSREF](#)
40. Kang YK, Kim KM, Sohn T, Choi D, Kang HJ, Ryu MH, et al. Clinical practice guideline for accurate diagnosis and effective treatment of gastrointestinal stromal tumor in Korea. *J Korean Med Sci* 2010;25:1543-1552.
[PUBMED](#) | [CROSSREF](#)
41. Yücel Z, Akal F, Oltulu P. Mitotic cell detection in histopathological images of neuroendocrine tumors using improved YOLOv5 by transformer mechanism. *Signal Image Video Process*. Forthcoming 2023.
[CROSSREF](#)
42. Pantanowitz L, Hartman D, Qi Y, Cho EY, Suh B, Paeng K, et al. Accuracy and efficiency of an artificial intelligence tool when counting breast mitoses. *Diagn Pathol* 2020;15:80.
[PUBMED](#) | [CROSSREF](#)
43. Zehra T, Anjum S, Mahmood T, Shams M, Sultan BA, Ahmad Z, et al. A novel deep learning-based mitosis recognition approach and dataset for uterine leiomyosarcoma histopathology. *Cancers (Basel)* 2022;14:3785.
[PUBMED](#) | [CROSSREF](#)
44. Choi S, Cho SI, Ma M, Park S, Pereira S, Aum BJ, et al. Artificial intelligence-powered programmed death ligand 1 analyser reduces interobserver variation in tumour proportion score for non-small cell lung cancer with better prediction of immunotherapy response. *Eur J Cancer* 2022;170:17-26.
[PUBMED](#) | [CROSSREF](#)
45. Baxi V, Lee G, Duan C, Pandya D, Cohen DN, Edwards R, et al. Association of artificial intelligence-powered and manual quantification of programmed death-ligand 1 (PD-L1) expression with outcomes in patients treated with nivolumab ± ipilimumab. *Mod Pathol* 2022;35:1529-1539.
[PUBMED](#) | [CROSSREF](#)
46. Wang X, Wang L, Bu H, Zhang N, Yue M, Jia Z, et al. How can artificial intelligence models assist PD-L1 expression scoring in breast cancer: results of multi-institutional ring studies. *NPJ Breast Cancer* 2021;7:61.
[PUBMED](#) | [CROSSREF](#)

47. Wu J, Liu C, Liu X, Sun W, Li L, Gao N, et al. Artificial intelligence-assisted system for precision diagnosis of PD-L1 expression in non-small cell lung cancer. *Mod Pathol* 2022;35:403-411.
[PUBMED](#) | [CROSSREF](#)
48. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 2012;12:252-264.
[PUBMED](#) | [CROSSREF](#)
49. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science* 2011;331:1565-1570.
[PUBMED](#) | [CROSSREF](#)
50. Shitara K, Özgüroğlu M, Bang YJ, Di Bartolomeo M, Mandalà M, Ryu MH, et al. Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial. *Lancet* 2018;392:123-133.
[PUBMED](#) | [CROSSREF](#)
51. Janjigian YY, Shitara K, Moehler M, Garrido M, Salman P, Shen L, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. *Lancet* 2021;398:27-40.
[PUBMED](#) | [CROSSREF](#)
52. Koelzer VH, Gisler A, Hanhart JC, Griss J, Wagner SN, Willi N, et al. Digital image analysis improves precision of PD-L1 scoring in cutaneous melanoma. *Histopathology* 2018;73:397-406.
[PUBMED](#) | [CROSSREF](#)
53. Widmaier M, Wiestler T, Walker J, Barker C, Scott ML, Sekhavati F, et al. Comparison of continuous measures across diagnostic PD-L1 assays in non-small cell lung cancer using automated image analysis. *Mod Pathol* 2020;33:380-390.
[PUBMED](#) | [CROSSREF](#)
54. Puladi B, Ooms M, Kintsler S, Houschyar KS, Steib F, Modabber A, et al. Automated PD-L1 scoring using artificial intelligence in head and neck squamous cell carcinoma. *Cancers (Basel)* 2021;13:4409.
[PUBMED](#) | [CROSSREF](#)
55. Park JH, Lee KS, Choi E, Jung W, Aum J, Pereira S, et al. Pathologic validation of artificial intelligence-powered prediction of combined positive score of PD-L1 immunohistochemistry in urothelial carcinoma. *J Clin Oncol* 2021;39:e16518.
[CROSSREF](#)
56. Gravalos C, Jimeno A. HER2 in gastric cancer: a new prognostic factor and a novel therapeutic target. *Ann Oncol* 2008;19:1523-1529.
[PUBMED](#) | [CROSSREF](#)
57. Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* 2010;376:687-697.
[PUBMED](#) | [CROSSREF](#)
58. Bartley AN, Washington MK, Colasacco C, Ventura CB, Ismaila N, Benson AB 3rd, et al. HER2 testing and clinical decision making in gastroesophageal adenocarcinoma: guideline from the College of American Pathologists, American Society for Clinical Pathology, and the American Society of Clinical Oncology. *J Clin Oncol* 2017;35:446-464.
[PUBMED](#) | [CROSSREF](#)
59. Wu S, Yue M, Zhang J, Li X, Li Z, Zhang H, et al. The role of artificial intelligence in accurate interpretation of HER2 immunohistochemical scores 0 and 1+ in breast cancer. *Mod Pathol* 2023;36:100054.
[PUBMED](#) | [CROSSREF](#)
60. Zakrzewski F, de Back W, Weigert M, Wenke T, Zeugner S, Mantey R, et al. Automated detection of the HER2 gene amplification status in Fluorescence in situ hybridization images for the diagnostics of cancer tissues. *Sci Rep* 2019;9:8231.
[PUBMED](#) | [CROSSREF](#)
61. Qaiser T, Mukherjee A, Reddy Pb C, Munugoti SD, Tallam V, Pitkäaho T, et al. HER2 challenge contest: a detailed assessment of automated HER2 scoring algorithms in whole slide images of breast cancer tissues. *Histopathology* 2018;72:227-238.
[PUBMED](#) | [CROSSREF](#)
62. Farahmand S, Fernandez AI, Ahmed FS, Rimm DL, Chuang JH, Reisenbichler E, et al. Deep learning trained on hematoxylin and eosin tumor region of Interest predicts HER2 status and trastuzumab treatment response in HER2+ breast cancer. *Mod Pathol* 2022;35:44-51.
[PUBMED](#) | [CROSSREF](#)

63. Tsai PC, Lee TH, Kuo KC, Su FY, Lee TM, Marostica E, et al. Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. *Nat Commun* 2023;14:2102.
[PUBMED](#) | [CROSSREF](#)
64. 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;72:2000-2013.
[PUBMED](#) | [CROSSREF](#)
65. Howard FM, Dolezal J, Kochanny S, Khramtsova G, Vickery J, Srisuwananukorn A, et al. Integration of clinical features and deep learning on pathology for the prediction of breast cancer recurrence assays and risk of recurrence. *NPJ Breast Cancer* 2023;9:25.
[PUBMED](#) | [CROSSREF](#)
66. Foersch S, Eckstein M, Wagner DC, Gach F, Woerl AC, Geiger J, et al. Deep learning for diagnosis and survival prediction in soft tissue sarcoma. *Ann Oncol* 2021;32:1178-1187.
[PUBMED](#) | [CROSSREF](#)
67. Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med* 2015;21:449-456.
[PUBMED](#) | [CROSSREF](#)
68. Sanjeevaiah A, Cheedella N, Hester C, Porembka MR. Gastric cancer: recent molecular classification advances, racial disparity, and management implications. *J Oncol Pract* 2018;14:217-224.
[PUBMED](#) | [CROSSREF](#)
69. Wang Y, Hu C, Kwok T, Bain CA, Xue X, Gasser RB, et al. DEMoS: a deep learning-based ensemble approach for predicting the molecular subtypes of gastric adenocarcinomas from histopathological images. *Bioinformatics* 2022;38:4206-4213.
[PUBMED](#) | [CROSSREF](#)
70. Flinner N, Gretser S, Quaas A, Bankov K, Stoll A, Heckmann LE, et al. Deep learning based on hematoxylin-eosin staining outperforms immunohistochemistry in predicting molecular subtypes of gastric adenocarcinoma. *J Pathol* 2022;257:218-226.
[PUBMED](#) | [CROSSREF](#)
71. Kim ST, Cristescu R, Bass AJ, Kim KM, Odegaard JI, Kim K, et al. Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. *Nat Med* 2018;24:1449-1458.
[PUBMED](#) | [CROSSREF](#)
72. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018;24:1559-1567.
[PUBMED](#) | [CROSSREF](#)
73. Chen RJ, Lu MY, Williamson DF, Chen TY, Lipkova J, Noor Z, et al. Pan-cancer integrative histology-genomic analysis via multimodal deep learning. *Cancer Cell* 2022;40:865-878.e6.
[PUBMED](#) | [CROSSREF](#)
74. Saldanha OL, Loeffler CM, Niehues JM, van Treeck M, Seraphin TP, Hewitt KJ, et al. Self-supervised attention-based deep learning for pan-cancer mutation prediction from histopathology. *NPJ Precis Oncol* 2023;7:35.
[PUBMED](#) | [CROSSREF](#)
75. Paijens ST, Vledder A, de Bruyn M, Nijman HW. Tumor-infiltrating lymphocytes in the immunotherapy era. *Cell Mol Immunol* 2021;18:842-859.
[PUBMED](#) | [CROSSREF](#)
76. Hegde PS, Karanikas V, Evers S. The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clin Cancer Res* 2016;22:1865-1874.
[PUBMED](#) | [CROSSREF](#)
77. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature* 2017;541:321-330.
[PUBMED](#) | [CROSSREF](#)
78. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 2018;554:544-548.
[PUBMED](#) | [CROSSREF](#)
79. Park S, Ock CY, Kim H, Pereira S, Park S, Ma M, et al. Artificial intelligence-powered spatial analysis of tumor-infiltrating lymphocytes as complementary biomarker for immune checkpoint inhibition in non-small-cell lung cancer. *J Clin Oncol* 2022;40:1916-1928.
[PUBMED](#) | [CROSSREF](#)
80. Oya Y, Hayakawa Y, Koike K. Tumor microenvironment in gastric cancers. *Cancer Sci* 2020;111:2696-2707.
[PUBMED](#) | [CROSSREF](#)

81. Kemi N, Ylitalo O, Väyrynen JP, Helminen O, Junttila A, Mrena J, et al. Tertiary lymphoid structures and gastric cancer prognosis. *APMIS* 2023;131:19-25.
[PUBMED](#) | [CROSSREF](#)
82. Ham IH, Lee D, Hur H. Role of cancer-associated fibroblast in gastric cancer progression and resistance to treatments. *J Oncol* 2019;2019:6270784.
[PUBMED](#) | [CROSSREF](#)
83. Zhou Q, Wu X, Wang X, Yu Z, Pan T, Li Z, et al. The reciprocal interaction between tumor cells and activated fibroblasts mediated by TNF- α /IL-33/ST2L signaling promotes gastric cancer metastasis. *Oncogene* 2020;39:1414-1428.
[PUBMED](#) | [CROSSREF](#)
84. Taylor-Weiner A, Pokkalla H, Han L, Jia C, Huss R, Chung C, et al. A machine learning approach enables quantitative measurement of liver histology and disease monitoring in NASH. *Hepatology* 2021;74:133-147.
[PUBMED](#) | [CROSSREF](#)
85. Yu Y, Wang J, Ng CW, Ma Y, Mo S, Fong EL, et al. Deep learning enables automated scoring of liver fibrosis stages. *Sci Rep* 2018;8:16016.
[PUBMED](#) | [CROSSREF](#)
86. Li Z, Jiang Y, Li B, Han Z, Shen J, Xia Y, et al. Development and validation of a machine learning model for detection and classification of tertiary lymphoid structures in gastrointestinal cancers. *JAMA Netw Open* 2023;6:e2252553.
[PUBMED](#) | [CROSSREF](#)
87. Xu H, Cong F, Hwang TH. Machine learning and artificial intelligence-driven spatial analysis of the tumor immune microenvironment in pathology slides. *Eur Urol Focus* 2021;7:706-709.
[PUBMED](#) | [CROSSREF](#)
88. Campanella G, Hanna MG, Geneslaw L, Mirafior A, Werneck Krauss Silva V, Busam KJ, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* 2019;25:1301-1309.
[PUBMED](#) | [CROSSREF](#)
89. Kann BH, Hosny A, Aerts HJ. Artificial intelligence for clinical oncology. *Cancer Cell* 2021;39:916-927.
[PUBMED](#) | [CROSSREF](#)
90. Ueyama H, Kato Y, Akazawa Y, Yatagai N, Komori H, Takeda T, et al. Application of artificial intelligence using a convolutional neural network for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. *J Gastroenterol Hepatol* 2021;36:482-489.
[PUBMED](#) | [CROSSREF](#)
91. Homeyer A, Lotz J, Schwen LO, Weiss N, Romberg D, Höfener H, et al. Artificial intelligence in pathology: from prototype to product. *J Pathol Inform* 2021;12:13.
[PUBMED](#) | [CROSSREF](#)
92. Tizhoosh HR, Pantanowitz L. Artificial intelligence and digital pathology: challenges and opportunities. *J Pathol Inform* 2018;9:38.
[PUBMED](#) | [CROSSREF](#)
93. Kim I, Kang K, Song Y, Kim TJ. Application of artificial intelligence in pathology: trends and challenges. *Diagnostics (Basel)* 2022;12:2794.
[PUBMED](#) | [CROSSREF](#)
94. Cheng JY, Abel JT, Balis UG, McClintock DS, Pantanowitz L. Challenges in the development, deployment, and regulation of artificial intelligence in anatomic pathology. *Am J Pathol* 2021;191:1684-1692.
[PUBMED](#) | [CROSSREF](#)
95. Carter SM, Rogers W, Win KT, Frazer H, Richards B, Houssami N. The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *Breast* 2020;49:25-32.
[PUBMED](#) | [CROSSREF](#)
96. Amann J, Vetter D, Blomberg SN, Christensen HC, Coffee M, Gerke S, et al. To explain or not to explain?—Artificial intelligence explainability in clinical decision support systems. *PLOS Digit Health* 2022;1:e0000016.
[PUBMED](#) | [CROSSREF](#)
97. Shmatko A, Ghaffari Laleh N, Gerstung M, Kather JN. Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. *Nat Can* 2022;3:1026-1038.
[PUBMED](#) | [CROSSREF](#)
98. Ayorinde JO, Citterio F, Landrò M, Peruzzo E, Islam T, Tilley S, et al. Artificial intelligence you can trust: what matters beyond performance when applying artificial intelligence to renal histopathology? *J Am Soc Nephrol* 2022;33:2133-2140.
[PUBMED](#) | [CROSSREF](#)

99. Stenzinger A, Alber M, Allgäuer M, Jurmeister P, Bockmayr M, Budczies J, et al. Artificial intelligence and pathology: from principles to practice and future applications in histomorphology and molecular profiling. *Semin Cancer Biol* 2022;84:129-143.
[PUBMED](#) | [CROSSREF](#)
100. Howard FM, Dolezal J, Kochanny S, Schulte J, Chen H, Heij L, et al. The impact of site-specific digital histology signatures on deep learning model accuracy and bias. *Nat Commun* 2021;12:4423.
[PUBMED](#) | [CROSSREF](#)