

Artificial Intelligence in *Gastroenterology*

Artif Intell Gastroenterol 2022 December 28; 3(5): 117-162



A

I

G

Artificial Intelligence in Gastroenterology

Contents

Bimonthly Volume 3 Number 5 December 28, 2022

REVIEW

- 117 Artificial intelligence in gastroenterology: A narrative review
Galati JS, Duve RJ, O'Mara M, Gross SA
- 142 Artificial intelligence applications in predicting the behavior of gastrointestinal cancers in pathology
Yavuz A, Alpsoy A, Gedik EO, Celik MY, Bassorgun CI, Unal B, Elpek GO

ABOUT COVER

Editorial Board Member of *Artificial Intelligence in Gastroenterology*, Alina Ioana Tantau, FACE, PhD, Academic Research, Chief Doctor, Professor, Senior Researcher, Department of Internal Medicine and Gastroenterology, "Iuliu Hatieganu" University of Medicine and Pharmacy, 4th Medical Clinic, Cluj-Napoca 400012, Romania. alitantau@gmail.com

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastroenterology (AIG, Artif Intell Gastroenterol)* is to provide scholars and readers from various fields of artificial intelligence in gastroenterology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIG mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastroenterology and covering a wide range of topics, including artificial intelligence in gastrointestinal cancer, liver cancer, pancreatic cancer, hepatitis B, hepatitis C, nonalcoholic fatty liver disease, inflammatory bowel disease, irritable bowel syndrome, and *Helicobacter pylori* infection.

INDEXING/ABSTRACTING

The *AIG* is now abstracted and indexed in Reference Citation Analysis, China Science and Technology Journal Database.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Si Zhao*; Production Department Director: *Xiang Li*; Editorial Office Director: *Jin-Li Wang*.

NAME OF JOURNAL

Artificial Intelligence in Gastroenterology

ISSN

ISSN 2644-3236 (online)

LAUNCH DATE

July 28, 2020

FREQUENCY

Bimonthly

EDITORS-IN-CHIEF

Rajvinder Singh, Ferruccio Bonino

EDITORIAL BOARD MEMBERS

<https://www.wjgnet.com/2644-3236/editorialboard.htm>

PUBLICATION DATE

December 28, 2022

COPYRIGHT

© 2022 Baishideng Publishing Group Inc

INSTRUCTIONS TO AUTHORS

<https://www.wjgnet.com/bpg/gerinfo/204>

GUIDELINES FOR ETHICS DOCUMENTS

<https://www.wjgnet.com/bpg/GerInfo/287>

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH

<https://www.wjgnet.com/bpg/gerinfo/240>

PUBLICATION ETHICS

<https://www.wjgnet.com/bpg/GerInfo/288>

PUBLICATION MISCONDUCT

<https://www.wjgnet.com/bpg/gerinfo/208>

ARTICLE PROCESSING CHARGE

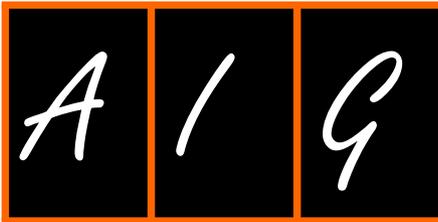
<https://www.wjgnet.com/bpg/gerinfo/242>

STEPS FOR SUBMITTING MANUSCRIPTS

<https://www.wjgnet.com/bpg/GerInfo/239>

ONLINE SUBMISSION

<https://www.f6publishing.com>



Artificial intelligence applications in predicting the behavior of gastrointestinal cancers in pathology

Aysen Yavuz, Anil Alpsoy, Elif Ocak Gedik, Mennan Yigitcan Celik, Cumhuriyet Ibrahim Bassorgun, Betul Unal, Gulsum Ozlem Elpek

Specialty type: Pathology

Provenance and peer review:

Invited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): 0
Grade D (Fair): 0
Grade E (Poor): 0

P-Reviewer: Tsoulfas G, Greece;
Wan XH, China

Received: October 16, 2022

Peer-review started: October 16, 2022

First decision: November 15, 2022

Revised: November 25, 2022

Accepted: December 13, 2022

Article in press: December 13, 2022

Published online: December 28, 2022



Aysen Yavuz, Anil Alpsoy, Elif Ocak Gedik, Mennan Yigitcan Celik, Cumhuriyet Ibrahim Bassorgun, Betul Unal, Gulsum Ozlem Elpek, Department of Pathology, Akdeniz University Medical School, Antalya 07070, Turkey

Corresponding author: Gulsum Ozlem Elpek, MD, Professor, Department of Pathology, Akdeniz University Medical School, Dumlupinar bulvarı, Antalya 07070, Turkey.
elpek@akdeniz.edu.tr

Abstract

Recent research has provided a wealth of data supporting the application of artificial intelligence (AI)-based applications in routine pathology practice. Indeed, it is clear that these methods can significantly support an accurate and rapid diagnosis by eliminating errors, increasing reliability, and improving workflow. In addition, the effectiveness of AI in the pathological evaluation of prognostic parameters associated with behavior, course, and treatment in many types of tumors has also been noted. Regarding gastrointestinal system (GIS) cancers, the contribution of AI methods to pathological diagnosis has been investigated in many studies. On the other hand, studies focusing on AI applications in evaluating parameters to determine tumor behavior are relatively few. For this purpose, the potential of AI models has been studied over a broad spectrum, from tumor subtyping to the identification of new digital biomarkers. The capacity of AI to infer genetic alterations of cancer tissues from digital slides has been demonstrated. Although current data suggest the merit of AI-based approaches in assessing tumor behavior in GIS cancers, a wide range of challenges still need to be solved, from laboratory infrastructure to improving the robustness of algorithms, before incorporating AI applications into real-life GIS pathology practice. This review aims to present data from AI applications in evaluating pathological parameters related to the behavior of GIS cancer with an overview of the opportunities and challenges encountered in implementing AI in pathology.

Key Words: Digital pathology; Colorectal cancer; Gastric cancer; Machine learning; Deep learning; Prognosis

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: This review outlines the potential of artificial intelligence applications for evaluating pathological parameters related to the behavior of gastrointestinal cancers. The role of these methods in determining the behavior of esophageal cancers remains to be investigated. On the other hand, the results are promising, supporting that these models can assist in the determination of conventional pathological parameters and perform molecular subtyping in gastric and colorectal cancers. Furthermore, these applications encourage digital prognostic biomarker discovery by revealing predictions that are impossible when using traditional visual methods. However, further studies are needed to overcome the obstacles to implementing these applications into pathology practice.

Citation: Yavuz A, Alpsoy A, Gedik EO, Celik MY, Bassorgun CI, Unal B, Elpek GO. Artificial intelligence applications in predicting the behavior of gastrointestinal cancers in pathology. *Artif Intell Gastroenterol* 2022; 3(5): 142-162

URL: <https://www.wjgnet.com/2644-3236/full/v3/i5/142.htm>

DOI: <https://dx.doi.org/10.35712/aig.v3.i5.142>

INTRODUCTION

Gastrointestinal (GIS) cancers, including tumors of the esophagus, stomach, colon, and rectum, are an important health problem worldwide. Although the incidence of esophageal cancer (EC) is relatively low, gastric cancer (GC) and colorectal cancer (CRC) are among the most common types of cancer (fifth and third, respectively)[1]. They are also responsible for a substantial proportion of cancer mortality, with GC being the third and CRC the second most common cause of cancer-related death[2]. Although various predictive and prognostic parameters are currently available, the mortality rates for patients with GIS cancer are, unfortunately, still very high[2]. It has been shown that rectifying this situation may depend on paving the way for more personalized treatment strategies that lead to a better prognosis and/or fewer treatment side effects[3,4]. Therefore, the meticulous and complete evaluation of patients to determine the appropriate treatment is critical.

In this context, in addition to providing a definitive diagnosis, the role of an accurate evaluation of pathological parameters related to the behavior and proper treatment of GIS tumors cannot be ignored. However, pathology, a morphology-based specialty, is susceptible to subjectivity regarding intraobserver and interobserver variations, particularly in oncology. That is why, in recent years, the search for more objective criteria to eliminate bias, as well as to reduce the growing workload and to contribute time-saving, has allowed the improvement of image analysis-based digital pathology (DP), which has an important place in modern pathological applications[5,6].

In particular, significant advances in slide scanner technology, which can rapidly digitize all pathological slides at high resolution whole slide images (WSIs), has enabled not only the analysis of a wide range of morphological parameters but also the detection of biomarkers/genetic changes in many types of tumors[7-9]. The ability of computer-based analysis to detect prognostic and predictive markers from these images, depending on the fact that they are composed of number matrices containing a large amount of information that is not accessible to the human eye, has led to the adoption of artificial intelligence (AI) for DP[10,11]. Accordingly, the number of studies on AI applications associated with the diagnosis, follow-up, and treatment of many tumors has increased significantly over time. Regarding GIS, data from previous studies evaluating pathological prognostic parameters with various AI models suggest that using these methods may be beneficial. Unfortunately, these encouraging results have not overcome the wide range of challenges to be solved, from laboratory infrastructure to improving the robustness of algorithms, before incorporating AI applications into real-life pathology practice.

This review presented the applications of AI in the evaluation of pathological parameters related to the behavior of GIS cancer, along with a brief overview of the opportunities and challenges encountered in its implementation in pathology.

GENERAL VIEW OF AI IN PATHOLOGY LABORATORIES

In parallel with technological developments, the evolution of whole slide imaging (WSI) has provided remote diagnosis, consultation, and education[12-14]. In the recent past, it was suggested that the use of WSI is comparable to, or even better than, conventional microscopic examination for decision-making in pathology[15-17]. On the other hand, WSIs are also crucial in applying AI methods in pathological practice. They not only provide quick access to the archive without loss of image quality, but they can also render gigabit images, which are very difficult to process, suitable for processing by "tessellation"

[18]. This preprocessing is based on cutting a large image into nonoverlapping smaller patches called "tiles," making them amenable to computational analysis. It should be noted that although some pathological studies use selected images captured manually with a camera, WSI is currently recommended as a standard for AI applications, especially in tumors where heterogeneity is frequent, such as those of the GIS[19].

To achieve reliable results with WSIs, many steps, from preserving the structure of the tissue to the preparation of sections, must be carried out with care in the pathological laboratory. In particular, it is imperative to evaluate and check slides for artifacts (tears, floating contamination, thickness) that have the potential to adversely affect digitization and, thus, AI applications[20,21]. However, it should be noted that even with optimal protocols and slide scanner standardization, the importance of color normalization to ensure consistency in WSI databases should not be overlooked, as it can affect the robustness of deep learning (DL) models. Accordingly, histogram-matching color transfer and spectral matching methods can be applied[22-24]. However, as these methods depend on the expertise of pathologists and are impractical for manual adjustment, various algorithms have been proposed by researchers capable of performing this normalization. Although promising results have been obtained, there is a need for future studies on the performance of AI models using color normalization systems[25, 26].

The gradual evolution of traditional pathology into DP has led to the development of powerful and user-friendly WSI analysis software tools with the ability to manage substantial WSIs and metadata from different hardware manufacturers, as well as interactive drawing annotation capabilities to facilitate decision-making and reporting. Moreover, a significant proportion of them is freely available [27-29]. In addition, the high costs of hardware required for high-performance computation in software development have become more affordable, leading to the implementation of DP in major medical centers[16,30-32]. Increasing the number of centers capable of using DP will allow for the generation of large and high-quality WSI databases, enabling the acquisition of large datasets and the design of algorithms for AI. However, the requirement of a significant investment is still an obstacle to overcome for the widespread application of these technologies[33]. In addition, the problem of proprietary datasets persists, limiting the repeatability of the proposed methodologies and hindering advancement in this field.

As mentioned above, the ability of AI to extract meaningful information from images that the naked human eye cannot discriminate makes it an attractive tool in the field of image processing and analysis in pathology. Therefore, contemporary AI models have evolved from expert systems to different types, such as machine learning (ML) and DL (Table 1). In brief, ML is a subtype of AI that provides a computer system to automatically learn and develop from datasets on its own and solve problems without explicit programming[34-36]. DL is a subfield of ML that employs sophisticated algorithmic structures inspired by the neural network of the human brain (artificial neural network, ANN) in which statistical models are established from input training data[37-39]. Therefore, DL requires large, annotated datasets to develop its algorithms. At present, the annotation of datasets is a complex task in model development[9,40]. In practice, the time-consuming and challenging nature of annotation, especially in systems where heterogeneous lesions are common, such as GIS, may affect the accuracy of the model being trained[41]. Another limitation is that the dataset obtained by a study group does not show the same performance when compared to external validation sets from other institutions. Recently, studies have been conducted to overcome the hindering properties of annotation[42-44]. It has also been suggested that the adoption of DP for diagnosis could indirectly facilitate the generation of valuable datasets for future algorithm development by enabling pathologists to describe areas of interest during evaluation and reporting[45].

It has often been emphasized that the validation of AI-based technologies requires an evidence-based approach[42,46]. This should also be considered in a laboratory-based medical specialty such as pathology. On the other hand, analyzing the performance of AI techniques to that of pathologists is a significant challenge regarding interobserver and interobserver heterogeneity. Currently, the problems related to establishing "ground truth" in AI methods should not be overlooked[40,47]. It should be noted that this requires repeated testing of the effectiveness and consistency of AI applications in many different patient populations. The relative lack of a validation cohort in developing AI-powered DP applications is also related to the possible drawbacks of sharing histopathological slides. Despite interobserver heterogeneity and variability in pathological assessment also demonstrating the uncertainty of "ground truth" in this regard, multi center assessments involving multiple pathologists and datasets may be the best way to overcome this obstacle.

Before the integration of AI into the pathology workflow, the need to validate its benefits and address ethical recommendations increases the importance of AI-based tools being transparent and interpretable, resulting in an increasing demand for more explainable AI models. In this respect, there is a dilemma about the application of AI. Because most algorithms developed use DL, ensemble methods called "black box" models to tackle multidimensional problems are very complex. However, more straightforward methods that are not complex are not powerful enough to achieve the expected results [48]. For this reason, model interpretability, ethical concerns, and potential regulatory barriers should also be considered in newly developed AI tools to meet these expectations.

Table 1 General features of machine learning methods in the development of artificial intelligence models in gastrointestinal pathology

| AI models | Strengths | Weaknesses |
|-----------------------------|---|---|
| ML, Traditional, Supervised | Data output can be produced from the previously labeled training set | Labeling big data takes a considerable amount of time and can be challenging |
| | Allows users to reflect domain knowledge features | Feature extraction quality significantly affects the accuracy |
| ML, Traditional, Supervised | Users do not supervise the model or label any data | Input data is unknown and not labeled |
| | Patterns are detected automatically | Precise information related to data sorting is not provided |
| | Save time | Interpretation is challenging |
| SVM | Suitable for more efficient regression and classification analysis with high-dimensional data | Not suitable for large data sets. Requires more time for training; Low performance in overlapping classes |
| CNN | No labeling is required for important information and features | Lack of interpretability due to black boxes |
| | The performance capacity in image recognition is high | |
| FCN | Provides computational speed | A large amount of labeled data for training is required |
| | The background noise is automatically eliminated | The labeling cost is high |
| RNN | Able to decide which information to remember from past experiences | The model is hard to train |
| | A suitable deep learning model for sequential data | The computational cost is high |
| MIL | A detailed annotation is not required | A large amount of training data is required |
| | Suitable to be performed on large datasets | The computational cost is high |
| GAN | The potential to produce new realistic data that resembles the original data | The model is hard to train |

AI: Artificial intelligence; ML: Machine learning; SVM: Support Vector Machine; CNN: Convolutional neural networks; FCN: Fully convolutional neural networks; RNN: Recurrent neural networks; MIL: Multi-instance learning; GAN: Generative adversarial networks.

AI IN THE PATHOLOGICAL DETERMINATION OF PRENEOPLASTIC LESIONS IN GIS

Barrett's esophagus

The majority of AI studies in EC consist of imaging studies. In pathology, there have been recent studies on the diagnosis of Barrett's esophagus (BE) and the evaluation of dysplasia in these lesions to predict the risk of EC[49,50]. A proposed attention-based deep NN framework for detecting BE and adenocarcinoma (ADC) was found to be reliable with a mean accuracy of 0.83[49]. Unlike existing methods based on the region of interest, this model is based on tissue-level annotations, suggesting that it may provide a new approach for applying DL in pathology. On the other hand, the fact that the study was performed in a single center and on a relatively small data set necessitates the development of the proposed model with further studies. Since trefoil factor 3 expression is the key finding of BE, a DL model (VGG16) using immunohistochemically stained sections showed significant adaptability, with an area under the curve (AUC) of 0.88[50]. Although the proposed approach reduced the pathologist workload by 57%, the underlying ML model still needs further optimization.

Colorectal polyp classification

In CRC, unlike GC, the classification of polyps is an important task to determine the risk of CRC and the future surveillance needs of patients[51]. In routine examinations, high-risk polyps are evaluated based on their histopathological features with considerable interobserver variability among pathologists[52, 53]. However, a precise diagnosis of high-risk polyps is required for efficient and early detection of cancer. In addition, the recommendation for endoscopic screening of these lesions for an early diagnosis of CRC, especially in elderly individuals, increases the workload of daily pathology practice[54].

Therefore, AI applications have been developed to classify high-risk colorectal polyps and/or adenomas with high-grade dysplasia. In studies on the classification of these lesions and the identification of CRC, datasets of three to six specific categories and five models were used[55-62] (Table 2). Although most studies showed good performance with generally high AUCs and accuracies, because of the following restrictions, the evidence level of each model needed to be improved. The number of patches and WSIs that make up the datasets are different. Accordingly, in some studies, the number of datasets may affect the reliability of the results. In various studies, the annotation process is not delineated in detail. In addition, the fact that each model has a different focus and characteristics makes

Table 2 AI-based applications in pathology for the determination of tumor behavior in colorectal carcinomas

| Ref. | Task | Data sets | Algorithm/Model | Performance | Comments |
|------------------------------------|-----------------------------|---|-----------------|--|---|
| Xu <i>et al</i> [55] | NL/ADC/MC/SC/PC/CCTA | 717 patches | AlexNet | Accuracy: 97% | The model provides the classifications of tumor subtypes |
| Korbar <i>et al</i> [56] | NL/HP/SSP/TSA/TA/TVA-VA | Training set: 458 WSIs; Test set: 239 WSIs | ResNET | F1 Score: 88.8%; Accuracy: 93%; Precision: 89.7%; Recall: 88.3% | The model may reduce the workload of pathologists in the assessment of colorectal polyps |
| Haj-Hassan <i>et al</i> [57] | NL/AD/ADC | 30 patients, Multispectral image patches | CNN | Accuracy: 99.2% | CNN allows the classification of CRC tissue types using pre-segmented regions of interest |
| Ponzio <i>et al</i> [58] | NL/AD/ADC | 27 WSIs | VGG16 | Accuracy: 96% | TL considerably outperforms the CNN fully trained on CRC samples on the same test dataset |
| Sena <i>et al</i> [59] | NL/HP/AD/ADC | 393 images | CNN | Accuracy: 80% | DL may provide a valuable tool to assist pathologists in the histological classification of CR tumors |
| Iizuka <i>et al</i> [60] | NL/AD/ADC | 4036 WSIs + 500WSIs | CNN/RNN | AUCs: 0.96-0.99 | Integrating DL models in pathology workflow would be of high benefit for easing the workload of pathologists |
| Wei <i>et al</i> [61] | NL//TA/TVA/VA/HP | 1182 WSIs | ResNet | Accuracy: 93.5% (Internal test set); Accuracy: 87% (External test set) | This model may assist pathologists by improving the accuracy of CRC screening |
| Awan <i>et al</i> [62] | NL/Low GR/High GR | 139 images | CNN | Accuracy: 97% (two-class), 91% (three-class) | The model provides the classifications of tumor subtypes based on the shape of glands |
| Sirinukunwattana <i>et al</i> [97] | Prediction of MSTs | 510 WSIs (FOCUS), 431 WSIs (TCGA), 265 WSIs (GRAMPIAN cohort) | Inception V3 | AUCs: 0.9 (FOCUS); 0.94 (TCGA), 0.85 (GRAMPIAN cohort) | RNA expression classifiers can predict from H-E stained images, opening the door to cheap and reliable biological stratification within routine workflows |
| Echle <i>et al</i> [98] | MSI vs MSS | 6406 WSIs (Training); 771 WSIs (External validation) | ShuffleNet | AUC: 0.92 (Training); AUC: 0.96 (External validation) | The model provides a low-cost evaluation of MSI without molecular testing |
| Kather <i>et al</i> [80] | MSI vs MSS | 60894 patches (TCGA-CRC-KR); 93408 patches (TCGA-CRC-DX) | ResNet18 | AUC: 0.84 (TCGA-CRC-KR); AUC: 0.77 (TCGA-CRC-DX) | This method may lead to improvements in molecular subtype screening workload in pathology |
| Kather <i>et al</i> [77] | Prediction of molecular Als | 426 patients (TCGA-CRC); 379 patients (DACHS) | ShuffleNet | AUROC: 0.76 | The algorithm predicts a wide range of molecular alterations from routine, H-E stained slides |
| Kruger <i>et al</i> [99] | Prediction of MSTs | 919 WSIs | ResNet 34 | AUCs: Mean: 0.87; CMS1: 0.85; CMS2: 0.92, CMS3: 0.85; CMS4: 0.86 | The MIL framework can identify morphological features indicative of different molecular subtypes |
| Popovici <i>et al</i> [100] | Prediction of MSTs | 300 WSIs | VGG-F | Accuracy: 0.84; | The image-based |

| | | | | | |
|-------------------------------------|---|---|------------------------------|--|--|
| | | | | Recall: 0.85; Precision: 0.84 | classifier shows a significant prognostic value similar to the molecular counterparts |
| Cao <i>et al</i> [101] | MSI vs MSS | 429 patients (TCGA-COAD); 785 patients (Asian-CRC) | EPLA | AUC: 0.88 (TCGA-COAD); AUC: 0.85 (Asian-CRC) | This pathomics-based model provides MSI estimation directly from images without molecular testing |
| Bilal <i>et al</i> [102] | Prediction of molecular Als | 502 slides (TCGA-CRC-DX); 47 slides (PAIP) | ResNet18, ResNet34, HoVerNet | AUROC: HM (0.81 vs 0.71); MSI (0.86 vs 0.74); CIN (0.83 vs 0.73), BRAFmut (0.79 vs 0.66), TP53mut (0 vs 0.64), KRASmut (0.60), CIMP (0.79) | This algorithm is based on non-annotated images and uses only slide-level labels to predict the status of CRC pathways and mutations |
| Kwak <i>et al</i> [110] | LNM prediction | 164 patients | CNN, U-Net | AUROC: 67% | PTS score is a potential prognostic parameter for LNM in CRC |
| Pai <i>et al</i> [111] | LNM prediction | 230 patients (training), (136 testing) | CNN | AUROC: 79% | The model allows to identify and quantify a broad spectrum of histological features, including LNM in CRC |
| Kiehl <i>et al</i> [112] | LNM prediction | 3013 patients | ResNET18 | AUROC: 74.1% | DL-based analysis may help predict the LNM of patients with CRC using routine HE-stained slides |
| Weis <i>et al</i> [120] | Tumor Budding (Pan-CK) | 381 patients | CNN | Spatial clusters of tumor buds correlates to N status (P: 0.003) | The model is a feasible and valid assessment tool for tumor budding on WSIs and can predict prognosis |
| Kather <i>et al</i> [121] | ADI, DEB, LYM, MUC, SM | 86 slides (Training), 25 slides (Testing); 862 slide (TCGA-COAD) | VGG19 | AUC: 98.7% HR: 2.29 (OS); 1.92 (RFS); Deep stroma score HR: 1.99 (P: 0.002), Shorter OS | This model can assess the human TME and predict prognosis directly from histopathological images |
| Shapcott <i>et al</i> [122] | TME (EC/IC/FC/MC) | 853 patches, 142 images (TCGA-COAD) | CNN | Accuracy: 76% (detection), 65% (classification) | The model provides the assessment of TME in CRC slides |
| Sirinukunwattana <i>et al</i> [123] | a-4 tissues classes; b- prediction of DM | 102 cases | Spatially Constrained CNN | a-AUROC: 90.4-99.9%; b-AUROC: 58.6-63.8% | The algorithm provides a digital marker for estimating the risk of DM |
| Swiderska-Chadaj <i>et al</i> [124] | TME Detection of ICs | 28 WSIs | FCN/LSM/U-Net | F1-score of 0.80; Sensitivity: 74%; Precision: 86% | DL approaches are reliable for automatically detecting lymphocytes in IHC-stained CRC tissue sections |
| Geessink <i>et al</i> [115] | TSR | 129 slides | CNN | HR: 2.48 (DSS); 2.05 (DFS) | CNN defined TSR as an independent prognosticator |
| Zhao <i>et al</i> [125] | TSR | 499 patients (Discovery cohort); 315 patients (Validation cohort) | CNN | TSR, independent prognostic parameter. HRs: 2.48 (Discovery cohort); 2.08 (Validation cohort) | CNN allows objective evaluation of TSR |
| Zhao <i>et al</i> [126] | Mucus tumor ratio low vs mucus tumor ratio high | 814 patients | CNN | HRs: 1.88 (Discovery cohort); 2.09 (Validation cohort) | The DL quantified mucus tumor ratio is an independent prognostic factor in CRC |
| Bychkov <i>et al</i> [132] | Prognosis LR vs HR | 420 TMA | VGG-16 | HR: 2.3 | The model extracts |

| | | | | | |
|---------------------------|---------------------------------------|--|-------------------|---|---|
| | | | | | more prognostic information from the tissue morphology than the experienced human observer |
| Skrede <i>et al</i> [133] | Prognosis (CSS) | 1122 patients (Validation cohort) | DoMorev1 | HRs: 1.89 (uncertain vs good); 3.84 (poor vs good) | The digital marker has the potential to identify patients at LR and HR and provides the selection of treatment |
| Jiang <i>et al</i> [134] | a-HRR vs LRR b-Poor vs good prognosis | 101 patients (Traning); 67 patients (Validation); 47 (TCGA-COAD) | InceptionResNetV2 | a-HRs: 8.98 (training); 10.69 (other 2 test groups); b-HRs: 10.687 (training); 5.03 (other 2 test groups) | The selected model offers an independent prognostic predictor which allows stratification of stage III CRC into risk groups |

NL: Normal; ADC: Adenocarcinoma; MC: Mucinous carcinoma; SC: Serrated carcinoma; PC: Papillary carcinoma; CCTA: Cribriform comedo-type adenocarcinoma; CRC: Colorectal cancer; HP: Hyperplastic polyp; SSP: Sessile serrated polyp; TSA: Traditional serrated adenoma; TA: Tubular adenoma; TVA: Tubulovillous adenoma; VA: Villous adenoma; AD: Adenoma; CNN: Convolutional neural networks; WSIs: Whole slide images; RNN: Recurrent neural networks; AUC: Area under the curve; GR: Grade; MSI: Microsatellite instable; MSS: Microsatellite stable; TCGA-CRC: Tumor Cancer Genome Atlas-Colorectal cancer; KR: Frozen tissues; DX: Formalin fixed paraffin embedded tissues; COAD: Colon adenocarcinoma; CRC: Colorectal carcinoma; EPLA: Ensemble patch likelihood aggregation; MST: Molecular subtype; CMS1: Tumor with MSI; CMS2: Tumors exhibiting epithelial gene expression, activated WNT and MYC signaling; CMS3: Tumors with metabolic dysregulations; CMS4: Tumors that possess TGF-β; MIL: Multi instance learning; Als: Alterations; AUROC: Area under the receiver operating characteristics; PAIP: Pathology artificial intelligence platform; HM: Hypermutation; CIN: Chromosomally unstable; CIMP: CpG island methylator phenotype; CK: Cytokeratin; ML: Machine learning; ADI: Adipocyte; DEB: Debris; LYM: Lymphocytes; MUC: Mucus; SM: Smooth muscle; HR: Hazard ratio; OS: Overall survival; RFS: Recurrence free-survival; TME: Tumor microenvironment; ICs: Immune cells; FCN: Fully convolutional network; LSM: Liquid state machine; IHC: Immunohistochemistry; EC: Epithelial cell; FC: Fibroblast; MC: Miscellaneous; TSR: Tumor stroma ratio; LR: Low risk; HR: High risk; TMA: Tissue microarray; CSS: Cancer specific survival; HRR: High recurrence risk; LRR: Low recurrence risk; DM: Distant metastasis; LNM: Lymph node metastasis; PTS: The predictive value of the peritumoral stroma score.

their comparison across studies impossible. One of the most striking examples of these studies is Korbar *et al*[56], where a DL model (ResNet-152) trained with over 400 WSIs showed a high overall accuracy in subtyping polyps. In another study, Wei *et al*[61], who ensembled five layers of ResNet, could classify these lesions with WSIs from a single institution, even in external datasets with a performance comparable to that of histopathological evaluation. This data indicates that further manual annotations by various qualified GI pathologists may be required to decrease classification problems in future AI systems for colorectal polyp detection.

AI IN THE PATHOLOGICAL DETERMINATION OF TUMOR BEHAVIOR IN GIS

In this section relevant data on GC and CRC will be discussed. Unfortunately, no AI studies have identified the parameters that are important in determining tumor behavior and survival in EC. Similarly, studies of EC concerning molecular characterization have not been found. Therefore, in EC, a tumor with extremely high mortality, it is clear that additional pathology studies are necessary to reveal the effectiveness of AI applications in predicting tumor behavior.

TUMOR SUBTYPING

Gastric cancer

Although nearly all GC are ADC, the clinicopathological features and behaviors show considerable variation depending on the histopathological diversity of tumor cells[63,64]. In recent years, it has been reported that the survival of patients with GC at the same stage differs significantly among the different subtypes. Therefore, accurate histopathological classification is critical in determining their prognosis, monitoring, and treatment.

GC is often classified based on the ADC differentiation grade, including well-differentiated ADC and poorly differentiated ADC. The grading depends on the presence or absence of glandular structure formation. ADCs are divided into intestinal and diffuse subtypes based on the Lauren classification[65]. While the diffuse form comprises a poorly differentiated type and signet ring cell carcinoma (SRCC), the intestinal type exhibits glands with papillae, tubules, or solid regions. Diffuse-type carcinomas are commonly confused with other nonneoplastic diseases. Because they usually consist of solitary dispersed cells in a desmoplastic stroma and inflammation.

In most of the reported studies, the adenocarcinoma differentiation grade is judged through manual identification by pathologists. Although there have been many studies on AI applications in the pathological diagnosis of GC in the recent past, there are few studies regarding tumor subclassification (Table 3). Yasuda *et al*[66] investigated the features and classification of GC tissues by using supervised ML algorithms. The results showed that this method reliably identifies morphological changes in tumors with different grades. Interestingly, PD-L1 expression levels have been found to serve as a morphological classification in hematoxylin and eosin (HE)-stained slides and correlate with histological grades. Therefore, quantitative analyses of tissue morphology may reveal molecular alterations in malignancies, and molecular analyses may aid in the pathological evaluation of cancer tissues. In another study, four different DL models were used to classify GC into diffuse ADC *vs* other ADC subtypes[67]. From biopsy WSIs, the trained model performed well at identifying both poorly differentiated ADC and SRCC cells. The authors pointed out that while higher magnification can reduce the false positive rate in classification, applying an RNN model with a more comprehensive dataset yields good results even at low magnifications. Hybrid models such as StoHisNet have also distinguished tubular, mucinous, and papillary subtypes of GC. This model showed a higher performance for multiclassification of pathological images of GC than other CNN-based models[68]. Although the model performed well in the four classifications of gastric pathological images, the study group does not include SRCC and other types. Also, the inability of the supervised network in the study to use unlabeled data and the lack of information on which combination maximizes the performance of the model performance warrant further studies. More recently, Su *et al*[69] demonstrated that DL models constructed using a pre-trained ResNet-18 model based on ImageNet27 achieved tumor differentiation recognition or poorly differentiated ADC and well-differentiated ADC classes, respectively. Although these results suggest that AI may be useful in GC classification, the scarcity of data and the differences in classification parameters used in these studies make it difficult to come to any solid conclusions.

Recently, GC has also been classified by the Tumor Cancer Genome Atlas (TCGA) into four molecular subtypes that are also included in the latest World Health Organization classification: Epstein-Barr virus (EBV)-positive (9%), microsatellite unstable (MSI) (22%), genomically stable (GS) (19%) and chromosomally unstable (CIN) (50%)[70,71]. The clinical significance of this classification comes from the fact that various factors, such as the prognosis and treatment response, differ among these subtypes [72,73]. In particular, among all subclasses of GC, tumors with MSI and positive EBV are associated with a better response to immunotherapy[72]. Consequently, recognizing these subtypes is crucial for categorizing patients who benefit from these treatments. Nevertheless, such classification requires the application of costly techniques, such as immunohistochemistry, and molecular testing, such as polymerase chain reaction, into pathological practice.

On the other hand, these two types have known characteristic histopathological findings. While EBV-positive GCs show prominent infiltration of lymphocytes into the neoplastic epithelium and the stroma, MSI subtype shows significant lymphocytic infiltration, intestinal-type histology, and expanding growth characteristics[63,74,75]. Therefore, these morphological features could be used to make predictions about the molecular subtype. In recent years, it has been suggested that molecular findings can be detected with AI *via* WSIs from HE-stained sections produced for pathological assessment[76-78]. Various models have been applied for molecular subtyping of GIS cancers. However, most of these studies have been conducted on CRCs (see below), whereas relatively few studies are available for GC (Table 3). For the detection of GC subtypes, Muti *et al*[79] demonstrated that DL could detect MSI and EBV positivity independently from each other in GC directly from HE-stained tissues in multi center pooled cohorts. They observed a high classification performance for the detection of MSI and EBV status. The relatively limited number of cases with positive findings and the fact that the ground truth methods for MSI were developed in CRC are presented as potential limitations of this study. On the other hand, their findings align with previous observations[69,80,81]. In addition, large-scale and multicenter validation broadens their work, which has considerable potential for integration into clinical procedures, suggesting that the application of DL could be a substitute for molecular techniques in the classification of GC. Furthermore, because these two subtypes share common morphological features and they are immunotherapy-sensitive tumors, Hinata *et al*[82] combined MSI and EBV in DL models and found they had a higher detection accuracy. This finding has been interpreted based on the possibility that these subtypes have similar distinctive pathological features, such as abundant stromal lymphocytic infiltration and intraepithelial lymphocytosis. On the other hand, the use of tissue microarray and manual labeling of tumor regions for TCGA presented as sources of bias compared to whole tissue slides, given the heterogeneity of tumor tissue. It was also emphasized that manual annotation by a pathologist might be a challenge to overcome by some weakly supervised methods (for example, attention-based deep multi instance learning) in the field of DL for the broad application of the proposed model.

Recently, a DL model called EBVNet that assists pathologists in predicting EBV from HE-stained slides has been introduced in GC[83]. The results suggested that human-machine fusion dramatically enhances the diagnostic ability of both EBVNet and the pathologist. However, this study has some limitations regarding its retrospective evaluation of training and validation. Additionally, the logistic regression model applied in the assessment is still an indirect way to interpret the model. More importantly, as in many DL models, the EBVNet decision-making procedure by the neural network is

Table 3 Artificial intelligence-based applications in pathology for the determination of tumor behavior in gastric cancer

| Ref. | Task | Data sets | Algorithm/Model | Performance | Comments |
|----------------------------|---|---|---|---|---|
| Yasuda <i>et al</i> [66] | NC, GR1, GR2, GR3; PDL-1, ATF7IP/MCAF1 | 66 WSIs | SV, ML, wndchrm | AUCs: 0.98-0.99 | The model allows grading emphasizing a correlation between molecular expression and tissue structures |
| Kanavati <i>et al</i> [67] | NC, ADC-D, ADC-O | 1-stage training: 1950 WSIs, 2-stage training: 874 WSIs | CNN and RNN | AUCs: 0.95-0.99 | The tool can aid pathologists by potentially accelerating their diagnostic workflow |
| Fu <i>et al</i> [68] | NC, TC, MC, PC | Training 2938 WSIs, Testing 980 WSIs | StoHisNet | The accuracy: 94.69%, F1 score: 94.96%, Recall: 94.95%, Precision: 94.97% | The model has high performance in the multi-classification on gastric images and shows strong generalization ability on other pathological datasets |
| Su <i>et al</i> [69] | NC, WD, PD, MSS vs MSI | GR: Training 348 WSIs, Testing 88 WSIs MSS: Training 212 WSIs, Testing: 52 WSIs, MSI: Training 136 WSIs, Testing: 36 WSIs | ResNet-18 | PD vs WD, F1 score: 0.8615, PD vs WD vs NC, F1 score: 0.8977; MSI vs MSS accuracy: 0.7727 | The proposed system integrated the tumor GR and MSI status recognition problems into the same workflow and was suitable for exploring the relationships between pathological features and molecular status |
| Muti <i>et al</i> [79] | MSI vs MSS; EBV (+) vs EBV (-) | 2823 patients with known MSI status; 2685 patients with known EBV status | CNN, Shufflenet | MSI vs MSS, AUROCs: 0.723-0.863; EBV (+) vs EBV (-), AUROCs: 0.672-0.859 | DL-based classifiers have the potential to provide faster decisions for pathologists and to offer therapeutic options tailored to the molecular profile of the individual patient |
| Kather <i>et al</i> [80] | MSI vs MSS | Training 81 patients +216 patients (TCGA-STAD) | ResNet-18 | AUC: 0.84 | This system provides significant improvements in molecular alterations screening workflow |
| Kather <i>et al</i> [81] | EBV (+) vs. EBV (-) | Training 317 patients (TCGA-STAD) | CNN, VGG19 | AUC: 0.80 | This workflow enables a fast and low-cost method to identify EBV and enables pathologists to check the plausibility of computer-based image classification (the black box of DL) |
| Hinata <i>et al</i> [82] | EBV+MSI/dMMR vs EBV- non MSI/dMMR | UTokyo training cohort: 326 patients; TCGA training cohort: 48 patients | CNNs,VGG16, VGG19, ResNet50, EfficientNetB0 | AUCs: 0.901-0.992 (Utokyo cohort); AUCs: 0.809-0.931 (TCGA cohort) | The model detects immunotherapy-sensitive GC subtypes from histological images at a lower cost and in a shorter time than the conventional methods |
| Zheng <i>et al</i> [83] | EBV (+) vs EBV (-) | EBV (+) 203 WSIs; EBV (-) 803 WSIs | EBVNet | AUROC: 0.969, Internal validation; AUROC: 0.941, External dataset AUROC: 0.895, TCGA dataset | The human-machine fusion significantly improves the diagnostic performance of both the EBVNet and the pathologist, provides an approach for the identification of EBV(+) GC, and may help effectively select patients for immunotherapy |
| Flinner <i>et al</i> [87] | EBV, MSI, GS, CIN | Training 84 WSIs (TCGA-STAD); Testing: 133 WSIs (TCGA-STAD) | CNN, DenseNet161 | AUC: 0.76 for four classes | The simplified molecular TCGA and GC subclasses could be predicted by DL directly based on H-E staining |
| Jang <i>et al</i> [88] | CDH1, ERBB2, KRAS, PIK3CA, TP53 mutations | 425 FF slides (TCGA-STAD); 320 FT slides (TCGA-STAD) | CNN, Inception-v3 | AUCs (FF-FI): CDH1 (0.667-0.778), ERBB2(0.63-0.833), KRAS (0.657-0.838); PIK3CA (0.688-0.761), TP53 (0.572-0.775) | When trained with appropriate tissue data, DL could predict genetic mutations in H-E-stained tissue slides |

| | | | | | |
|-------------------------------|-----------------------------------|---|---------------------------------------|---|--|
| Huang <i>et al</i> [109] | Metastatic LNs | 983 WSIs | ESCNN | AUC: 0.9936 | ESCNN improves the accuracy of pathologists in identifying metastatic LNs, micrometastases, and isolated tumor cells, allowing for shortening the review time |
| Hu <i>et al</i> [107] | Metastatic LNs | 222 patients | RCNN, Xception and DenseNet-121 | Accuracy 97.13%; PPV: 93.53, NPV: 97.99% | The system can be implemented into clinical workflow to assist pathologists in preliminary screening for LN metastases in GC patients |
| Matsushima <i>et al</i> [108] | Metastatic LNs | 827 lymph nodes | CNN | AUROC: 0.9994 | This DL-based diagnosis-aid system can assist pathologists in detecting LN metastasis in GC and reduce their workload |
| Wang <i>et al</i> [106] | Metastatic LNs, T/LNM | 9366 slides (7736 with metastasis) | Resnet-50 | LNM (+) vs (-): Sensitivity 98.5%, Specificity 96.1%; T/LNM: HR: 2.05 (univariate analysis); 1.39 (multivariate analysis) | This system can assist pathologists in detecting LN metastasis in GC and reduce their workload. Besides, T/LNM is prognostic of OS in GC patients |
| Hong <i>et al</i> [116] | dTSR (HE and CK7) | Training 13 WSIs; Testing 358 WSIs | cGAN | Kappa value: 0.623 (dTSR and vTSR); AUROC: 0.907; OS (P: 0.0024) | By diagnosing TSR in GC, this model predicts OS in the advanced stage of GC |
| Meier <i>et al</i> [127] | TME + Ki-67 | 248 patients | CNN | HRs: Ki67&CD20: 1.364, CD20&CD68: 1.338; Ki67&CD68: 1.473 | In combination with a panel of IHC markers, this model predicts the prognosis of patients with GC |
| Huang <i>et al</i> [128] | OS | Training: 2261 pictures; Internal validation: 960 pictures | GastroMIL | HR: 2.414 (univariate analysis), 1.843 (multivariate analysis) | The risk score computed by MIL-GC was proved to be the independent prognostic value of GC |
| Jiang <i>et al</i> [129] | 5-YS, 5-YDFS | 786 patients | ML, SVM | AUCs: 5-YS: 0.834; 5-YDFS: 0.828 | The classifier can accurately distinguishes GC patients with different OS and DFS and identifies a subgroup of patients with stage II and III disease who could benefit from adjuvant chemotherapy |
| Jiang <i>et al</i> [130] | Low SVM vs High SVM, 5-YS, 5-YDFS | Training: 223 patients; Internal validation: 218 patients External validation: 227 patients | ML, SVM | AUCs: 5-YS: 0.818; 5-YDFS: 0.827 | SVM signature distinguish GC patients with different OS and DFS and identifies a subgroup of patients with stage II and III disease who could benefit from adjuvant chemotherapy |
| Wang <i>et al</i> [131] | TME | 172 patients | CG _{signature} powered by AI | AUROC: 0.960 ± 0.01 (binary classification), 0.771 ± 0.024 to 0.904 ± 0.012 (ternary classification) | Digital grade cancer staging produced by CG _{signature} predicts the prognosis of GC and significantly outperforms the AJCC 8 th edition Tumor Node Metastasis staging system |

NC: Non cancer; GR: Grade; ATF7IP/MCAF1: Activating Transcription Factor 7 Interacting Protein; WSIs: Whole slide images; SV: Supervised; ML: Machine learning; wndchrm: weighted neighbor distances using a compound hierarchy of algorithms representing morphology; AUC: Area under the curve; ADC-D: Diffuse adenocarcinoma; ADC-O: Adenocarcinoma other; DL: Deep learning; CNN: Convolutional neural networks; RNN: Recurrent neural network; TC: Tubular carcinoma; MC: Mucinous carcinoma; PC: Papillary carcinoma; WD: Well differentiated; PD: Poorly differentiated; MSI: Microsatellite instable; MSS: Microsatellite stable EBV: Epstein-Barr virus; TCGA-STAD: Tumor Cancer Genome Atlas, Stomach adenocarcinoma dMMR: Deficient mismatch repair; GC: Gastric cancer; AUROC: Area under the receiver operating characteristics; GS: Genomically stable; CIN: Chromosomally unstable; LN: Lymph node; ESCNN: Enhanced streaming CNN; RCNN: Region based CNN; PPV: Positive predictive value; NPV: Negative predictive value; T/LNM: Tumor area-to-metastatic LN-area ratio; dTSR: Digital tumor-stroma ratio; HE: Hematoxylin and eosin; CK7: Cytokeratin 7; cGAN: Conditional generative adversarial network; vTSR: Visual tumor-stroma ratio; OS: Overall survival; TME: Tumor microenvironment; HR: Hazard ratio; 5-YS: Five year survival; 5-YDFS: Five year disease free survival; SVM: Support vector machine; AI: Artificial intelligence.

nontransparent (black boxes). Since various methods have been proposed to solve black boxes in DL in the recent past, additional studies applying these methods will contribute to the determination of the molecular subtypes of AI models of GC[84-86]. In a more recent study, Flinner *et al*[87], in their study emphasizing the error-proneness of the morphological and staining methods used to determine GC subtypes for subclassification, found that DL could be more effective in this regard. On the other hand, they also pointed out that image tiles labeled with false ground truth associated with GC heterogeneity may reduce the accuracy of DL but this can be overcome by first experimentally defining the test data.

Recently, the feasibility of a DL approach has also been evaluated in the classification of GC for mutations in the CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes[88]. High AUCs observed in both frozen and formalin-fixed tissues highlight that DL-based classifiers could predict the mutational status of these tumors. Although these results are promising for the application of AI to subtyping GC, additional studies are necessary, with further refinement of these methods.

Colorectal cancer

Similar to GC, molecular subtyping of CRC is essential for targeted treatment against critical oncogenic signaling pathways. CRCs are divided by molecular consensus into four types (CMS): 1. CMS1: Tumors with MSI that have a good prognosis in non metastatic stages; CMS2: Tumors with intermediate prognosis exhibiting epithelial gene expression, activated WNT and MYC signaling; CMS3: Tumors with intermediate prognosis demonstrating metabolic dysregulations; CMS4: Tumors with a poor prognosis that possess transforming growth factor beta (TGF- β) activation[89-91]. The identification of CRC with MSI is paramount because this group is susceptible to immunomodulating therapies[92,93]. Although some findings, such as tissue architecture, growth pattern, cellular morphology, and distributions of tumor stroma ratio (TSR) and tumor microenvironment (TME) provide some clues about the subclassification of these tumors, molecular stratification of patients necessitates RNA analyses that are expensive and difficult to standardize[94-96]. Accordingly, some studies have investigated the contribution of AI to tumor subclassification from HE-stained tissue sections by DL models (Table 2). Sirinukunwattana *et al*[97] demonstrated that a CNN-based model could detect CMS subtypes. At the same time, they criticized the potential over fitting of the computational model to the training cohort as a limitation of the study. In a more recent study, Echele *et al*[98] developed a DL model in a large series of 8836 cases of CRC to predict MSI tumors. In the international validation of the study group, the algorithm achieved a high performance [area under the receiver operating curve (AUROC) of 0.96][80]. Other investigators have also reported similar results, pointing out the potential use of DL models for detecting molecular subtypes of CRC[77,99-101]. In a retrospective study, a DL pipeline method was developed based on experimental setups similar to previous studies[102]. Three models were used to predict mutation density (low *vs* high), MSI, CIN, and GpG island methylator phenotype. The mutated and wild-type BRAF, TP53, and KRAS types were also investigated. This method showed higher AUROCs for the prediction of hypermutation, MSI, CIN, BRAF, and TP53 compared to previously reported data, suggesting that AI methods may provide the stratification of patients with CRC for targeted therapies. However, further large-scale validations with multicenter datasets are required before their implementation in pathological practice.

LYMPH NODE METASTASIS

Gastric cancer

Another important parameter that predicts GC behavior and treatment is lymph node metastasis (LNM) [103]. However, identifying LNM is still a challenging and tedious task in pathological practice, making the implementation of AI an attractive tool to reduce the workload[104,105]. Although numerous studies have demonstrated that DL-based algorithms can detect metastatic lymph nodes in GC with a similar level of accuracy to human specialists, these algorithms have not yet been implemented into pathology practice[106-108] (Table 3). The failure to integrate these algorithms is related to the characteristics of WSIs, the excessive effort required to apply the annotation, and the limited associated data. Recently, Huang *et al*[109] developed a weakly supervised end-to-end technique termed enhanced streaming CNN (ESCNN). Their results revealed that the routine pathological evaluation benefitted from the AI-assisted LN assessment workflow regarding review time, sensitivity, and consistency. On the other hand, AI-attributable false alarms that misled the pathologists on negative results led to a decrease in specificity from 94% to 84%, which needs more large-scale or multicenter studies to check the effectiveness of the workflow.

Colorectal cancer

Recent evidence indicates that features extracted by DL models from routine histologic slides can predict LNM in CRC[110-112] (Table 2). For example, Kwak *et al*[110] detected LNM by generating a score based on the ratio of peritumoral stroma to tumor tissue on a test set. In another study, the presence of LNM was detected with a model which segmented WSIs into areas such as tumor budding or poorly differentiated clusters[111]. More recently, Kiehl *et al*[112] performed an approach that uses

DL-based image analysis (slide-based artificial intelligence predictor) in association with patient data to estimate LNM in CRC patients. Their results indicated that LNM could be predicted in patients with CRC through AI applications from histological slides to a similar level to using a classifier containing clinical data.

THE TUMOR STROMA RATIO, TUMOR MICROENVIRONMENT AND TUMOR BUDDING

Gastric cancer

In recent years, it has been shown that the TSR in many organ tumors is an important clue to the course of the disease. In particular, stromal dominance has been observed to be an independent prognostic factor in many tumors, including GIS[113,114]. However, TSRs are not included in pathology report protocols because of the lack of a standard procedure among different methodologies and a low reproducibility related to the high interobserver variation[115]. Recently, a DL pipeline has been introduced to facilitate the automated assessment of TSR in GC[116]. Although this model has been shown to be effective in detecting survival according to the low and high TSR rates in advanced GC, it was emphasized that some limitations, such as the nonautomatic selection of hot spots and the use of a single test, should be eliminated. Therefore, there is a need for many studies on the use of AI applications in TSR determination of GC.

In a recent study, a DL model determined the tumor-to-metastatic lymph node-area ratio in metastatic lymph nodes in patients with GC[106]. Statistical analysis also revealed that this ratio is an independent prognostic factor warranting further investigation.

Colorectal cancer

In CRC, recent studies have demonstrated that lymphocytes and fibroblasts profoundly shape the TME and significantly impact tumor behavior[117-119]. In addition, it has been shown that CRC may have a poor prognosis due to tumor budding (1-5 cells in the invasive area)[120]. In the literature, seven studies of AI methods have been identified to determine these parameters in a more objective and time-saving manner (Table 2). However, many of them used different methods. Three models focused on the classification of the cell types, such as epithelial, inflammatory, fibroblast, lymphocytes, and others (mucus, smooth muscle, normal mucosa, stroma, and cancer epithelium)[121-123]. In an elegant study, a DL algorithm was proposed for estimating the risk of distant metastasis by analyzing the TME[123]. Cell detection and cell classification were evaluated in two CNNs used to build a cell network. In each tumor, a tissue phenotype signature was obtained by proportioning the area of tissue phenotypes to the total tissue area. Statistical analysis revealed that the connection frequency (CF) of the smooth muscle ratio, the CF of the inflammation ratio, and the appearance (AP) based on inflammation could independently estimate the development of distant metastasis. Distant metastasis-free survival analysis indicated that CF smooth muscle and AP inflammation ratios were potential prognosticators. Although the hazard ratios for CF of the smooth muscle ratio and AP inflammation were 2.11 and 0.39, respectively, the AUC values for distant metastasis prediction were 0.59 for the CF of the smooth muscle ratio and 0.64 for AP based on inflammation. As emphasized by the authors, specific immunohistochemical staining can improve the prediction of distant metastases by increasing the informative value of histological slides. Another limitation of this study is the small number of metastatic cases. Another recent study was performed to detect CD3- and CD8-positive immune cells on WSIs of slides stained by immunohistochemistry in a multicenter cohort by four different methods[124]. U-Net obtained the highest performance and highest agreement with manual evaluation (0.72), which was higher than that of pathologists (K = 0.64), supporting that DL models are helpful for automatically detecting lymphocytes in immunohistochemically stained tissue sections.

In CRC, the automatic tumor budding evaluation on immunohistochemical pankeratin-stained slides revealed that the absolute number of buds per image was significantly correlated with manually segmented ground truth (R: 0.86)[120]. Interestingly, the number of spatial clusters of buds in hot spots was significantly correlated with the prognosis. In three studies, the impact of detecting the TSR or deep stroma score in CRC by DL algorithms was found to be an independent parameter to predict tumor behavior[115,121,125] (Table 2).

Recently, Zhao *et al*[126] demonstrated that the ratio of the mucinous component in the tumor area (MTR) quantified by AI is an independent prognostic factor in CRC. On the other hand, the most invasive part of primary tumors was selected for evaluation. As noted by the authors, measuring the exact proportion and prognostic value of mucus in the entire tumor is still worthy of further investigation.

SURVIVAL OUTCOMES

Gastric cancer

Another continuing research topic is evaluating survival outcomes in GC with AI models[127-129] (Table 3). Recently, support vector machine (SVM), one of the popular algorithms in ML, has been applied to predict the survival of GC. Jiang *et al*[129] demonstrated that SVM could be useful in predicting the outcome and identifying patients with GC who might benefit from adjuvant therapy. In this study, the classifier incorporated patient gender, carcinoembryonic antigen levels, LNM, and the protein expression level of eight features, composed of CD3 invasive margin (IM), CD3 center of the tumor (CT), CD8IM, CD45ROCT, CD57IM, CD66bIM, CD68CT, and CD34. There were significant variations between the high- and low-GC-SVM classifiers. Recently, Huang *et al*[128] designed MIL-GC (a DL-based model) to predict overall survival (OS) in patients with GC. They observed C-indices of 0.728 and 0.671 in the training and internal validation sets, respectively. The external validation likewise exhibited strong prognostic prediction performance (C-index = 0.657), confirming the resilience of the two models. Furthermore, univariate and multivariate Cox analyses demonstrated that the risk score derived by MIL-GC has independent prognostic significance, indicating the potential of AI approaches to predict GC behavior. Additionally, tumor progression includes complex interactions between malignant cells and their surrounding microenvironment (TME)[130]. TME targeting and reprogramming is, in fact, can be a potential strategy to achieve antitumor effects in many cancers. Several AI studies involving the TME have recently demonstrated that these methods can determine the prognosis of GIS cancers. Regarding GC, Wang *et al*[131], suggested a graph NN-based solution, CellGraph Signature powered AI, for the digital staging of TME and the exact prediction of patient survival by combining and converting multiplexed immunohistochemistry (mIHC) images as Cell-Graphs. The survival prediction achieved outstanding model performance for both binary and ternary classifications. Furthermore, survival analysis revealed that this method outperforms the AJCC 8th edition Tumor Node Metastasis staging system in discriminating both binary and ternary classes with statistical significance (P value < 0.0001), implying the effectiveness and advantages of such an AI-powered digital staging system in DP and precision oncology.

These data demonstrate that AI-based models allow prognosis prediction in GC. However, developing efficient models requires training on large sets reflecting scanning and staining protocols variability.

Colorectal cancer

Regarding prognostic evaluations from HE-stained slides by AI in CRC, some DL models have been developed for prognostication (Table 2). Bychkov *et al*[132] combined a CNN and a recurrent NN model to estimate the disease-specific five-year survival from tumor tissue microarray samples without tissue classification. The model classified patients into a low- or high-risk group (AUC of 0.69). This result was more significant than the AUC of the visual evaluation of the pathologist (AUC of 0.58) or the histological grade determined at the time of the original diagnosis (AUC of 0.57). However, an external dataset was not included. In another study by Skrede *et al*[133], diverse data from four different cohorts were used to develop an automatic prognostic marker to predict the outcome. The model included a CNN used to separate tumor tissue and two other CNN ensembles that identified individuals as having a favorable or poor survival. Patients were assigned as uncertain when the two CNN ensembles predicted different outcomes. In an external test group, the classifier was a strong predictor of survival. In addition, the output of the two CNN ensembles produced a strong predictive score related to patient outcome (AUC of 0.71). A generalization of this approach has been recommended, as an external test cohort from more than one medical center demonstrated similar hazard ratios.

Jiang *et al*[134], to achieve a shorter computational time, developed a hybrid model by synergizing ML algorithms with DL (InceptionResNetV2 and gradient boosting decision machine classifier) to predict the survival of patients with stage III CRC. While the internal test sets constituted a Chinese cohort, external testing was performed on the TCGA cohort. They revealed that the model stratifies patients with stage III colon cancer into high- and low-risk recurrence and poor and favorable prognostic groups directly from tissue sections. These data suggest that the analysis of H-E-stained tissue samples by AI methods could serve as a digital prognostic biomarker in CRC. However, additional studies are warranted to support the evaluation of the performance of these methods in larger patient series.

OVERALL LIMITATIONS OF AI-BASED APPLICATIONS IN REAL-LIFE PRACTICE

In the literature, there are some frequently discussed topics considering the general challenges of AI such as identification of the clinical need, ethical considerations, funding, optimization of data-sets, annotation of the dataset, regulation, validation, and implementation[46].

Recognizing the actual clinical need and defining a potential solution is the first stage in developing the AI application. However, there can be an imbalance between the benefits in daily pathological practice and the total cost of its implementation. As a result, the market for a particular AI tool may be too tiny and it may not be profitable.

Although patients can provide permission for data to be used for studies, constructing AI models may have issues if commercial use is not approved[135]. In order to develop a framework for global data sharing, patient consent should include the possibility of its commercial use for product development[40].

Training on huge datasets is necessary for developing AI systems with high performance in digital pathology. Changes related to differences in fixation, tissue thickness, and variations in staining and scanning protocols encountered in preanalytical and analytical phases may influence data accuracy[136, 137]. For example, it is difficult to convert a glass slide to WSI, and changing the hue of the slide could affect AI accuracy. Many AI algorithms have emerged for this purpose recently, including staining and color features[138,139]. In addition, a number of algorithms are presented to optimize WSI quality. These algorithms identify areas of the highest quality and exclude areas that are out of focus or affected by artifacts[140,141].

Concerning the implementation of AI, to enable users to shift the daily routine practice in the pathology laboratory, from glass slides to WSIs, the first step is to install an institutional IT infrastructure. In addition to these changes in infrastructure, pathology residency training might need to be adjusted in accordance with the availability of this new tool. Preventing residents from relying completely on AI while also allowing them to benefit from it as a helping instrument would require fine balancing and planning prior to its installation[142].

Similar to other clinical tests, quality assurance is crucial, hence it is urgently necessary to develop a plan for external quality assurance for applications. Furthermore, laboratory workers should also be familiar with the quality management system.

Although some algorithms and automated AI models are thought to perform better than pathologists, pathologists will always be required to audit technology and control mechanisms in AI implementation [143].

CONCLUSION

In this review, we outlined the potential of AI applications for evaluating pathological parameters related to the behavior of GIS cancers. Current data suggest the merit of AI-based approaches in assessing tumor grading, subtyping, detection of metastasis, and prognosis in GC and CRC. In addition, these methods encourage biomarker discovery by revealing predictions that are impossible when using traditional visual methods. Regarding EC, there is still much room for improvement in developing AI models to predict the behavior of these tumors in pathology. On the other hand, the enormous potential of AI in improving workflows, eliminating simple errors, and increasing objectivity during pathological evaluations to determine the behavior of GIS cancers should motivate researchers to overcome the many remaining hurdles. In algorithm development, variations in imaging data, interobserver variability during interpretations, model transparency, and interpretability are significant challenges to be solved. A large number of studies with external validation and quality controls implemented on large datasets are essential in meeting the standards of these methods. Thereby, AI applications that are practical, interpretable, manageable, and cost-effective can play a crucial role in the development of pathological evaluations to be performed in the prognosis and treatment of GIS tumors.

FOOTNOTES

Author contributions: Yavuz A, Alpsoy A, Gedik EO, Celik MY performed the data acquisition; Bassorgun CI, Unal B and Elpek GO designed the outline and coordinated the writing of the paper; all authors equally contributed to the writing of the paper and preparation of the tables.

Conflict-of-interest statement: All the authors declare that they have no conflict of interest.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

Country/Territory of origin: Turkey

ORCID number: Aysen Yavuz 0000-0001-9991-5515; Anil Alpsoy 0000-0003-4978-7652; Elif Ocak Gedik 0000-0003-2618-498X; Mennan Yigitcan Celik 0000-0001-8769-5156; Cumhuri Ibrahim Bassorgun 0000-0003-2440-511X; Betul Unal 0000-

0001-7680-1808; Gulsum Ozlem Elpek 0000-0002-1237-5454.

S-Editor: Liu JH**L-Editor:** A**P-Editor:** Liu JH

REFERENCES

- 1 **Arnold M**, Abnet CC, Neale RE, Vignat J, Giovannucci EL, McGlynn KA, Bray F. Global Burden of 5 Major Types of Gastrointestinal Cancer. *Gastroenterology* 2020; **159**: 335-349.e15 [PMID: 32247694 DOI: 10.1053/j.gastro.2020.02.068]
- 2 **Bray F**, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]
- 3 **Johdi NA**, Sukor NF. Colorectal Cancer Immunotherapy: Options and Strategies. *Front Immunol* 2020; **11**: 1624 [PMID: 33042104 DOI: 10.3389/fimmu.2020.01624]
- 4 **Li K**, Zhang A, Li X, Zhang H, Zhao L. Advances in clinical immunotherapy for gastric cancer. *Biochim Biophys Acta Rev Cancer* 2021; **1876**: 188615 [PMID: 34403771 DOI: 10.1016/j.bbcan.2021.188615]
- 5 **Niazi MKK**, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol* 2019; **20**: e253-e261 [PMID: 31044723 DOI: 10.1016/S1470-2045(19)30154-8]
- 6 **Abels E**, Pantanowitz L, Aeffner F, Zarella MD, van der Laak J, Bui MM, Vemuri VN, Parwani AV, Gibbs J, Agosto-Arroyo E, Beck AH, Kozlowski C. Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. *J Pathol* 2019; **249**: 286-294 [PMID: 31355445 DOI: 10.1002/path.5331]
- 7 **Dangott B**, Parwani A. Whole slide imaging for teleconsultation and clinical use. *J Pathol Inform* 2010; **1** [PMID: 20805956 DOI: 10.4103/2153-3539.65342]
- 8 **Evans AJ**, Depeiza N, Allen SG, Fraser K, Shirley S, Chetty R. Use of whole slide imaging (WSI) for distance teaching. *J Clin Pathol* 2021; **74**: 425-428 [PMID: 32646928 DOI: 10.1136/jclinpath-2020-206763]
- 9 **Saillard C**, Schmauch B, Laifa O, Moarii M, Toldo S, Zaslavskiy M, Pronier E, Laurent A, Amaddeo G, Regnault H, Sommacale D, Ziolo M, Pawlotsky JM, Mulé S, Luciani A, Wainrib G, Clozel T, Courtiol P, Calderaro J. Predicting Survival After Hepatocellular Carcinoma Resection Using Deep Learning on Histological Slides. *Hepatology* 2020; **72**: 2000-2013 [PMID: 32108950 DOI: 10.1002/hep.31207]
- 10 **Wong ANN**, He Z, Leung KL, To CCK, Wong CY, Wong SCC, Yoo JS, Chan CKR, Chan AZ, Lacambra MD, Yeung MHY. Current Developments of Artificial Intelligence in Digital Pathology and Its Future Clinical Applications in Gastrointestinal Cancers. *Cancers (Basel)* 2022; **14** [PMID: 35954443 DOI: 10.3390/cancers14153780]
- 11 **Courtiol P**, Maussion C, Moarii M, Pronier E, Pilcer S, Sefta M, Manceron P, Toldo S, Zaslavskiy M, Le Stang N, Girard N, Elemento O, Nicholson AG, Blay JY, Galateau-Sallé F, Wainrib G, Clozel T. Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nat Med* 2019; **25**: 1519-1525 [PMID: 31591589 DOI: 10.1038/s41591-019-0583-3]
- 12 **Dun XP**, Parkinson DB. Visualizing peripheral nerve regeneration by whole mount staining. *PLoS One* 2015; **10**: e0119168 [PMID: 25738874 DOI: 10.1371/journal.pone.0119168]
- 13 **Kim D**, Pantanowitz L, Schüffler P, Yarlagadda DVK, Ardon O, Reuter VE, Hameed M, Klimstra DS, Hanna MG. (Re) Defining the High-Power Field for Digital Pathology. *J Pathol Inform* 2020; **11**: 33 [PMID: 33343994 DOI: 10.4103/jpi.jpi_48_20]
- 14 **Heffner S**, Colgan O, Doolan C. Digital Pathology. Available from: <https://www.leicabiosystems.com/en-br/knowledgepathway/digitalpathology/>
- 15 **Borowsky AD**, Glassy EF, Wallace WD, Kallichanda NS, Behling CA, Miller DV, Oswal HN, Feddersen RM, Bakhtar OR, Mendoza AE, Molden DP, Saffer HL, Wixom CR, Albro JE, Cessna MH, Hall BJ, Lloyd IE, Bishop JW, Darrow MA, Gui D, Jen KY, Walby JAS, Bauer SM, Cortez DA, Gandhi P, Rodgers MM, Rodriguez RA, Martin DR, McConnell TG, Reynolds SJ, Spiegel JH, Stepenaskie SA, Viktorova E, Magari R, Wharton KA, Qiu J, Bauer TW. Digital Whole Slide Imaging Compared With Light Microscopy for Primary Diagnosis in Surgical Pathology. *Arch Pathol Lab Med* 2020; **144**: 1245-1253 [PMID: 32057275 DOI: 10.5858/arpa.2019-0569-OA]
- 16 **Hanna MG**, Reuter VE, Ardon O, Kim D, Sirintrapun SJ, Schüffler PJ, Busam KJ, Sauter JL, Brogi E, Tan LK, Xu B, Bale T, Agaram NP, Tang LH, Ellenson LH, Philip J, Corsale L, Stamelos E, Friedlander MA, Ntiamoah P, Labasin M, England C, Klimstra DS, Hameed M. Validation of a digital pathology system including remote review during the COVID-19 pandemic. *Mod Pathol* 2020; **33**: 2115-2127 [PMID: 32572154 DOI: 10.1038/s41379-020-0601-5]
- 17 **Aloqaily A**, Polonia A, Campelos S, Alrefae N, Vale J, Caramelo A, Eloy C. Digital Versus Optical Diagnosis of Follicular Patterned Thyroid Lesions. *Head Neck Pathol* 2021; **15**: 537-543 [PMID: 33128731 DOI: 10.1007/s12105-020-01243-y]
- 18 **Yeh FC**, Ye Q, Hitchens TK, Wu YL, Parwani AV, Ho C. Mapping stain distribution in pathology slides using whole slide imaging. *J Pathol Inform* 2014; **5**: 1 [PMID: 24672736 DOI: 10.4103/2153-3539.126140]
- 19 **Calderaro J**, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. *Gut* 2021; **70**: 1183-1193 [PMID: 33214163 DOI: 10.1136/gutjnl-2020-322880]
- 20 **Taqi SA**, Sami SA, Sami LB, Zaki SA. A review of artifacts in histopathology. *J Oral Maxillofac Pathol* 2018; **22**: 279 [PMID: 30158787 DOI: 10.4103/jomfp.JOMFP_125_15]
- 21 **Salvi M**, Acharya UR, Molinari F, Meiburger KM. The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Comput Biol Med* 2021; **128**: 104129

- [PMID: 33254082 DOI: 10.1016/j.compbio.2020.104129]
- 22 **Tabesh A**, Teverovskiy M, Pang HY, Kumar VP, Verbel D, Kotsianti A, Saidi O. Multifeature prostate cancer diagnosis and Gleason grading of histological images. *IEEE Trans Med Imaging* 2007; **26**: 1366-1378 [PMID: 17948727 DOI: 10.1109/TMI.2007.898536]
 - 23 **Macenko M**, Niethammer M, Marron JS, Borland D, Woosley JT, Guan X, Schmitt C, Thomas NE. A method for normalizing histology slides for quantitative analysis. Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009 June 28–July 1; Boston, USA, 2009: 1107–1110 [DOI: 10.1109/ISBI.2009.5193250]
 - 24 **Yagi Y**. Color standardization and optimization in whole slide imaging. *Diagn Pathol* 2011; **6** Suppl 1: S15 [PMID: 21489185 DOI: 10.1186/1746-1596-6-S1-S15]
 - 25 **Lee JS**, Ma YX. Stain Style Transfer for Histological Images Using S3CGAN. *Sensors (Basel)* 2022; **22** [PMID: 35161789 DOI: 10.3390/s22031044]
 - 26 **Kausar T**, Kausar A, Ashraf MA, Siddique MF, Wang M, Sajid M, Siddique MZ, Haq AU, Riaz I. SA-GAN: Stain Acclimation Generative Adversarial Network for Histopathology Image Analysis. *Appl Sci* 2022; **12**: 288 [DOI: 10.3390/app12010288]
 - 27 **Schneider CA**, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012; **9**: 671-675 [PMID: 22930834 DOI: 10.1038/nmeth.2089]
 - 28 **Bankhead P**, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, James JA, Salto-Tellez M, Hamilton PW. QuPath: Open source software for digital pathology image analysis. *Sci Rep* 2017; **7**: 16878 [PMID: 29203879 DOI: 10.1038/s41598-017-17204-5]
 - 29 **Aubreville M**, Bertram C, Klopffleisch R, Maier A. SlideRunner. In *Bildverarbeitung für die Medizin* 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 309–314 [DOI: 10.007/978-3-662-56-537-7_81]
 - 30 **Tizhoosh HR**, Pantanowitz L. Artificial Intelligence and Digital Pathology: Challenges and Opportunities. *J Pathol Inform* 2018; **9**: 38 [PMID: 30607305 DOI: 10.4103/jpi.jpi_53_18]
 - 31 **Williams B**, Hanby A, Millican-Slater R, Verghese E, Nijhawan A, Wilson I, Besusparis J, Clark D, Snead D, Rakha E, Treanor D. Digital pathology for primary diagnosis of screen-detected breast lesions - experimental data, validation and experience from four centres. *Histopathology* 2020; **76**: 968-975 [PMID: 31994224 DOI: 10.1111/his.14079]
 - 32 **Eloy C**, Vale J, Curado M, Polónia A, Campelos S, Caramelo A, Sousa R, Sobrinho-Simões M. Digital Pathology Workflow Implementation at IPATIMUP. *Diagnostics (Basel)* 2021; **11** [PMID: 34829458 DOI: 10.3390/diagnostics11112111]
 - 33 **The Royal College of Pathologists**. Digital Pathology. (accessed on 20 March 2022). Available from: <https://www.rcpath.org/profession/digital-pathology.html>
 - 34 **Gurcan MN**, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng* 2009; **2**: 147-171 [PMID: 20671804 DOI: 10.1109/RBME.2009.2034865]
 - 35 **Rashidi HH**, Tran NK, Betts EV, Howell LP, Green R. Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. *Acad Pathol* 2019; **6**: 2374289519873088 [PMID: 31523704 DOI: 10.1177/2374289519873088]
 - 36 **Saxena S**, Gyanchandani M. Machine Learning Methods for Computer-Aided Breast Cancer Diagnosis Using Histopathology: A Narrative Review. *J Med Imaging Radiat Sci* 2020; **51**: 182-193 [PMID: 31884065 DOI: 10.1016/j.jmir.2019.11.001]
 - 37 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]
 - 38 **Wang X**, Chen H, Gan C, Lin H, Dou Q, Tsougenis E, Huang Q, Cai M, Heng PA. Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. *IEEE Trans Cybern* 2020; **50**: 3950-3962 [PMID: 31484154 DOI: 10.1109/TCYB.2019.2935141]
 - 39 **Silva-Rodríguez J**, Colomer A, Naranjo V. WeGleNet: A weakly-supervised convolutional neural network for the semantic segmentation of Gleason grades in prostate histology images. *Comput Med Imaging Graph* 2021; **88**: 101846 [PMID: 33485056 DOI: 10.1016/j.compmid.2020.101846]
 - 40 **Yoshida H**, Kiyuna T. Requirements for implementation of artificial intelligence in the practice of gastrointestinal pathology. *World J Gastroenterol* 2021; **27**: 2818-2833 [PMID: 34135556 DOI: 10.3748/wjg.v27.i21.2818]
 - 41 **Cosatto E**, Laquerre PF, Malon C, Graf HP, Saito A, Kiyuna T, Marugame A, Kamijo K. Automated gastric cancer diagnosis on H and E-stained sections; training a classifier on a large scale with multiple instance machine learning. Proceedings of SPIE - Progress in Biomedical Optics and Imaging. MI: 2013 [DOI: 10.1117/12.2007047]
 - 42 **Mattocks CJ**, Morris MA, Matthijs G, Swinnen E, Corveleyn A, Dequeker E, Müller CR, Pratt V, Wallace A; EuroGentest Validation Group. A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur J Hum Genet* 2010; **18**: 1276-1288 [PMID: 20664632 DOI: 10.1038/ejhg.2010.101]
 - 43 **Joshi, R.**; Kruger, A.J.; Sha, L.; Kannan, M.; Khan, A.A.; Stumpe, M.C. Learning relevant H&E slide morphologies for prediction of colorectal cancer tumor mutation burden using weakly supervised deep learning. *J Clin Oncol* 2020; **38**: e15244 [DOI: 10.1200/JCO.2020.38.15_suppl.e15244]
 - 44 **Lu MY**, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng* 2021; **5**: 555-570 [PMID: 33649564 DOI: 10.1038/s41551-020-00682-w]
 - 45 **van der Laak J**, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat Med* 2021; **27**: 775-784 [PMID: 33990804 DOI: 10.1038/s41591-021-01343-4]
 - 46 **Colling R**, Pitman H, Oien K, Rajpoot N, Macklin P; CM-Path AI in Histopathology Working Group, Snead D, Sackville T, Verrill C. Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. *J Pathol* 2019; **249**: 143-150 [PMID: 31144302 DOI: 10.1002/path.5310]
 - 47 **Elmore JG**, Longton GM, Carney PA, Geller BM, Onega T, Tosteson AN, Nelson HD, Pepe MS, Allison KH, Schnitt SJ, O'Malley FP, Weaver DL. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* 2015; **313**: 1122-1132 [PMID: 25781441 DOI: 10.1001/jama.2015.1405]

- 48 **Amann J**, Blasimme A, Vayena E, Frey D, Madai VI; Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak* 2020; **20**: 310 [PMID: [33256715](#) DOI: [10.1186/s12911-020-01332-6](#)]
- 49 **Tomita N**, Abdollahi B, Wei J, Ren B, Suriawinata A, Hassanpour S. Attention-Based Deep Neural Networks for Detection of Cancerous and Precancerous Esophagus Tissue on Histopathological Slides. *JAMA Netw Open* 2019; **2**: e1914645 [PMID: [31693124](#) DOI: [10.1001/jamanetworkopen.2019.14645](#)]
- 50 **Gehring M**, Crispin-Ortuzar M, Berman AG, O'Donovan M, Fitzgerald RC, Markowitz F. Triage-driven diagnosis of Barrett's esophagus for early detection of esophageal adenocarcinoma using deep learning. *Nat Med* 2021; **27**: 833-841 [PMID: [33859411](#) DOI: [10.1038/s41591-021-01287-9](#)]
- 51 **Vu HT**, Lopez R, Bennett A, Burke CA. Individuals with sessile serrated polyps express an aggressive colorectal phenotype. *Dis Colon Rectum* 2011; **54**: 1216-1223 [PMID: [21904135](#) DOI: [10.1097/DCR.0b013e318228f8a9](#)]
- 52 **Foss FA**, Milkins S, McGregor AH. Inter-observer variability in the histological assessment of colorectal polyps detected through the NHS Bowel Cancer Screening Programme. *Histopathology* 2012; **61**: 47-52 [PMID: [22486166](#) DOI: [10.1111/j.1365-2559.2011.04154.x](#)]
- 53 **Osmond A**, Li-Chang H, Kirsch R, Divaris D, Falck V, Liu DF, Marginean C, Newell K, Parfitt J, Rudrick B, Sapp H, Smith S, Walsh J, Wasty F, Driman DK. Interobserver variability in assessing dysplasia and architecture in colorectal adenomas: a multicentre Canadian study. *J Clin Pathol* 2014; **67**: 781-786 [PMID: [25004943](#) DOI: [10.1136/jclinpath-2014-202177](#)]
- 54 **US Preventive Services Task Force**, Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, Davis EM, Donahue KE, Doubeni CA, Krist AH, Kubik M, Li L, Ogedegbe G, Owens DK, Pbert L, Silverstein M, Stevermer J, Tseng CW, Wong JB. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* 2021; **325**: 1965-1977 [PMID: [34003218](#) DOI: [10.1001/jama.2021.6238](#)]
- 55 **Xu Y**, Jia Z, Wang LB, Ai Y, Zhang F, Lai M, Chang EI. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics* 2017; **18**: 281 [PMID: [28549410](#) DOI: [10.1186/s12859-017-1685-x](#)]
- 56 **Korbar B**, Olofson AM, Miraflor AP, Nicka CM, Suriawinata MA, Torresani L, Suriawinata AA, Hassanpour S. Deep Learning for Classification of Colorectal Polyps on Whole-slide Images. *J Pathol Inform* 2017; **8**: 30 [PMID: [28828201](#) DOI: [10.4103/jpi.jpi_34_17](#)]
- 57 **Haj-Hassan H**, Chaddad A, Harkouss Y, Desrosiers C, Toews M, Tanougast C. Classifications of Multispectral Colorectal Cancer Tissues Using Convolution Neural Network. *J Pathol Inform* 2017; **8**: 1 [PMID: [28400990](#) DOI: [10.4103/jpi.jpi_47_16](#)]
- 58 **Ponzo F**, Macii E, Ficarra E, Di Cataldo S. Colorectal Cancer Classification using Deep Convolutional Networks-An Experimental Study. Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2. Bioimaging, 2018: 58-66. Available from: <https://www.sciencepress.org/papers/2018/66431.pdf>
- 59 **Sena P**, Fioresi R, Faglioni F, Losi L, Faglioni G, Roncucci L. Deep learning techniques for detecting preneoplastic and neoplastic lesions in human colorectal histological images. *Oncol Lett* 2019; **18**: 6101-6107 [PMID: [31788084](#) DOI: [10.3892/ol.2019.10928](#)]
- 60 **Iizuka O**, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M. Deep Learning Models for Histopathological Classification of Gastric and Colonic Epithelial Tumours. *Sci Rep* 2020; **10**: 1504 [PMID: [32001752](#) DOI: [10.1038/s41598-020-58467-9](#)]
- 61 **Wei JW**, Suriawinata AA, Vaickus LJ, Ren B, Liu X, Lisovsky M, Tomita N, Abdollahi B, Kim AS, Snover DC, Baron JA, Barry EL, Hassanpour S. Evaluation of a Deep Neural Network for Automated Classification of Colorectal Polyps on Histopathologic Slides. *JAMA Netw Open* 2020; **3**: e203398 [PMID: [32324237](#) DOI: [10.1001/jamanetworkopen.2020.3398](#)]
- 62 **Awan R**, Sirinukunwattana K, Epstein D, Jefferyes S, Qidwai U, Aftab Z, Mujeeb I, Snead D, Rajpoot N. Glandular Morphometrics for Objective Grading of Colorectal Adenocarcinoma Histology Images. *Sci Rep* 2017; **7**: 16852 [PMID: [29203775](#) DOI: [10.1038/s41598-017-16516-w](#)]
- 63 **Arai T**, Sakurai U, Sawabe M, Honma N, Aida J, Ushio Y, Kanazawa N, Kuroiwa K, Takubo K. Frequent microsatellite instability in papillary and solid-type, poorly differentiated adenocarcinomas of the stomach. *Gastric Cancer* 2013; **16**: 505-512 [PMID: [23274922](#) DOI: [10.1007/s10120-012-0226-6](#)]
- 64 **Sugimura H**. Editorial: an obsession with subtyping gastric cancer. *Gastric Cancer* 2013; **16**: 451-453 [PMID: [23483302](#) DOI: [10.1007/s10120-013-0243-0](#)]
- 65 **Lauren P**. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An Attempt At A Histo-Clinical Classification. *Acta Pathol Microbiol Scand* 1965; **64**: 31-49 [PMID: [14320675](#) DOI: [10.1111/apm.1965.64.1.31](#)]
- 66 **Yasuda Y**, Tokunaga K, Koga T, Sakamoto C, Goldberg IG, Saitoh N, Nakao M. Computational analysis of morphological and molecular features in gastric cancer tissues. *Cancer Med* 2020; **9**: 2223-2234 [PMID: [32012497](#) DOI: [10.1002/cam4.2885](#)]
- 67 **Kanavati F**, Tsuneki M. A deep learning model for gastric diffuse-type adenocarcinoma classification in whole slide images. *Sci Rep* 2021; **11**: 20486 [PMID: [34650155](#) DOI: [10.1038/s41598-021-99940-3](#)]
- 68 **Fu B**, Zhang M, He J, Cao Y, Guo Y, Wang R. StoHisNet: A hybrid multi-classification model with CNN and Transformer for gastric pathology images. *Comput Methods Programs Biomed* 2022; **221**: 106924 [PMID: [35671603](#) DOI: [10.1016/j.cmpb.2022.106924](#)]
- 69 **Su F**, Li J, Zhao X, Wang B, Hu Y, Sun Y, Ji J. Interpretable tumor differentiation grade and microsatellite instability recognition in gastric cancer using deep learning. *Lab Invest* 2022; **102**: 641-649 [PMID: [35177797](#) DOI: [10.1038/s41374-022-00742-6](#)]
- 70 **Fukayama M**, Goldblum JR, Miettinen M, et al Digestive System Tumours, WHO Classification of Tumours (5th edn). International Agency for Research on Cancer: Lyon, 2019. Available from: <http://iarc.who.int>
- 71 **Cancer Genome Atlas Research Network**. Comprehensive molecular characterization of gastric adenocarcinoma.

- Nature* 2014; **513**: 202-209 [PMID: 25079317 DOI: 10.1038/nature13480]
- 72 **Kim HS**, Shin SJ, Beom SH, Jung M, Choi YY, Son T, Kim HI, Cheong JH, Hyung WJ, Noh SH, Chung H, Park JC, Shin SK, Lee SK, Koom WS, Lim JS, Chung HC, Rha SY, Kim H. Comprehensive expression profiles of gastric cancer molecular subtypes by immunohistochemistry: implications for individualized therapy. *Oncotarget* 2016; **7**: 44608-44620 [PMID: 27331626 DOI: 10.18632/oncotarget.10115]
 - 73 **Sohn BH**, Hwang JE, Jang HJ, Lee HS, Oh SC, Shim JJ, Lee KW, Kim EH, Yim SY, Lee SH, Cheong JH, Jeong W, Cho JY, Kim J, Chae J, Lee J, Kang WK, Kim S, Noh SH, Ajani JA, Lee JS. Clinical Significance of Four Molecular Subtypes of Gastric Cancer Identified by The Cancer Genome Atlas Project. *Clin Cancer Res* 2017; **23**: 4441-4449 [PMID: 28747339 DOI: 10.1158/1078-0432.CCR-16-2211]
 - 74 **Grogg KL**, Lohse CM, Pankratz VS, Halling KC, Smyrk TC. Lymphocyte-rich gastric cancer: associations with Epstein-Barr virus, microsatellite instability, histology, and survival. *Mod Pathol* 2003; **16**: 641-651 [PMID: 12861059 DOI: 10.1097/01.MP.0000076980.73826.C0]
 - 75 **Fukayama M**, Abe H, Kunita A, Shinozaki-Ushiku A, Matsusaka K, Ushiku T, Kaneda A. Thirty years of Epstein-Barr virus-associated gastric carcinoma. *Virchows Arch* 2020; **476**: 353-365 [PMID: 31836926 DOI: 10.1007/s00428-019-02724-4]
 - 76 **Fu Y**, Jung AW, Torne RV, Gonzalez S, Vöhringer H, Shmatko A, Yates LR, Jimenez-Linan M, Moore L, Gerstung M. Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nat Cancer* 2020; **1**: 800-810 [PMID: 35122049 DOI: 10.1038/s43018-020-0085-8]
 - 77 **Kather JN**, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, Krause J, Niehues JM, Sommer KAJ, Bankhead P, Kooreman LFS, Schulte JJ, Cipriani NA, Buelow RD, Boor P, Ortiz-Brüchle NN, Hanby AM, Speirs V, Kochanny S, Patnaik A, Srisuwananukorn A, Brenner H, Hoffmeister M, van den Brandt PA, Jäger D, Trautwein C, Pearson AT, Luedde T. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer* 2020; **1**: 789-799 [PMID: 33763651 DOI: 10.1038/s43018-020-0087-6]
 - 78 **Schmauch B**, Romagnoni A, Pronier E, Saillard C, Maillé P, Calderaro J, Kamoun A, Sefta M, Toldo S, Zaslavskiy M, Clozel T, Moarii M, Courtiol P, Wainrib G. A deep learning model to predict RNA-Seq expression of tumours from whole slide images. *Nat Commun* 2020; **11**: 3877 [PMID: 32747659 DOI: 10.1038/s41467-020-17678-4]
 - 79 **Muti HS**, Heij LR, Keller G, Kohlruss M, Langer R, Dislich B, Cheong JH, Kim YW, Kim H, Kook MC, Cunningham D, Allum WH, Langley RE, Nankivell MG, Quirke P, Hayden JD, West NP, Irvine AJ, Yoshikawa T, Oshima T, Huss R, GROSSER B, Roviello F, d'Ignazio A, Quaas A, Alakus H, Tan X, Pearson AT, Luedde T, Ebert MP, Jäger D, Trautwein C, Gaisa NT, Grabsch HI, Kather JN. 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 [PMID: 34417147 DOI: 10.1016/S2589-7500(21)00133-3]
 - 80 **Kather JN**, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, Marx A, Boor P, Tacke F, Neumann UP, Grabsch HI, Yoshikawa T, Brenner H, Chang-Claude J, Hoffmeister M, Trautwein C, Luedde T. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019; **25**: 1054-1056 [PMID: 31160815 DOI: 10.1038/s41591-019-0462-y]
 - 81 **Kather JN**, Schulte J, Grabsch HI, Loeffler C, Muti H, Dolezal J, Srisuwananukorn A, Agrawal N, Kochanny S, von Stillfried S, Boor P, Yoshikawa T, Jaeger D, Trautwein C, Bankhead P, Cipriani NA, Luedde T, Pearson AT. Deep learning detects virus presence in cancer histology. *bioRxiv* [DOI: 10.1101/690206]
 - 82 **Hinata M**, Ushiku T. Detecting immunotherapy-sensitive subtype in gastric cancer using histologic image-based deep learning. *Sci Rep* 2021; **11**: 22636 [PMID: 34811485 DOI: 10.1038/s41598-021-02168-4]
 - 83 **Zheng X**, Wang R, Zhang X, Sun Y, Zhang H, Zhao Z, Zheng Y, Luo J, Zhang J, Wu H, Huang D, Zhu W, Chen J, Cao Q, Zeng H, Luo R, Li P, Lan L, Yun J, Xie D, Zheng WS, Cai M. A deep learning model and human-machine fusion for prediction of EBV-associated gastric cancer from histopathology. *Nat Commun* 2022; **13**: 2790 [PMID: 35589792 DOI: 10.1038/s41467-022-30459-5]
 - 84 **Montavon G**, Samek W, Müller KR. Methods for interpreting and understanding deep neural networks. *Digit Signal Process* 2018; **73**: 1-15 [DOI: 10.1016/j.dsp.2017.10.011]
 - 85 **Yang JH**, Wright SN, Hamblin M, McCloskey D, Alcantar MA, Schrübbers L, Lopatkin AJ, Satish S, Nili A, Palsson BO, Walker GC, Collins JJ. A White-Box Machine Learning Approach for Revealing Antibiotic Mechanisms of Action. *Cell* 2019; **177**: 1649-1661.e9 [PMID: 31080069 DOI: 10.1016/j.cell.2019.04.016]
 - 86 **Tosun AB**, Pullara F, Becich MJ, Taylor DL, Fine JL, Chennubhotla SC. Explainable AI (xAI) for Anatomic Pathology. *Adv Anat Pathol* 2020; **27**: 241-250 [PMID: 32541594 DOI: 10.1097/PAP.0000000000000264]
 - 87 **Flinner N**, Gretser S, Quaas A, Bankov K, Stoll A, Heckmann LE, Mayer RS, Doering C, Demes MC, Buettner R, Rueschoff J, Wild PJ. Deep learning based on hematoxylin-eosin staining outperforms immunohistochemistry in predicting molecular subtypes of gastric adenocarcinoma. *J Pathol* 2022; **257**: 218-226 [PMID: 35119111 DOI: 10.1002/path.5879]
 - 88 **Jang HJ**, Lee A, Kang J, Song IH, Lee SH. Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach. *World J Gastroenterol* 2021; **27**: 7687-7704 [PMID: 34908807 DOI: 10.3748/wjg.v27.i44.7687]
 - 89 **Guinney J**, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, Simon IM, Gerster S, Fessler E, De Sousa E Melo F, Missiaglia E, Ramay H, Barras D, Homicsko K, Maru D, Manyam GC, Broom B, Boige V, Perez-Villamil B, Laderas T, Salazar R, Gray JW, Hanahan D, Tabernero J, Bernards R, Friend SH, Laurent-Puig P, Medema JP, Sadanandam A, Wessels L, Delorenzi M, Kopetz S, Vermeulen L, Tejpar S. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015; **21**: 1350-1356 [PMID: 26457759 DOI: 10.1038/nm.3967]
 - 90 **Dienstmann R**, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat Rev Cancer* 2017; **17**: 268 [PMID: 28332502 DOI: 10.1038/nrc.2017.24]
 - 91 **Ganesh K**, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, Diaz LA Jr. Immunotherapy in colorectal cancer:

- rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol* 2019; **16**: 361-375 [PMID: 30886395 DOI: 10.1038/s41575-019-0126-x]
- 92 **Li K**, Luo H, Huang L, Zhu X. Microsatellite instability: a review of what the oncologist should know. *Cancer Cell Int* 2020; **20**: 16 [PMID: 31956294 DOI: 10.1186/s12935-019-1091-8]
- 93 **André T**, Shiu KK, Kim TW, Jensen BV, Jensen LH, Punt C, Smith D, Garcia-Carbonero R, Benavides M, Gibbs P, de la Fouchardiere C, Rivera F, Elez E, Bendell J, Le DT, Yoshino T, Van Cutsem E, Yang P, Farooqui MZH, Marinello P, Diaz LA Jr; KEYNOTE-177 Investigators. Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. *N Engl J Med* 2020; **383**: 2207-2218 [PMID: 33264544 DOI: 10.1056/NEJMoa2017699]
- 94 **De Smedt L**, Lemahieu J, Palmans S, Govaere O, Tousseyn T, Van Cutsem E, Prenen H, Tejpar S, Spaepen M, Matthijs G, Decaestecker C, Moles Lopez X, Demetter P, Salmon I, Sagaert X. Microsatellite instable vs stable colon carcinomas: analysis of tumour heterogeneity, inflammation and angiogenesis. *Br J Cancer* 2015; **113**: 500-509 [PMID: 26068398 DOI: 10.1038/bjc.2015.213]
- 95 **Sepulveda AR**, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, Funkhouser WK, Kopetz SE, Lieu C, Lindor NM, Minsky BD, Monzon FA, Sargent DJ, Singh VM, Willis J, Clark J, Colasacco C, Bryan Rumble R, Temple-Smolkin R, B Ventura C, Nowak JA. Molecular Biomarkers for the Evaluation of Colorectal Cancer: Guideline From the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and American Society of Clinical Oncology. *Arch Pathol Lab Med* 2017; **141**: 625-657 [PMID: 28165284 DOI: 10.5858/arpa.2016-0554-CP]
- 96 **Punt CJ**, Koopman M, Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat Rev Clin Oncol* 2017; **14**: 235-246 [PMID: 27922044 DOI: 10.1038/nrclinonc.2016.171]
- 97 **Sirinukunwattana K**, Domingo E, Richman SD, Redmond KL, Blake A, Verrill C, Leedham SJ, Chatzipli A, Hardy C, Whalley CM, Wu CH, Beggs AD, McDermott U, Dunne PD, Meade A, Walker SM, Murray GI, Samuel L, Seymour M, Tomlinson I, Quirke P, Maughan T, Rittscher J, Koelzer VH; S:CORT consortium. Image-based consensus molecular subtype (imCMS) classification of colorectal cancer using deep learning. *Gut* 2021; **70**: 544-554 [PMID: 32690604 DOI: 10.1136/gutjnl-2019-319866]
- 98 **Echle A**, Grabsch HI, Quirke P, van den Brandt PA, West NP, Hutchins GGA, Heij LR, Tan X, Richman SD, Krause J, Alwers E, Jenniskens J, Offermans K, Gray R, Brenner H, Chang-Claude J, Trautwein C, Pearson AT, Boor P, Luedde T, Gaisa NT, Hoffmeister M, Kather JN. Clinical-Grade Detection of Microsatellite Instability in Colorectal Tumors by Deep Learning. *Gastroenterology* 2020; **159**: 1406-1416.e11 [PMID: 32562722 DOI: 10.1053/j.gastro.2020.06.021]
- 99 **Kruger AJ**, Sha L, Kannan M, Joshi RP, Leibowitz BD, Zhang R, Khan AA, Stumpe M. H&E Image-based Consensus Molecular Subtype Classification of Colorectal Cancer Using Weak Labeling. *J Clin Oncol* 2020; **38**: e16097 [DOI: 10.1200/JCO.2020.38.15_suppl.e16097]
- 100 **Popovici V**, Budinská E, Dušek L, Kozubek M, Bosman F. Image-based surrogate biomarkers for molecular subtypes of colorectal cancer. *Bioinformatics* 2017; **33**: 2002-2009 [PMID: 28158480 DOI: 10.1093/bioinformatics/btx027]
- 101 **Cao R**, Yang F, Ma SC, Liu L, Zhao Y, Li Y, Wu DH, Wang T, Lu WJ, Cai WJ, Zhu HB, Guo XJ, Lu YW, Kuang JJ, Huan WJ, Tang WM, Huang K, Huang J, Yao J, Dong ZY. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in Colorectal Cancer. *Theranostics* 2020; **10**: 11080-11091 [PMID: 33042271 DOI: 10.7150/thno.49864]
- 102 **Bilal M**, Raza SEA, Azam A, Graham S, Ilyas M, Cree IA, Snead D, Minhas F, Rajpoot NM. Development and validation of a weakly supervised deep learning framework to predict the status of molecular pathways and key mutations in colorectal cancer from routine histology images: a retrospective study. *Lancet Digit Health* 2021; **3**: e763-e772 [PMID: 34686474 DOI: 10.1016/S2589-7500(21)00180-1]
- 103 **Sun F**, Liu S, Song P, Zhang C, Liu Z, Guan W, Wang M. Impact of retrieved lymph node count on short-term complications in patients with gastric cancer. *World J Surg Oncol* 2020; **18**: 224 [PMID: 32838799 DOI: 10.1186/s12957-020-02000-9]
- 104 **Veta M**, Pluim JP, van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: a review. *IEEE Trans Biomed Eng* 2014; **61**: 1400-1411 [PMID: 24759275 DOI: 10.1109/TBME.2014.2303852]
- 105 **Bhargava R**, Madabhushi A. Emerging Themes in Image Informatics and Molecular Analysis for Digital Pathology. *Annu Rev Biomed Eng* 2016; **18**: 387-412 [PMID: 27420575 DOI: 10.1146/annurev-bioeng-112415-114722]
- 106 **Wang X**, Chen Y, Gao Y, Zhang H, Guan Z, Dong Z, Zheng Y, Jiang J, Yang H, Wang L, Huang X, Ai L, Yu W, Li H, Dong C, Zhou Z, Liu X, Yu G. Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nat Commun* 2021; **12**: 1637 [PMID: 33712598 DOI: 10.1038/s41467-021-21674-7]
- 107 **Hu Y**, Su F, Dong K, Wang X, Zhao X, Jiang Y, Li J, Ji J, Sun Y. Deep learning system for lymph node quantification and metastatic cancer identification from whole-slide pathology images. *Gastric Cancer* 2021; **24**: 868-877 [PMID: 33484355 DOI: 10.1007/s10120-021-01158-9]
- 108 **Matsushima J**, Sato T, Ohnishi T, Yoshimura Y, Mizutani H, Koto S, Ikeda JI, Kano M, Matsubara H, Hayashi H. The Use of Deep Learning-Based Computer Diagnostic Algorithm for Detection of Lymph Node Metastases of Gastric Adenocarcinoma. *Int J Surg Pathol* 2022; 10668969221113475 [PMID: 35898183 DOI: 10.1177/10668969221113475]
- 109 **Huang SC**, Chen CC, Lan J, Hsieh TY, Chuang HC, Chien MY, Ou TS, Chen KH, Wu RC, Liu YJ, Cheng CT, Huang YJ, Tao LW, Hwu AF, Lin IC, Hung SH, Yeh CY, Chen TC. Deep neural network trained on gigapixel images improves lymph node metastasis detection in clinical settings. *Nat Commun* 2022; **13**: 3347 [PMID: 35688834 DOI: 10.1038/s41467-022-30746-1]
- 110 **Kwak MS**, Lee HH, Yang JM, Cha JM, Jeon JW, Yoon JY, Kim HI. Deep Convolutional Neural Network-Based Lymph Node Metastasis Prediction for Colon Cancer Using Histopathological Images. *Front Oncol* 2020; **10**: 619803 [PMID: 33520727 DOI: 10.3389/fonc.2020.619803]
- 111 **Pai RK**, Hartman D, Schaeffer DF, Rosty C, Shivji S, Kirsch R, Pai RK. Development and initial validation of a deep learning algorithm to quantify histological features in colorectal carcinoma including tumour budding/poorly differentiated clusters. *Histopathology* 2021; **79**: 391-405 [PMID: 33590485 DOI: 10.1111/his.14353]
- 112 **Kiehl L**, Kuntz S, Höhn J, Jutzi T, Krieghoff-Henning E, Kather JN, Holland-Letz T, Kopp-Schneider A, Chang-Claude J,

- Brobeil A, von Kalle C, Fröhling S, Alwers E, Brenner H, Hoffmeister M, Brinker TJ. Deep learning can predict lymph node status directly from histology in colorectal cancer. *Eur J Cancer* 2021; **157**: 464-473 [PMID: [34649117](#) DOI: [10.1016/j.ejca.2021.08.039](#)]
- 113 **van Pelt GW**, Kjær-Frifeldt S, van Krieken JHJM, Al Dieri R, Morreau H, Tollenaar RAEM, Sørensen FB, Mesker WE. Scoring the tumor-stroma ratio in colon cancer: procedure and recommendations. *Virchows Arch* 2018; **473**: 405-412 [PMID: [30030621](#) DOI: [10.1007/s00428-018-2408-z](#)]
- 114 **Fu M**, Chen D, Luo F, Li M, Wang Y, Chen J, Li A, Liu S. Association of the tumour stroma percentage in the preoperative biopsies with lymph node metastasis in colorectal cancer. *Br J Cancer* 2020; **122**: 388-396 [PMID: [31787749](#) DOI: [10.1038/s41416-019-0671-7](#)]
- 115 **Geessink OGF**, Baidoshvili A, Klaase JM, Ehteshami Bejnordi B, Litjens GJS, van Pelt GW, Mesker WE, Nagtegaal ID, Ciompi F, van der Laak JAWM. Computer aided quantification of intratumoral stroma yields an independent prognosticator in rectal cancer. *Cell Oncol (Dordr)* 2019; **42**: 331-341 [PMID: [30825182](#) DOI: [10.1007/s13402-019-00429-z](#)]
- 116 **Hong Y**, Heo YJ, Kim B, Lee D, Ahn S, Ha SY, Sohn I, Kim KM. Deep learning-based virtual cytokeratin staining of gastric carcinomas to measure tumor-stroma ratio. *Sci Rep* 2021; **11**: 19255 [PMID: [34584193](#) DOI: [10.1038/s41598-021-98857-1](#)]
- 117 **Kather JN**, Poleszczuk J, Suarez-Carmona M, Krisam J, Charoentong P, Valous NA, Weis CA, Tavernar L, Leiss F, Herpel E, Klupp F, Ulrich A, Schneider M, Marx A, Jäger D, Halama N. In Silico Modeling of Immunotherapy and Stroma-Targeting Therapies in Human Colorectal Cancer. *Cancer Res* 2017; **77**: 6442-6452 [PMID: [28923860](#) DOI: [10.1158/0008-5472.CAN-17-2006](#)]
- 118 **Scheer R**, Baidoshvili A, Zoidze S, Elferink MAG, Berkel AEM, Klaase JM, van Diest PJ. Tumor-stroma ratio as prognostic factor for survival in rectal adenocarcinoma: A retrospective cohort study. *World J Gastrointest Oncol* 2017; **9**: 466-474 [PMID: [29290917](#) DOI: [10.4251/wjgo.v9.i12.466](#)]
- 119 **Kather JN**, Halama N, Jaeger D. Genomics and emerging biomarkers for immunotherapy of colorectal cancer. *Semin Cancer Biol* 2018; **52**: 189-197 [PMID: [29501787](#) DOI: [10.1016/j.semcancer.2018.02.010](#)]
- 120 **Weis CA**, Kather JN, Melchers S, Al-Ahmdi H, Pollheimer MJ, Langner C, Gaiser T. Automatic evaluation of tumor budding in immunohistochemically stained colorectal carcinomas and correlation to clinical outcome. *Diagn Pathol* 2018; **13**: 64 [PMID: [30153844](#) DOI: [10.1186/s13000-018-0739-3](#)]
- 121 **Kather JN**, Krisam J, Charoentong P, Luedde T, Herpel E, Weis CA, Gaiser T, Marx A, Valous NA, Ferber D, Jansen L, Reyes-Aldasoro CC, Zörnig I, Jäger D, Brenner H, Chang-Claude J, Hoffmeister M, Halama N. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS Med* 2019; **16**: e1002730 [PMID: [30677016](#) DOI: [10.1371/journal.pmed.1002730](#)]
- 122 **Shapcott M**, Hewitt KJ, Rajpoot N. Deep Learning With Sampling in Colon Cancer Histology. *Front Bioeng Biotechnol* 2019; **7**: 52 [PMID: [30972333](#) DOI: [10.3389/fbioe.2019.00052](#)]
- 123 **Sirinukunwattana K**, Snead D, Epstein D, Aftab Z, Mujeeb I, Tsang YW, Cree I, Rajpoot N. Novel digital signatures of tissue phenotypes for predicting distant metastasis in colorectal cancer. *Sci Rep* 2018; **8**: 13692 [PMID: [30209315](#) DOI: [10.1038/s41598-018-31799-3](#)]
- 124 **Swiderska-Chadaj Z**, Pinckaers H, van Rijthoven M, Balkenhol M, Melnikova M, Geessink O, Manson Q, Sherman M, Polonia A, Parry J, Abubakar M, Litjens G, van der Laak J, Ciompi F. Learning to detect lymphocytes in immunohistochemistry with deep learning. *Med Image Anal* 2019; **58**: 101547 [PMID: [31476576](#) DOI: [10.1016/j.media.2019.101547](#)]
- 125 **Zhao K**, Li Z, Yao S, Wang Y, Wu X, Xu Z, Wu L, Huang Y, Liang C, Liu Z. Artificial intelligence quantified tumour-stroma ratio is an independent predictor for overall survival in resectable colorectal cancer. *EBioMedicine* 2020; **61**: 103054 [PMID: [33039706](#) DOI: [10.1016/j.ebiom.2020.103054](#)]
- 126 **Zhao K**, Wu L, Huang Y, Yao S, Xu Z, Lin H, Wang H, Liang Y, Xu Y, Chen X, Zhao M, Peng J, Liang C, Li Z, Li Y, Liu Z. Deep learning quantified mucus-tumor ratio predicting survival of patients with colorectal cancer using whole-slide images. *Precis Clin Med* 2021; **4**: 17-24 [PMID: [35693123](#) DOI: [10.1093/pccmedi/pbab002](#)]
- 127 **Meier A**, Nekolla K, Hewitt LC, Earle S, Yoshikawa T, Oshima T, Miyagi Y, Huss R, Schmidt G, Grabsch HI. Hypothesis-free deep survival learning applied to the tumour microenvironment in gastric cancer. *J Pathol Clin Res* 2020; **6**: 273-282 [PMID: [32592447](#) DOI: [10.1002/cjp2.170](#)]
- 128 **Huang B**, Tian S, Zhan N, Ma J, Huang Z, Zhang C, Zhang H, Ming F, Liao F, Ji M, Zhang J, Liu Y, He P, Deng B, Hu J, Dong W. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: A retrospective multicentre study. *EBioMedicine* 2021; **73**: 103631 [PMID: [34678610](#) DOI: [10.1016/j.ebiom.2021.103631](#)]
- 129 **Jiang Y**, Xie J, Han Z, Liu W, Xi S, Huang L, Huang W, Lin T, Zhao L, Hu Y, Yu J, Zhang Q, Li T, Cai S, Li G. Immunomarker Support Vector Machine Classifier for Prediction of Gastric Cancer Survival and Adjuvant Chemotherapeutic Benefit. *Clin Cancer Res* 2018; **24**: 5574-5584 [PMID: [30042208](#) DOI: [10.1158/1078-0432.CCR-18-0848](#)]
- 130 **Jiang Y**, Xie J, Huang W, Chen H, Xi S, Han Z, Huang L, Lin T, Zhao LY, Hu YF, Yu J, Cai SR, Li T, Li G. Tumor Immune Microenvironment and Chemosensitivity Signature for Predicting Response to Chemotherapy in Gastric Cancer. *Cancer Immunol Res* 2019; **7**: 2065-2073 [PMID: [31615816](#) DOI: [10.1158/2326-6066.CIR-19-0311](#)]
- 131 **Wang Y**, Wang YG, Hu C, Li M, Fan Y, Otter N, Sam I, Gou H, Hu Y, Kwok T, Zalcborg J, Boussioutas A, Daly RJ, Montúfar G, Liò P, Xu D, Webb GI, Song J. Cell graph neural networks enable the precise prediction of patient survival in gastric cancer. *NPJ Precis Oncol* 2022; **6**: 45 [PMID: [35739342](#) DOI: [10.1038/s41698-022-00285-5](#)]
- 132 **Bychkov D**, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, Walliander M, Lundin M, Haglund C, Lundin J. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 2018; **8**: 3395 [PMID: [29467373](#) DOI: [10.1038/s41598-018-21758-3](#)]
- 133 **Skrede OJ**, De Raedt S, Kleppe A, Hveem TS, Liestøl K, Maddison J, Askautrud HA, Pradhan M, Nesheim JA, Albrechtsen F, Farstad IN, Domingo E, Church DN, Nesbakken A, Shepherd NA, Tomlinson I, Kerr R, Novelli M, Kerr

- DJ, Danielsen HE. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* 2020; **395**: 350-360 [PMID: 32007170 DOI: 10.1016/S0140-6736(19)32998-8]
- 134 **Jiang D**, Liao J, Duan H, Wu Q, Owen G, Shu C, Chen L, He Y, Wu Z, He D, Zhang W, Wang Z. A machine learning-based prognostic predictor for stage III colon cancer. *Sci Rep* 2020; **10**: 10333 [PMID: 32587295 DOI: 10.1038/s41598-020-67178-0]
- 135 **Kotsenas AL**, Balthazar P, Andrews D, Geis JR, Cook TS. Rethinking Patient Consent in the Era of Artificial Intelligence and Big Data. *J Am Coll Radiol* 2021; **18**: 180-184 [PMID: 33413897 DOI: 10.1016/j.jacr.2020.09.022]
- 136 **Inoue T**, Yagi Y. Color standardization and optimization in whole slide imaging. *Clin Diagn Pathol* 2020; **4** [PMID: 33088926 DOI: 10.15761/cdp.1000139]
- 137 **Yoshida H**, Yokota H, Singh R, Kiyuna T, Yamaguchi M, Kikuchi S, Yagi Y, Ochiai A. Meeting Report: The International Workshop on Harmonization and Standardization of Digital Pathology Image, Held on April 4, 2019 in Tokyo. *Pathobiology* 2019; **86**: 322-324 [PMID: 31707388 DOI: 10.1159/000502718]
- 138 **Janowczyk A**, Basavanhally A, Madabhushi A. Stain Normalization using Sparse AutoEncoders (StaNoSA): Application to digital pathology. *Comput Med Imaging Graph* 2017; **57**: 50-61 [PMID: 27373749 DOI: 10.1016/j.compmedimag.2016.05.003]
- 139 **Vahadane A**, Peng T, Sethi A, Albarqouni S, Wang L, Baust M, Steiger K, Schlitter AM, Esposito I, Navab N. Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images. *IEEE Trans Med Imaging* 2016; **35**: 1962-1971 [PMID: 27164577 DOI: 10.1109/TMI.2016.2529665]
- 140 **Janowczyk A**, Zuo R, Gilmore H, Feldman M, Madabhushi A. HistoQC: An Open-Source Quality Control Tool for Digital Pathology Slides. *JCO Clin Cancer Inform* 2019; **3**: 1-7 [PMID: 30990737 DOI: 10.1200/CCI.18.00157]
- 141 **Senaras C**, Niazi MKK, Lozanski G, Gurcan MN. DeepFocus: Detection of out-of-focus regions in whole slide digital images using deep learning. *PLoS One* 2018; **13**: e0205387 [PMID: 30359393 DOI: 10.1371/journal.pone.0205387]
- 142 **Arora A**, Arora A. Pathology training in the age of artificial intelligence. *J Clin Pathol* 2021; **74**: 73-75 [PMID: 33020175 DOI: 10.1136/jclinpath-2020-207110]
- 143 **Hekler A**, Utikal JS, Enk AH, Solass W, Schmitt M, Klode J, Schadendorf D, Sondermann W, Franklin C, Bestvater F, Flaig MJ, Krahl D, von Kalle C, Fröhling S, Brinker TJ. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *Eur J Cancer* 2019; **118**: 91-96 [PMID: 31325876 DOI: 10.1016/j.ejca.2019.06.012]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: <https://www.f6publishing.com/helpdesk>
<https://www.wjgnet.com>

