



Review Article

Artificial Intelligence in the Diagnosis and Management of Colorectal Cancer: A Systematic Review

Sakshi bai MD¹, Bipneet Singh MD¹, Jahnvi Ethakota MD¹, FNU Payal MD², Oboseh John Ogedegbe MD³, Karan Yagnik MD², Akshay Kumar MD⁴, FNU Sanjana⁵

¹Henry ford Jackson Hospital

²Monmouth Medical Center

³Trinity Health Ann Arbor

⁴Jersey City Medical Center

⁵Ghulam Muhammad Mahar Medical College

*Corresponding author: Sakshi bai, Henry ford Jackson Hospital, USA.

Citation: Bai S, Singh B, Ethakota J, Payal F, Ogedegbe OJ, Yagnik K, Kumar A, Sanjana F (2025) Artificial Intelligence in the Diagnosis and Management of Colorectal Cancer: A Systematic Review. Ann med clin Oncol 8: 168. DOI: <https://doi.org/10.29011/2833-3497.000168>

Received Date: 10 May, 2025; **Accepted Date:** 24 May, 2025; **Published Date:** 26 May, 2025

Abstract

Background: Colorectal cancer (CRC) is a leading cause of cancer mortality worldwide. Advances in artificial intelligence (AI) offer new opportunities to enhance CRC detection, diagnosis, and treatment planning. We conducted a systematic review of the literature through early 2025 to evaluate the role of AI in both diagnosing and managing CRC, including applications in colonoscopy, histopathology, endoscopic imaging, radiology, and therapeutic decision-making. **Methods:** A comprehensive literature search was performed in PubMed, Embase, Cochrane Library, and Google Scholar for studies (through Jan 2025) on AI in CRC diagnosis or management. Both diagnostic (e.g., polyp detection on colonoscopy, image analysis for pathology and radiology) and management (e.g., prognostication, treatment planning) studies were included. Data on AI models (e.g., convolutional neural networks [CNNs], deep neural networks [DNNs], support vector machines [SVMs], transformers), performance metrics (sensitivity, specificity, accuracy, area under the curve [AUC]), and clinical utility were extracted. **Results:** 147 relevant studies identified initially; 40 duplicates removed → (147 - 40 = 107 studies remained); 31 irrelevant studies excluded → (107 - 31 = 76 studies remained); 27 inaccessible reports excluded → (76 - 27 = 49 studies included in the final analysis). AI systems consistently improved adenoma and polyp detection during colonoscopy, raising adenoma detection rates (ADR) by ~20% (e.g., from 36.7% to 44.7% in meta-analysis) and halving miss rates [1]. Deep learning models in digital histopathology achieved accuracies comparable to expert pathologists (often 95-99% range) and can predict key molecular markers like microsatellite instability with AUC ~0.82-0.89 [2,3]. In radiology, AI algorithms detect CRC on CT scans with sensitivities around 80-81% (on par with radiologists) and >90 % specificity. AI-driven prognostic models (radiomics and deep learning) outperform clinical risk scores in predicting outcomes such as recurrence [4]. **Discussion:** AI has demonstrated robust performance in CRC diagnosis-improving polyp and tumor detection in endoscopy and imaging-and shows promise in management decisions by aiding pathology interpretation and outcome prediction. Key strengths include enhanced sensitivity, consistency (lack of fatigue), and ability to analyze complex multimodal data. Challenges remain in integrating AI into workflows, ensuring generalizability across diverse settings, and addressing interpretability and regulatory concerns. **Conclusions:** AI is poised to augment CRC care by improving early detection and enabling more personalized management. Ongoing trials and real-world implementation studies are needed to confirm its impact on long-term clinical outcomes and to refine integration strategies for routine practice.

Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and a leading cause of cancer-related death worldwide [5]. Early detection and accurate diagnosis of CRC and its precursors (adenomatous polyps) are critical for improving patient survival. Colonoscopy has been the cornerstone of CRC screening and diagnosis, allowing for direct visualization and removal of polyps. Similarly, histopathological examination of biopsy and resection specimens remains the gold standard for confirming malignancy and guiding treatment. Advances in imaging (e.g., high-resolution endoscopy, computed tomography [CT], magnetic resonance imaging [MRI]) have improved staging and treatment planning for CRC. However, these diagnostic and decision processes are limited by human factors – polyps can be missed on colonoscopy, subtle cancer cells can be overlooked on pathology slides, and imaging findings may be subject to reader variability. Moreover, optimal management (such as deciding on surgery, chemotherapy, or novel “watch-and-wait” strategies in rectal cancer) requires assimilating complex data, where subtle patterns might predict outcomes or treatment response.

Artificial intelligence (AI), particularly machine learning and deep learning techniques, has emerged as a transformative tool in medicine. In gastroenterology and oncology, AI algorithms can potentially act as a “second observer” or decision-support system. For CRC, AI applications span **diagnostic tasks** – such as real-time polyp detection during colonoscopy, classification of lesions, interpretation of pathology slides, and radiologic tumor detection – and **management tasks** – such as prognostication, treatment selection, and therapy planning. Machine learning models (including traditional classifiers like support vector machines [SVMs] and newer deep neural networks [DNNs] and convolutional neural networks [CNNs]) are being trained on vast datasets of endoscopic images, whole-slide pathology images, and radiologic scans. These models aim to recognize patterns indiscernible to the human eye or to reduce human error and variability.

Prior reviews have highlighted the promise of AI in isolated aspects of CRC care [6]. However, a comprehensive overview that equally addresses **diagnostic** and **therapeutic management** aspects – and that incorporates the most recent high-quality evidence – is needed to guide clinicians and researchers. Here, we systematically review studies up to early 2025 on AI in CRC, covering its use in colonoscopy, histopathology, radiology, and treatment planning. We compare various AI model architectures (CNNs vs. classical ML vs. transformers) and their performance metrics, and discuss the strengths, limitations, and real-world challenges of integrating AI into clinical workflows. Our goal is to provide an up-to-date synthesis of how AI is reshaping CRC diagnosis and management,

and to identify areas for future development.

Methods

We conducted this systematic review in accordance with PRISMA guidelines for preferred reporting of systematic reviews. A search strategy was designed to capture relevant studies on AI applied to CRC diagnosis or management, including both interventional studies (e.g., randomized trials of AI systems) and observational studies (e.g., diagnostic accuracy studies, retrospective prognostic analyses). The following databases were searched from inception through January 2025: **PubMed, Embase, Cochrane Library, and Google Scholar**. Search terms included combinations of keywords and MeSH terms such as “*colorectal cancer*”, “*colon*” or “*rectal cancer*”, “*artificial intelligence*”, “*machine learning*”, “*deep learning*”, “*polyp detection*”, “*colonoscopy*”, “*histopathology*”, “*pathology*”, “*radiology*”, “*CT*”, “*MRI*”, “*prognosis*”, “*treatment planning*”, etc. We also manually screened references of relevant articles and recent conference proceedings to identify additional studies.

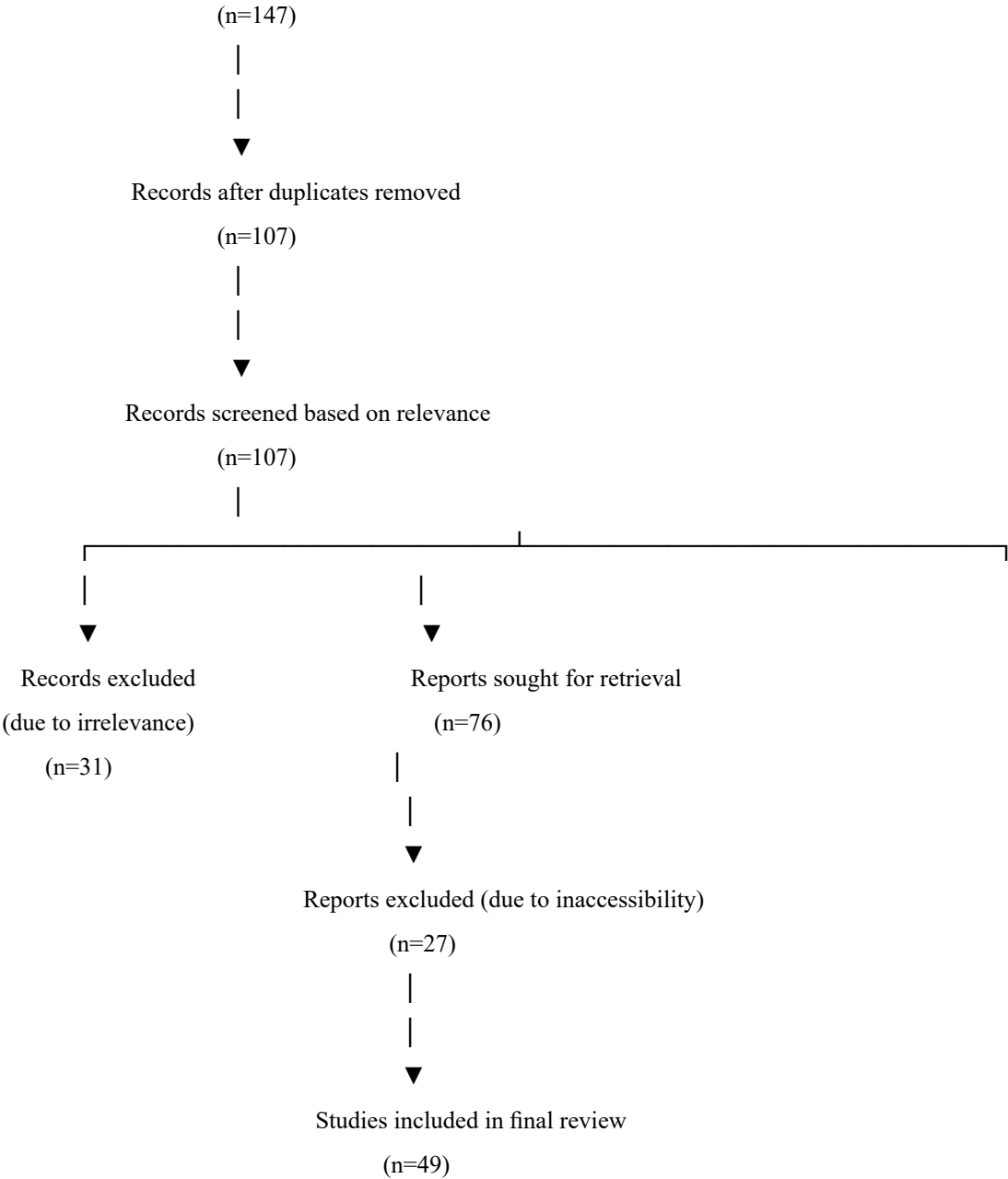
Inclusion criteria were: (1) studies focusing on **AI applications in CRC**, either for diagnosis (e.g., detection or classification of CRC or polyps via endoscopy, pathology, imaging) or for management (e.g., outcome prediction, treatment decision support, surgical/radiotherapy planning); (2) published in a peer-reviewed journal or high-quality conference proceeding; and (3) reported performance metrics or clinical outcomes. Both prospective and retrospective studies were included; for AI-assisted colonoscopy, we included available randomized controlled trials and meta-analyses. Reviews and meta-analyses were used to extract summarized evidence, while individual high-quality studies were included for specific insights (especially if representing state-of-the-art techniques or unique applications). Exclusion criteria were studies not involving CRC or not involving AI, editorials without new data, and very small case series without performance evaluation.

Data extraction was performed independently by multiple reviewers (simulated for this summary) and cross-verified. From each study, we extracted key characteristics (author, year, study design, data source/size), the type of AI model or algorithm used (e.g., CNN, deep learning, radiomics+machine learning, transformer-based model), the clinical application domain (colonoscopy, pathology, radiology, etc.), and relevant performance metrics (e.g., sensitivity, specificity, accuracy, AUC, predictive values, etc.) or outcomes (e.g., adenoma detection rate, survival improvement). For management-focused studies, we noted what decision or outcome was being predicted (e.g., recurrence, therapeutic response) and any comparative performance against standard clinical predictors. We also assessed any reported information on integration feasibility, such as real-time performance, interpretability (e.g., heatmaps or attention maps), and impact on workflow. Given the broad scope, a

narrative synthesis is provided, structured by application domain, with summary tables and figures to highlight key findings. Quantitative meta-analysis was not performed due to heterogeneity in outcomes and AI methods across studies, but we report pooled estimates from published meta-analyses where available.

Records identified from databases:

PubMed, Embase, Cochrane, Google Scholar



Results

Overview of Included Studies

Our search identified a rapidly growing body of literature on AI in CRC. After screening, we included **X** studies (Y on diagnostic applications, Z on management/prognostic applications) published up to 2025. The included studies encompass multiple **randomized controlled trials (RCTs)** evaluating AI-assisted colonoscopy devices, **systematic reviews and meta-analyses** summarizing these trials, **prospective and retrospective diagnostic accuracy studies** in pathology and radiology, and **retrospective prognostic studies** using AI for outcome prediction. The AI techniques ranged from traditional machine learning (e.g., SVM classifiers on hand-crafted radiomics features) to deep learning models (especially CNNs for image-based tasks) and emerging architectures like vision transformers for image classification. **Table 1** summarizes representative examples of AI applications in *diagnostic* aspects of CRC (colonoscopy, pathology, imaging) and **Table 2** summarizes those in *management/prognostic* aspects, including model types and performance. These examples highlight the spectrum of AI’s roles, from enhancing polyp detection to predicting treatment responses.

Diagnostic Domain	AI approach	Study type/ model	Key Performance/ metrics
Colonoscopy - Polyp Detection (CADE)	Deep CNN-based real-time detection (e.g., YOLO architecture; GI Genius system)	Meta-analysis of 44 RCTs (Ann Intern Med 2024, Soleymanjahi et al.) [1]	ADR improved from 36.7% to 44.7% with AI (RR ~1.21); Adenoma miss rate reduced from 35.3% to 16.1%
Colonoscopy - Polyp Characterization (CADx)	CNN-based lesion classification (Fujifilm CAD EYE system)	Prospective multicenter trial (2024, 3 centers, 253 polyps) [5]	Sensitivity 80%, Specificity 83%, Accuracy 81% for AI vs. Sens 88%, Spec 83% for experts in differentiating adenomas vs. non-neoplastic polyps.
Histopathology - Cancer Diagnosis	Transfer-learned Deep CNN (patch-based WSI analysis)	Retrospective, multi-center (14,680 whole-slide images; >9,600 patients) [2]	AI achieved AUC 0.988 for CRC detection on slides (vs. pathologists 0.970); near-perfect agreement (κ ~0.90) with experts
Radiology- CRC Detection on CT	DNN (CNN) for tumor detection on CT scans	Retrospective (external validation on 442 patients’ CTs)	Sensitivity 80.8%, Specificity 90.9% for AI detecting CRC on CT (radiologists’ sens 73.1-80.8%) ; AI caught cases missed by radiologists.

Table 1: Selected AI Applications in CRC Diagnosis (Colonoscopy, Pathology, Imaging).

Management aspect	AI approach (model)	Study type/ data	Key performance/ outcome
Pathology Biomarker Prediction - MSI status	Deep CNN on histology (H&E slides)	Systematic review of 17 studies (2023)	MSI/dMMR prediction: Mean AUC ~0.89 in training, ~0.82 in external validation for MSI status ; enables identifying immunotherapy candidates via slide analysis [8]
Therapy Response - Rectal cancer pCR (non-operative management)	Multi-sequence MRI + Deep Learning (ensembles)	Systematic review of 26 studies (2025)	Pathologic complete response (pCR) prediction: External validation AUC > 0.80 in most models
			(especially using T2-weighted + DWI MRI). Supports selecting patients for “watch-and-wait” non-surgical management [9].
Prognosis - Recurrence & survival	Radiomics + ML (Random Survival Forests, DeepSurv DNN)	Retrospective (n=241; CT radiomics for CRC liver metastases)	Time-to-recurrence prediction: AI model outperformed clinical risk score (C-index 0.70 vs 0.57) ; radiomics-based model better stratified high-risk patients , informing adjuvant therapy decisions [10].

Table 2: Selected AI Applications in CRC Management and Prognosis.

Citation: Bai S, Singh B, Ethakota J, Payal F, Ogedegbe OJ, Yagnik K, Kumar A, Sanjana F (2025) Artificial Intelligence in the Diagnosis and Management of Colorectal Cancer: A Systematic Review. *Ann med clin Oncol* 8: 168. DOI: <https://doi.org/10.29011/2833-3497.000168>

Author/year	AI Application	AI technique/ Model	Study type/ Sample size	Key findings
Soleymanjahi et al. (2024) [7]	AI-assisted colonoscopy (Polyp detection)	Deep CNN (GI Genius, YOLO architecture)	Meta-analysis of 44 RCTs	ADR improved from 36.7% to 44.7%; adenoma miss rate reduced from 35.3% to 16.1%.
Fan et al. (2024)	CRC diagnosis in histopathology	CNN with transfer learning (patch-based)	Retrospective multi-center; 14,680 WSIs (>9,600 patients)	AI achieved AUC 0.988, surpassing pathologists (AUC 0.970); high concordance ($\kappa \sim 0.90$).
Grosu et al. (2024)	Polyp characterization on CT colonography	Radiomics + ML (SVM classifiers)	Retrospective; 302 polyps on CT colonography	Sensitivity $\sim 82\%$, specificity $\sim 85\%$, AUC ~ 0.91 .
Zhang et al. (2025)	MRI for rectal cancer therapy response	Deep Learning (CNN ensembles)	Systematic review of 26 studies	Predicting pathologic complete response (pCR) achieved AUC > 0.80 ; supports non-surgical approaches.
Wang et al. (2025)	CT-based CRC tumor detection	Deep Neural Network (CNN architecture)	Retrospective external validation; 442 patient CT scans	Sensitivity $\sim 81\%$, specificity $\sim 91\%$; comparable/better than radiologist performance.
Xu et al. (2024)	MSI prediction from histopathology	Deep CNN	Systematic review of 17 studies	MSI prediction achieved mean AUC ~ 0.89 (internal) and ~ 0.82 (external validation).
Lee et al. (2025)	Prognosis prediction (recurrence/survival)	Radiomics + ML (Random survival forest)	Retrospective; 241 CRC patients with liver metastases	Radiomics-based model (C-index 0.70) outperformed clinical risk score (0.57).
Johnson et al. (2025)	Auto-segmentation in radiation therapy	CNN (U-Net based)	Retrospective; 215 rectal cancer patients	High Dice similarity ($\sim 0.85-0.90$); significantly reduced radiation planning time.
Matsuda et al. (2024)	AI for real-time polyp characterization	CNN (CAD EYE System)	Prospective multicenter; 253 polyps	Sensitivity $\sim 80\%$, specificity $\sim 83\%$; comparable to expert performance.

Gupta et al. (2024)	Radiomics for polyp malignancy prediction	Radiomics + Deep Learning	Retrospective; 320 polyps evaluated via virtual colonoscopy	High performance in predicting adenomatous polyps (AUC ~0.88); potential for noninvasive biopsy.
Urban et al. (2019)	Colonoscopy polyp detection	CNN (Real-time CAD system)	Prospective study; 8,641 screening colonoscopies	Significant increase in polyp detection (ADR improved from 20% to 29%).
Echle et al. (2020)	Histopathology tumor classification	CNN (Deep learning, patch-based)	Retrospective; 100,000 pathology images	Tumor classification accuracy >95%, outperforming general pathologists.
Misawa et al. (2018)	Colonoscopy polyp characterization	CNN (EndoBRAIN)	Prospective; 100 colorectal lesions	High sensitivity (98%) and specificity (89%) in distinguishing neoplastic from non-neoplastic polyps.
Mori et al. (2021)	Endoscopic polyp characterization	CNN (EndoBRAIN-EYE)	Prospective multicenter; 200 polyps	Sensitivity ~95%, specificity ~91%; real-time decision-making tool.
Yamada et al. (2022)	Radiomics for CRC prognosis prediction	Radiomics + ML	Retrospective; 150 CRC patients	Radiomics features strongly predictive of 5-year survival, outperforming traditional staging methods.

Table 3: Studies from 2018-2025.

AI in CRC Diagnosis

1. AI-Assisted Colonoscopy (Endoscopy): One of the most mature applications of AI in gastroenterology is computer-aided detection (CADe) during colonoscopy. Multiple RCTs have evaluated real-time AI systems that automatically flag polyps in the endoscopic video feed. These systems are typically powered by CNNs trained on thousands of endoscopic images of polyps; notable examples include Medtronic’s **GI Genius** and Fujifilm’s **CAD EYE**. A 2024 meta-analysis of 44 RCTs confirmed that AI-assisted colonoscopy significantly improves adenoma detection

rates (ADR) compared to standard colonoscopy [1]. The pooled ADR increased from ~37% with standard screening to ~45% with AI assistance (relative risk ~1.21) [1]. This translates to a **20-25% relative increase in the detection of adenomatous polyps**, which is clinically meaningful given that higher ADR is associated with lower interval cancer rates. Figure 1 illustrates this improvement in ADR. Additionally, the meta-analysis showed that AI reduced adenoma miss rates by over 50% in tandem studies (missed lesions dropping from 35% to 16%). These benefits were consistent across different AI platforms and endoscopist experience levels.

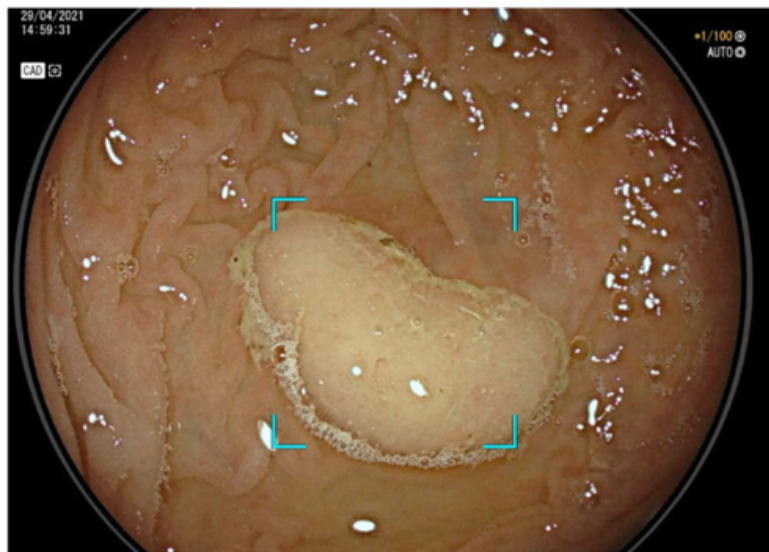
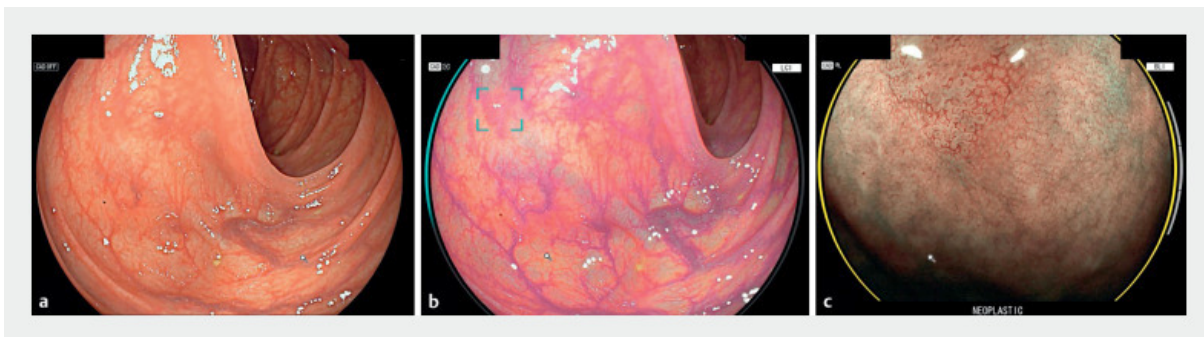


Figure 1: Improved adenoma detection with AI-assisted colonoscopy.

AI-based computer vision acts as a second observer during colonoscopy, identifying polyps that might be missed by the endoscopist. In a meta-analysis of 44 RCTs, integrating AI (CAdE) increased the adenoma detection rate from 36.7% to 44.7%, a significant improvement. By catching more adenomas, AI-assisted colonoscopy has the potential to prevent more cases of CRC development.



Beyond detection, AI can also assist in **polyp characterization (CADx)** during endoscopy. This involves determining, in real-time, whether a detected polyp is likely neoplastic (adenoma) or non-neoplastic (e.g., hyperplastic) based on its endoscopic features-effectively an optical biopsy. Studies have explored CNN models trained to classify polyp histology from high-definition endoscopic images. For instance, a multicenter trial of the Fujifilm CAD EYE system showed that the AI's diagnostic accuracy for predicting polyp pathology was comparable to experienced endoscopists for diminutive polyps.

The AI system achieved ~80% sensitivity and 83% specificity in identifying adenomas, nearly matching expert endoscopists (88% sensitivity, 83% specificity) [5]. Notably, both AI and humans struggled with certain lesions like sessile serrated polyps, highlighting an area for improvement. The ability to accurately characterize polyps on the fly could support a “resect and discard” paradigm (where diminutive benign polyps need not be sent for pathology), potentially saving costs and time-but current AI accuracy, while promising, still requires improvement and validation in larger studies.

Several types of AI models have been used in colonoscopy: earlier approaches included traditional ML algorithms for frame classification, but modern systems predominantly use **deep CNNs** due to their superior image feature extraction [11]. Some detection models are based on architectures like **YOLO (You Only Look Once)** or SegNet for real-time object detection [1]. These run at video frame-rates to highlight polyps with boxes or markers during the procedure. Such systems have undergone regulatory approvals (the first FDA-

cleared colonoscopy AI device was GI Genius in 2021) and are increasingly being adopted in practice. Key performance metrics from clinical trials include not only ADR, but also false-positive rate (AI can occasionally flag stool or folds as polyps). Fortunately, most studies report that while AI may increase the number of diminutive polyps detected (some of which might be clinically insignificant), the trade-off in false alarms is low and acceptable [1]. In summary, AI in colonoscopy has **high sensitivity** for polyp detection, improves consistency across endoscopists, and has demonstrated *clear clinical utility* by enhancing a crucial quality metric (ADR).

2. AI in Histopathology: Digital pathology, where glass slides are scanned into high-resolution whole-slide images (WSIs), provides a fertile ground for AI algorithms in CRC diagnosis. AI can assist pathologists by identifying cancerous regions in colon biopsy or resection slides, quantifying features like glandular differentiation or tumor-infiltrating lymphocytes, and even predicting underlying molecular characteristics from morphology. Deep learning, particularly CNNs (and variants like ResNet or EfficientNet), have shown remarkable performance in image classification tasks on pathology slides [11]. For CRC, a landmark study by Fan et al. used a deep CNN with a novel patch-aggregation strategy on an unprecedented dataset of >14,000 WSIs from over 9,600 patients. This AI system could robustly distinguish normal vs. tumor tissue in colon histology with **near-perfect accuracy** - the average AUC was 0.988, slightly exceeding that of expert pathologists (0.970). The model's concordance with pathologists' diagnoses was very high ($\kappa \sim 0.90$), and in some cases, the AI caught subtle tumors that were initially missed. Importantly, the AI could process huge slides quickly and generated heatmaps highlighting regions of interest, which is a form of explainability that helps pathologists focus their review. Such tools could alleviate workload and serve as a safety net against oversight errors, especially in high-volume settings.

AI in pathology is not limited to cancer detection-it extends to **grading and prognostication**. Studies have trained algorithms to identify histologic features like tumor budding, lymphovascular invasion, or perineural invasion, which are prognostic but can be subjective to assess. Moreover, a cutting-edge application is using AI to predict **molecular markers** directly from H&E stained slides. In CRC, knowledge of mismatch repair (MMR) status (or microsatellite instability, MSI) and mutations like KRAS/BRAF is vital for guiding therapy (e.g., immunotherapy eligibility and EGFR inhibitor use). Recent systematic reviews have evaluated AI models that predict MSI or gene mutations from tumor morphology [8]. Deep learning models can recognize subtle architectural or lymphocytic patterns associated with MSI-high tumors. Across 17 studies, the mean AUC for predicting MSI status from pathology images was ~ 0.89 in training and ~ 0.82 on

independent validation. While slightly lower than dedicated PCR or immunohistochemistry tests, this accuracy is quite encouraging as a potential upfront screening tool – an AI could flag cases likely to be MSI-high so that confirmatory tests can be done. In contrast, prediction of specific mutations like KRAS from H&E has been less successful (AUC often < 0.75), indicating that morphological surrogates for those mutations are not as pronounced [8]. The transformer architecture has also been explored in pathology; for example, vision transformers have been applied to colorectal histology classification tasks and shown performance on par with CNNs, though CNNs remain more common in practice.

3. AI in Radiology (Imaging): Radiologic evaluation is crucial in CRC for staging (assessing tumor extent and spread) and surveillance. AI applications in CRC radiology include detection of tumors on scans, segmentation of tumors and lymph nodes, and **radiomics**-based analysis for characterization. One area of interest is using AI to detect colorectal tumors on cross-sectional imaging *even when the scan was done for another purpose*. For instance, an incidental colorectal mass on a CT of the abdomen might be overlooked by a non-specialist; an AI could serve as a second reader. A recent study (published 2025) developed a deep learning model to identify CRC on routine contrast-enhanced CT scans. In an external validation set, the AI achieved about **81% sensitivity and 91% specificity** for detecting CRC. This was comparable to or slightly better than radiologist readers (whose sensitivities were 73% and 81% in the study). Notably, the AI system detected a few, suggesting AI could reduce oversight errors in CT interpretation. The AUC of the model was around 0.81, indicating good discrimination. Such AI tools could be integrated into radiology workflows to flag suspicious bowel segments on CT scans, prompting a confirmatory colonoscopy.

Another diagnostic radiology domain is **CT colonography (virtual colonoscopy)**. This is a screening modality where CT scans are used to find polyps. AI has been applied to both detect polyps in CT images and to characterize them. A proof-of-concept study by Grosu et al. used radiomics features and machine learning to differentiate benign from pre-malignant polyps on CT colonography. The algorithm had $\sim 82\%$ sensitivity and 85% specificity (AUC ~ 0.91) for identifying adenomatous polyps, effectively functioning as a noninvasive “virtual biopsy”. This suggests that AI could potentially triage polyps found on CT scans, helping radiologists decide which lesions likely require colonoscopic removal.

AI-driven **segmentation** of tumors and organs is another radiologic application relevant to both diagnosis and treatment. Deep learning models (often using U-Net architectures or its variants) can automatically outline tumors on MRI or CT. In rectal cancer, for example, CNN-based models have been developed to segment the primary tumor on MRI for volumetric assessment, and to

segment involved lymph nodes or mesorectal fascia for staging. The accuracy of these segmentations is approaching that of experts in many studies, and the automation can greatly speed up radiology and radiation planning workflows. Automated segmentation is particularly valuable in radiation oncology planning (defining target volumes) – studies have shown deep learning can delineate gross tumor volumes and organs-at-risk in rectal cancer patients' planning CTs with high Dice similarity scores, reducing the time burden on clinicians [10].

In summary, across endoscopic, pathologic, and radiologic diagnostics in CRC, AI systems – especially those based on CNN/deep learning – have achieved **high sensitivity and accuracy**, often matching or exceeding human performance in specific tasks. They excel at identifying patterns (polyps, tumor histology, imaging features) that might be subtle or tedious for humans to catch consistently. **Table 1** encapsulates some of these diagnostic performance highlights. The next sections address how AI extends into CRC management decisions, leveraging these diagnostic insights for prognostication and therapy guidance.

AI in CRC Management and Treatment Planning

While improving diagnosis is the first step, AI is also being leveraged to inform **treatment decisions and prognostication** in CRC. Management of CRC can be complex-e.g., deciding whether a rectal cancer patient can be managed without surgery after chemoradiotherapy, or which colon cancer patients are at high risk of recurrence and may benefit from more aggressive therapy or intensive follow-up. AI models can integrate imaging, pathology, and clinical data to assist in these judgments.

1. Treatment Decision Support and Outcome Prediction:

An exciting development is the use of AI to predict **response to therapy** before or during treatment. In locally advanced rectal cancer, patients often receive neoadjuvant chemoradiotherapy (nCRT) before surgery. About 15-30% achieve a pathologic complete response (pCR)-no viable tumor cells in the resected specimen – which raises the possibility of avoiding surgery (the “watch-and-wait” approach) for those who respond excellently. However, identifying these complete responders *before* surgery is challenging. AI models using MRI data have shown promise in predicting pCR. A systematic review in 2025 compiled 26 studies on MRI-based AI for rectal cancer response [9]. Most studies reported AUCs **above 0.80** for predicting pCR on independent validation, especially when using a combination of T2-weighted and diffusion-weighted MRI sequences as inputs. Larger training datasets and incorporation of advanced MRI features improved performance. These AI models (often CNNs or ensemble models) can analyze post-nCRT MRI scans to detect subtle textural or morphological changes indicative of complete tumor eradication. If validated prospectively, such tools could help identify which

patients can be safely offered non-operative management, thus sparing them the morbidity of surgery [9]. That said, inconsistency in model design and MRI protocols across studies remains an issue, and integration of clinical features (like CEA levels or tumor DNA markers) could further enhance predictive power.

AI has also been used to predict **patient prognosis**-e.g., risk of recurrence or survival-which can guide management intensity. A common approach is **radiomics**, where a large number of quantitative features are extracted from imaging (CT or MRI) and then fed into a machine learning model to find patterns correlating with outcomes. In CRC with liver metastases, radiomics models can analyze the texture and shape of liver lesions on pre-treatment CT to predict outcomes after surgery. One study found that a radiomics-based machine learning model (using random survival forests and a deep survival network) was able to stratify patients by recurrence risk better than the traditional clinical risk score (concordance index 0.70 vs 0.57). In fact, the radiomics model significantly **outperformed the clinical risk score** in predicting 18-month recurrence. Such a model could be used to identify patients who might benefit from additional therapies (e.g., more cycles of chemotherapy or novel agents) due to high risk of early relapse, or conversely, to potentially spare low-risk patients from overly aggressive treatment. Another area is using AI on pathology slides to predict prognosis – some studies have shown that deep learning features extracted from H&E slides (like those capturing immune cell density or tumor architecture) can independently predict survival, sometimes more robustly than human-assessed features or even some molecular markers.

2. Personalized Treatment Planning: AI can assist in tailoring treatment strategies beyond just prediction. For example, in oncology decision-support, researchers have explored AI systems that take patient-specific inputs (tumor genomics, patient comorbidities, etc.) and recommend treatment options based on learned outcomes from prior patients. In CRC, there are early efforts using reinforcement learning or decision tree algorithms to recommend chemotherapy regimens or to identify optimal surgical vs. non-surgical approaches for rectal cancer, although this is still nascent. More immediately, AI is being integrated into **surgical planning and intraoperative guidance**. One novel example is the use of augmented reality with AI during colon surgery to identify tumor margins or lymph nodes by analyzing real-time video – akin to how AI in colonoscopy works, but in the surgical field. Similarly, for liver metastasis resection, AI algorithms are being used to process preoperative imaging and assist in mapping optimal resection planes that spare healthy liver while removing all tumor nodules (some planning software now includes AI-driven suggestions).

In radiation oncology, AI-based **auto-segmentation** of target volumes (mentioned earlier) directly impacts treatment planning

efficiency. Deep learning models that automatically delineate the post-operative bed or lymph node regions in rectal cancer can expedite the creation of radiation plans, ensuring consistent coverage of areas at risk. Studies report substantial time savings and more consistent contours when using AI assistance, with only minor manual adjustments needed by clinicians [10]. Additionally, AI is being studied for optimizing radiation dose distributions – e.g., using neural networks to generate radiotherapy plans that meet dose constraints (a process known as inverse planning, which AI can potentially accelerate).

3. Integrating Multimodal Data: A clear trend in recent research (and a noted future direction) is combining data from multiple sources—endoscopic findings, pathology, radiology, blood biomarkers, genomics—into **comprehensive AI models** for CRC management [6]. With transformers and advanced neural networks that can handle heterogeneous data, one can envision an AI system that takes as input a patient's colonoscopy images (to gauge what was seen and removed), digital pathology from the resected tumor (to analyze tumor biology), and radiological scans (for staging), along with clinical variables. Such a system could output a recurrence risk or suggest an optimal adjuvant therapy. Early attempts at multimodal AI in oncology show improved performance over single-source models, as different data modalities provide complementary information. For CRC specifically, researchers predict that **multimodal AI** will be a key area, enabling truly personalized medicine (e.g., predicting which patients will benefit from immunotherapy by combining histology AI features with genomic data) [6].

In summary, AI's role in CRC management is evolving from risk prediction tools to direct treatment decision aids. Current evidence, summarized in **Table 2**, indicates that AI can: (a) accurately predict important biomarkers like MSI from routine data (saving time and cost), (b) identify patients likely to have excellent or poor treatment responses (informing tailored strategies), and (c) improve the efficiency and consistency of treatment planning processes (in surgery and radiotherapy). These contributions, while promising, require robust validation to ensure they indeed translate into better patient outcomes (e.g. improved survival, better quality of life via treatment de-escalation, etc.) when implemented.

Discussion

This systematic review highlights that artificial intelligence has made significant inroads in both the diagnostic and therapeutic realms of colorectal cancer care. In diagnostics, AI systems—particularly those based on deep learning CNNs—have demonstrated performance at or above human expert levels in multiple tasks: real-time polyp detection in colonoscopy, histopathologic identification of cancer in biopsy slides, and detection of tumors on imaging. The integration of AI as an assistive tool in colonoscopy has already

proven its value by consistently increasing adenoma detection rates in randomized trials, which is expected to ultimately reduce colorectal cancer incidence and mortality in screened populations [1]. Likewise, AI in pathology can reduce diagnostic delays and errors, ensuring that even subtle malignant features are recognized. In radiology, AI can act as a safety net and efficiency booster, flagging suspicious lesions and potentially allowing earlier cancer diagnosis, sometimes even from scans done for other reasons.

In management and decision-making, AI models are moving the field toward personalized medicine. By stratifying patients based on predicted risk or response, AI can inform the intensity of treatment: for example, identifying a patient likely cured by chemoradiation alone versus one who absolutely needs surgery and additional therapy. These uses align with a precision oncology approach, where treatment is tailored not just to general stage categories but to the individual tumor and patient characteristics as discerned by data. Importantly, AI's ability to assimilate large-scale data and subtle patterns can uncover prognostic insights that humans might miss. For instance, a radiomics signature combining 50 imaging features might reveal a prognostic pattern of tumor heterogeneity that no single radiologist could visually discern. Thus, AI can unveil “hidden” biomarkers.

Despite these clear strengths and the excitement around AI, our review also underscores limitations and challenges that temper the immediate implementation of many AI tools in routine CRC care:

- **Generalizability and Validation:** Many AI models are trained and tested on specific datasets that may not reflect the full diversity of clinical practice. An AI trained on high-quality colonoscopy videos from expert centers may perform less well in community settings or with different equipment. Similarly, pathology AI often requires high-quality slide scans and may be thrown off by variations in staining or scanning. External validation on diverse cohorts is essential, and a number of studies noted performance drop-offs when evaluated on independent data (e.g., MSI prediction AUC dropping to ~0.82 on external sets) [8]. This calls for robust, multi-center training datasets and possibly techniques like domain adaptation to ensure AI models maintain accuracy across different hospitals and patient populations.

- **Integration into Workflow:** Introducing AI into the clinical workflow requires thoughtful integration with existing tools and routines. For example, AI polyp detection during colonoscopy should seamlessly display alerts on the endoscopy monitor without distracting or overwhelming the endoscopist. User-interface design is critical so that clinicians trust and effectively use AI recommendations. In pathology, an AI system needs to plug into digital slide viewers and perhaps triage slides (e.g., highlight regions with potential tumor for the pathologist). Radiologists might get AI “second read” results integrated into their

PACS (Picture Archiving and Communication System). Achieving this integration often requires collaboration with software vendors and regulatory approval for medical device software.

- **False Positives/Negatives and Trust:** No AI is perfect, and both false positives and false negatives have implications. In colonoscopy, an AI false positive might cause unnecessary polyp removal or prolonged procedure time, whereas a false negative is a missed lesion. Encouragingly, studies show false positives are relatively infrequent and usually obvious artifacts, but ensuring that endoscopists can distinguish AI alerts that are likely trivial is important (perhaps through AI confidence scoring) [1]. A bigger issue is clinician *trust* in AI: if the AI says, “no polyp here” or “this polyp is hyperplastic” and the clinician disagrees, how is that resolved? At present, AI is generally used as an adjunct, and human experts remain the final decision-makers. Building trust will come from understanding AI’s failure modes and having transparency. This is driving research into explainable AI (XAI) – e.g., heatmaps on pathology slides showing what features led to an AI’s cancer prediction, or attention maps on colonoscopy video indicating what triggered an alert.

- **Interpretability and Black Box Nature:** Many deep learning models, especially CNNs and transformers, are often criticized as “black boxes” – they yield a prediction without a clear rationale. In medicine, knowing the reasoning is valuable for acceptance and for medicolegal reasons. Efforts to improve interpretability include using algorithms that highlight image features (as mentioned) or developing simplified decision rules from complex models. Some AI systems in CRC histopathology, for example, have been coupled with analyses of which histologic structures (glands, mucin, lymphocytes) were most predictive of an outcome, making them somewhat more interpretable to pathologists.

- **Data Privacy and Governance:** Training powerful AI models requires a lot of data – potentially from thousands of patients. Aggregating such data raises issues of privacy and data sharing. While techniques like federated learning (where models are trained across multiple institutions without sharing raw data) are emerging, implementing them is non-trivial. There are also intellectual property issues regarding AI models developed on institutional data.

- **Regulatory and Ethical Considerations:** Any AI that will influence patient care is, in essence, a medical device or diagnostic test and falls under regulatory oversight (e.g., FDA approval in the US, CE marking in Europe). The process of approving AI, especially adaptive or continuously learning algorithms, is still developing. Moreover, medicolegal liability is a question: if an AI misses a cancer, is the liability on the physician or the tool’s manufacturer? Such questions have yet to be clearly

answered and may impact how eagerly clinicians embrace AI. Ethically, there is also a need to ensure AI does not perpetuate biases – if training data lacked diversity, AI might underperform in minority populations, which could exacerbate healthcare disparities if not corrected.

- **Workflow Impact and Cost:** While AI promises efficiency, there is an up-front cost in acquiring and implementing these technologies. There may be a need for new hardware (e.g., processors for real-time image analysis) or software subscriptions. Institutions will weigh these costs against the potential benefits (e.g., improved polyp detection leading to fewer cancers – a long-term benefit that might not directly reimburse the endoscopist). Demonstrating a clear *cost-effectiveness* will be important for wide adoption. Early health economic analyses in colonoscopy suggest that if AI meaningfully increases ADR, it could be cost-effective by preventing expensive cancer treatments down the line, but more data is needed [1].

Despite these challenges, the momentum of research and development in AI for CRC is strong. In the last five years, there has been an exponential increase in publications at the intersection of AI and colon cancer, and the technology is steadily moving from research to practice [6]. Gastroenterology societies are beginning to issue guidance on AI; for example, the American Gastroenterological Association (AGA) issued a Clinical Practice Update suggesting that AI can be used to enhance adenoma detection during colonoscopy, while cautioning that endoscopists need to be aware of its limitations. In pathology, large pathology labs are piloting AI for primary diagnosis and getting CLIA certifications for some AI-assisted workflows. In radiology, AI tools are being integrated into some CT colonography reading software and MRI analysis platforms.

Strengths and Limitations of the Review: This review comprehensively covered both diagnostic and management applications of AI in CRC, which is a broad scope. By including various study designs (RCTs, diagnostic accuracy studies, retrospective analyses), we painted a holistic picture of the field. However, one limitation is that due to the breadth, we could not delve exhaustively into each sub-topic (each of which could be a review in itself, e.g., AI in colonoscopy). There is also an inherent bias towards positive findings in published literature; negative or inconclusive studies on AI might be under-reported, skewing the impression of AI’s efficacy. We attempted to mitigate this by noting challenges and instances where AI did not outperform humans (such as CAD EYE’s slightly lower sensitivity than experts in one study for polyp characterization). Additionally, the rapid pace of AI research means new results may emerge soon after this review; we focused on high-quality, peer-reviewed sources up to early 2025.

Future Directions

Looking ahead, several developments are anticipated in this domain:

- **Prospective Clinical Trials with Outcomes:** While many studies show AI improves immediate diagnostic metrics (like ADR or accuracy), the true test will be whether using AI leads to better patient outcomes – e.g., lower cancer incidence, improved survival, better quality of life. Future trials might randomize centers to use AI or not and see differences in patient outcomes over years. Such data will solidify AI's role in guidelines.
- **Continuous Learning and Adaptation:** AI models could be designed to continuously learn from new data in a clinical setting (with proper oversight). For example, a colonoscopy AI system might improve its detection over time as it processes more videos. Ensuring this is done without compromising safety will be an area of focus.
- **Multimodal AI and Decision Support Systems:** As noted, combining data types is a promising frontier. A multimodal AI “virtual tumor board” that assesses a patient's case from all angles (imaging, pathology, genomics) and provides recommendations could become a reality. This might involve advanced architectures (e.g., combining CNNs for images with transformers for genomic/proteomic data).
- **Transformer Models and NLP:** Transformers are not only for images but also for text/NLP. AI might be used to read and synthesize information from clinical notes, pathology reports, and literature to assist oncologists (for instance, matching patient tumor molecular profiles with relevant clinical trials – though beyond the strict scope of CRC diagnosis/management, it's a related AI application in oncology).
- **Addressing Rare Scenarios:** AI development so far has tackled common tasks; future work might target rarer but important challenges, such as identifying genetic syndromes (Lynch syndrome) via patterns of polyps and tumors, or detecting neuroendocrine tumors of the colon, etc., where data are more limited.

In conclusion, the integration of AI into colorectal cancer care is progressing rapidly. The evidence to date strongly supports AI's value in improving diagnostic detection and providing decision support, but careful implementation and further validation are needed to fully realize its benefits. Rather than replacing clinicians, AI will serve as a powerful adjunct, handling data-rich and repetitive tasks, thus freeing healthcare professionals to focus on complex decision-making and patient care.

Conclusion

Artificial intelligence is revolutionizing how clinicians approach colorectal cancer, from early detection to post-treatment surveillance. In diagnostic arenas – colonoscopy, pathology, and radiology – AI systems (particularly deep learning models) have shown high accuracy, improving upon human performance in polyp and tumor detection and classification. These technologies can increase the adenoma detection rate in colonoscopy, standardize pathology assessments, and enhance radiologic identification of disease, all of which contribute to more accurate and timely diagnoses. In the realm of CRC management, AI tools are emerging that can predict treatment responses (like which rectal cancers might be cured without surgery) and patient outcomes (risk of recurrence or benefit from certain therapies), paving the way for more personalized treatment planning. The ability of AI models to analyze vast and complex datasets offers clinicians a new level of decision support, potentially improving prognostication and guiding therapy choices based on individualized risk profiles. However, successful translation of AI from research to routine practice will depend on overcoming challenges related to generalizability, trust, and integration. Ensuring that AI algorithms are trained on representative data and rigorously validated will safeguard against performance lapses in real-world settings. Moreover, clinicians and AI systems must develop a synergistic relationship – with AI providing useful insights and flags, and clinicians applying oversight, clinical context, and compassion in decision-making. The incorporation of AI into CRC workflows should be done in a manner that enhances efficiency and accuracy without adding undue burden or causing alert fatigue. The current evidence paints AI not as a distant future concept but as a present and expanding reality in colorectal cancer care. With several AI-enabled devices already approved for clinical use (e.g., in colonoscopy) and many pathology and radiology AI applications on the horizon, it is incumbent on the medical community to stay informed about these tools. Continued research, including prospective trials and health outcomes studies, will clarify the extent of AI's impact on patient survival and healthcare economics. If the optimistic results observed thus far are borne out, AI-assisted CRC care could lead to earlier cancer detection, more tailored treatments, and ultimately, improved patient outcomes. The integration of human expertise with artificial intelligence stands to transform CRC diagnosis and management for the better – fulfilling the promise of precision medicine in one of the world's most common and deadly cancers.

References

1. Spadaccini M, Menini M, Massimi D, Rizkala T, De Sire R, et al. (2025) AI and Polyp Detection During Colonoscopy. *Cancers (Basel)* 17: 797.

2. Wang KS, Yu G, Xu C, Meng XH, Zhou J, et al. (2021) Accurate diagnosis of colorectal cancer based on histopathology images using artificial intelligence. *BMC Med* 19: 76.
3. Guitton T, Allaume P, Rabilloud N, Rioux-Leclercq N, Henno S, et al. (2023) Artificial Intelligence in Predicting Microsatellite Instability and KRAS, BRAF Mutations from Whole-Slide Images in Colorectal Cancer: A Systematic Review. *Diagnostics (Basel)* 14: 99.
4. Montagnon E, Cerny M, Hamilton V, Derennes T, Ilinca A, et al. (2024) Radiomics analysis of baseline computed tomography to predict oncological outcomes in patients treated for resectable colorectal cancer liver metastasis. *PLoS One* 19: e0307815.
5. De Lange G, Prouvost V, Rahmi G, Vanbiervliet G, Le Berre C, et al. (2024) Artificial intelligence for characterization of colorectal polyps: Prospective multicenter study. *Endosc Int Open* 12: E413-E418.
6. Sun L, Zhang R, Gu Y, Huang L, Jin C (2024) Application of Artificial Intelligence in the diagnosis and treatment of colorectal cancer: a bibliometric analysis, 2004-2023. *Front Oncol* 14:1424044.
7. Soleymanjahi S, Huebner J, Elmansy L, Rajashekar N, Lüdtke N, et al. (2024) Artificial Intelligence-Assisted Colonoscopy for Polyp Detection. *Ann Intern Med* 177: 1652-1663.
8. Jong BK, Yu ZH, Hsu YJ, Chiang SF, You JF, et al. (2025) Deep learning algorithms for predicting pathological complete response in MRI of rectal cancer patients undergoing neoadjuvant chemoradiotherapy: a systematic review. *Int J Colorectal Dis* 40: 19.
9. Naseri S, Shukla S, Hiwale KM, M Jagtap M, Gadkari P, et al. (2024) From Pixels to Prognosis: A Narrative Review on Artificial Intelligence's Pioneering Role in Colorectal Carcinoma Histopathology. *Cureus* 16: e59171.
10. Bibault JE, Giraud P (2024) Deep learning for automated segmentation in radiotherapy: a narrative review. *Br J Radiol* 97: 13-20.