



대장암 위험 예측 모델에서의 인공지능 활용

서정국

중앙대학교 의과대학 소화기내과

The Use of Artificial Intelligence in Colorectal Cancer Risk Prediction Models: A Review

Jeongkuk Seo

Department of Internal Medicine, Chung-Ang University College of Medicine, Seoul, Korea

Received November 5, 2025
Revised November 20, 2025
Accepted November 26, 2025

Corresponding author:

Jeongkuk Seo

E-mail: tonari@gmail.com

<https://orcid.org/0000-0003-2915-2203>

The application of artificial intelligence (AI) and machine learning (ML) in colorectal cancer (CRC) risk prediction models has been increasingly used in clinical and research settings. AI-based models have been used to predict the risk of CRC and high-risk polyps by integrating diverse data, including clinical information, genetic factors, lifestyle factors, and imaging data. Various ML algorithms, such as random forest, artificial/deep neural networks, and support vector machines, are currently applied and often demonstrate a higher sensitivity and specificity than traditional statistical prediction models. In particular, AI has been shown to improve prediction accuracy by combining multimodal data sources, such as electronic health records, genomic information, and pathology images, thereby supporting personalized screening strategies. Furthermore, this study aimed to investigate the latest trends, performance, external validation results, integration into clinical workflow settings, and the future challenges of AI models for CRC risk prediction.

Key Words: Colorectal cancer; Artificial intelligence; Machine learning; Risk prediction; Screening

INTRODUCTION

대장암은 전 세계적으로 높은 발생률과 사망률을 보이는 주요 공중 보건 문제 중 하나이다[1]. 대장암은 조기 발견 시 5년 생존율이 90% 이상으로 높기 때문에, 효과적인 선별검사(screening) 전략이 매우 중요하다. 최근 인공지능(artificial intelligence, AI)과 머신러닝(machine learning, ML) 기술이 의료 분야에 도입되면서, 대장암 위험 예측 및 선별검사 영역에서도 혁신적인 변화가 일어나고 있다[2,3].

전통적인 대장암 위험 평가는 연령, 성별, 가족력, 생활습관 등 임상적 위험 요인에 의존해왔으나, 이러한 방식은 개인의 복합적인 위험을 정확히 반영하는 데 한계

가 있었다. AI 기반 모델은 임상 정보, 유전적 요인, 생활습관, 영상 데이터, 전자의무기록(electronic health record, EHR) 등 방대하고 이질적인 데이터를 통합하여 대장암 및 고위험 용종의 발생 위험을 보다 정밀하게 예측하는 데 사용된다[4,5].

AI의 적용은 대장암 선별검사(예: 대장내시경, 분변잠혈검사[fecal immunochemical test, FIT])의 효율성 향상, 고위험군 선별, 불필요한 검사 감소, 의료 자원 최적화에 기여할 수 있다[6,7]. 본 종설의 목적은 대장암 위험 예측 분야에서 AI 모델의 최신 개발 동향, 성능 검증 결과, 실제 임상 적용 사례를 검토하고, 향후 임상 워크플로우 통합을 위한 과제와 전망을 논의하는 것이다.



MAIN SUBJECTS

AI 기반 예측 모델의 개발 및 성능

AI 기반 대장암 위험 예측 모델은 다양한 데이터 소스와 ML 알고리즘을 기반으로 개발되고 있다. 대표적으로 random forest, 인공/심층 신경망(artificial/deep neural network, ANN/DNN), 서포트 벡터 머신(support vector machine, SVM)을 비롯해 여러 기본 모델을 결합하여 성능을 높이는 앙상블 학습(ensemble learning, EML) 등이 널리 적용되고 있으며, 이들 모델은 기존의 통계적 예측 모델(예: 로지스틱 회귀)보다 높은 민감도와 특이도를 보이는 경우가 많다[2,4,7].

특히 AI 모델은 EHR과 유전체 정보, 병리 이미지 등 다중 데이터 소스(multi-modal data)를 결합하여 위험 예측의 정확도를 높이는 데 강점이 있다. 예를 들어, AI 기반 모델은 35-50세의 젊은 성인에서 조기 대장암 위험을 예측하여 선별검사 대상자를 효율적으로 선정하는 데 활용될 수 있다[8]. 또한, 병리 슬라이드와 임상 기록을 통합한 딥러닝 모델은 5년 내 진행 위험 예측에 있어 기존 임상 변수만을 사용한 모델보다 더 우수한 성능을 보였다[9].

외부 검증 및 실제 임상 적용

개발된 AI 모델이 실제 임상 현장에서 유용성을 갖기 위해서는 다양한 인구 집단과 환경에서의 외부 검증(external validation)이 필수적이다. 최근 연구들은 AI 모델이

기존 통계적 방법보다 대장암 위험 예측에서 우수한 성능을 보임을 반복적으로 확인했다.

특히, 임상 데이터와 FIT 결과를 결합한 AI 모델은 외부 코호트에서 area under the curve (AUC) 0.93-0.98 수준의 높은 진단 정확도를 보였으며, 기존 위험 변수만을 사용한 모델 대비 통계적으로 유의하게 우수한 결과를 나타냈다[10]. 또한, 병리 이미지와 임상 기록을 통합한 딥러닝 기반 모델은 5년 내 진행 위험 예측에서 단일 임상 변수 모델보다 향상된 성능(AUC 0.672 vs. 0.666, $p = 0.002$)을 보였다[9].

실제 임상 적용 사례에서 AI를 활용한 예측 모델은 대장암 선별검사 대상자 선정, 고위험군 분류, 불필요한 내시경 검사 감소 등에서 의료 자원 효율화에 기여하고 있다. 미국의 일부 헬스케어 시스템에서는 독자적인 AI 기반 대장암 위험 예측 및 진단 보조 시스템을 개발하여 임상 현장에 도입하고 있으며, 실제 환자 데이터와 장기 추적 결과를 활용한 모델의 임상적 유용성이 점차 입증되고 있다[2,5,11]. 본 종설에서 검토된 주요 연구들의 알고리즘, 데이터 특성 및 성능은 Table 1에 요약하였다.

임상 워크플로우 통합 및 비용-효과성

AI 기반 예측 모델의 임상 워크플로우 통합은 불필요한 대장내시경 감소, 조기 진단율 증가, 비용 절감, 장기 추적 결과 개선 등에서 긍정적인 영향을 미칠 잠재력을 가지고 있다.

AI를 활용한 위험 예측 및 선별 전략은 임상 정보, 생체

Table 1. Summary of key studies on artificial intelligence-based colorectal cancer risk prediction

Study (reference)	Key algorithm	Input data	Dataset size	Key performance
Malik et al. [4] Abuzuhri et al. [5]	RF, SVM, neural networks, EML	Various clinical, genomic, lifestyle data (systematic review)	Meta-analysis of > 30 studies	AUROC 0.83-0.99 (varies by model/data; superior performance with EML and molecular data)
Hussan et al. [8]	Machine learning	EHR	Single-center large cohort	Development of early-onset CRC risk prediction model (see paper for metrics)
Jiang et al. [9]	Deep learning (multimodal)	Pathology images + clinical records	Multi-center data	Prediction of progression within 5 years AUC 0.672 (vs. clinical variables only 0.666)
Eskelinen et al. [10]	ANN, RF, gradient boosting	Clinical data + FIT results	Includes external validation cohort	Improved diagnostic accuracy vs. FIT alone AUC 0.93-0.98

RF, random forest; SVM, support vector machine; EML, ensemble learning; AUROC, area under the receiver operating characteristic curve; EHR, electronic health record; CRC, colorectal cancer; AUC, area under the curve; ANN, artificial neural network; FIT, fecal immunochemical test.

지표, 생활습관, 유전체 등 다양한 데이터를 통합하여 고 위험군을 선별하여 대장내시경을 권유함으로써 의료 자원 효율성을 높인다[6]. 실제로 AI 기반 모델은 기존 FIT 단독 전략 대비, FIT 양성 후 AI 보조 대장내시경을 시행할 경우 불필요한 내시경 감소와 더불어 조기 대장암 진단율이 증가하며, 예방된 대장암 사례 수와 생존 연수 증가, 비용 대비 효과(incremental cost-effectiveness ratio) 개선이 확인되었다.

특히, 아시아 인구 집단을 대상으로 한 마르코프 모델 분석에서 FIT + AI 대장내시경 전략이 가장 비용 효과적이었으며, 기존 전략 대비 추가 비용 없이 더 많은 생존 연수와 암 예방 효과를 보였다[12]. 이처럼 AI 기반 예측 모델은 조기 진단율을 높여 치료 성적과 생존율을 개선하는데 기여하며, 장기 추적 연구에서는 대장암 발생률 및 사망률 감소에 긍정적 영향을 미칠 것으로 기대된다.

다기관 연구, 재현성 및 향후 과제

AI 기반 대장암 위험 예측 모델의 임상 적용이 확대됨에 따라, 다기관 연구를 통한 외부 검증과 재현성(reproducibility) 확보가 임상 워크플로우 통합을 위한 핵심 과제로 인식되고 있다.

최근 메타분석과 체계적 문헌고찰에서는 random forest, ANN/DNN, SVM 등 다양한 모델이 여러 코호트에서 높은 진단 성능(area under the receiver operating characteristic curve 0.83–0.88, 민감도 0.83, 특이도 0.80)을 보였다[4]. 그러나 데이터셋의 이질성과 검증 방법의 다양성으로 인해 성능의 일관성 확보가 필요하다는 점이 강조된 바 있다[11].

특히, 미국 New Hampshire Colonoscopy Registry 등 실제 임상 데이터와 장기 추적 결과를 활용한 다기관 연구에서는 임상 기록과 병리 이미지 등 다중 데이터 소스를 통합한 AI 모델이 5년 내 진행 위험 예측에서 기존 임상 변수 단독 모델보다 통계적으로 유의하게 우수한 성능을 보였다는 점에서, 이러한 다중 모달리티(multi-modality) 접근은 다양한 인종, 성별, 생활습관을 포함한 인구 집단에서의 재현성을 높이는 데 기여할 것으로 생각된다[9].

그럼에도 불구하고, AI 모델이 실제 임상 현장에서 고위험군 선별 및 의료 자원 최적화에 기여하기 위해서는 표준화된 외부 검증 절차뿐만 아니라, 모델의 의사결정 과

정을 투명하게 보여주는 설명 가능한 AI (explainable AI, XAI) 기술의 도입이 필수적이다.

특히, 딥러닝이나 EML과 같이 고성능을 보이는 모델들은 내부 연산 과정을 알기 어려운 '블랙박스(black box)' 특성을 가진다. 이는 의료진이 AI의 예측 결과를 맹목적으로 신뢰하거나 반대로 불신하게 만드는 원인이 될 수 있다. 이를 극복하기 위해 최근에는 Shapley Additive exPlanations이나 Local Interpretable Model-agnostic Explanations과 같은 XAI 방법론이 주목받고 있다. 이러한 기술은 각 환자의 데이터에서 어떤 변수(예: 특정 유전 인자, 연령, 과거 병력 등)가 대장암 위험 예측에 긍정적 혹은 부정적 영향을 미쳤는지 시각화하여 제공한다. 결과적으로 의료진은 AI 모델의 예측이 생물학적, 임상적 지식과 일치하는지(biological plausibility)를 검증할 수 있게 되며, 이는 AI 모델을 실제 임상적 의사결정 지원 도구(clinical decision support system)로 통합하는 데 있어 의료진의 신뢰를 확보하는 데 중요한 역할을 할 것이다[11,13,14].

CONCLUSION

AI 기반 대장암 위험 예측 모델은 임상, 유전체, 병리 이미지 등 다중 데이터를 통합하여 기존의 통계적 모델보다 우수한 예측 성능을 보인다. 이러한 모델들은 외부 검증 연구를 통해 높은 진단 정확도를 입증하고 있으며, 근시일 내 실제 임상 워크플로우에 통합되어 불필요한 대장내시경을 줄이고 조기 진단율을 높이며, 비용-효과성을 개선하는 등 긍정적인 영향을 미칠 것으로 생각된다.

그러나 데이터셋의 다양성 부족, 모델의 외부 검증 및 재현성 확보 문제, 표준화된 성능 평가의 부재, 그리고 임상 현장에서의 수용성을 높이기 위한 XAI의 도입은 여전히 해결해야 할 주요 과제로 남아있다.

결론적으로, AI 기반 대장암 위험 예측 모델은 향후 개인 맞춤형 선별검사 및 예방 전략의 핵심 도구로 자리잡을 것으로 전망된다. 이를 위해 지속적인 다기관 협력 연구, 표준화된 검증 프로토콜 개발, 그리고 임상 현장과의 긴밀한 통합 노력이 필요하다.

FUNDING

None.

CONFLICTS OF INTEREST

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

REFERENCES

1. Eng C, Yoshino T, Ruíz-García E, et al. Colorectal cancer. *Lancet* 2024;404:294-310. [https://doi.org/10.1016/S0140-6736\(24\)00360-X](https://doi.org/10.1016/S0140-6736(24)00360-X)
2. Farhoudian A, Heidari A, Shahhosseini R. A new era in colorectal cancer: artificial Intelligence at the forefront. *Comput Biol Med* 2025;196(Pt C):110926. <https://doi.org/10.1016/j.compbimed.2025.110926>
3. Hassan C, Bisschops R, Sharma P, Mori Y. Colon cancer screening, surveillance, and treatment: novel artificial intelligence driving strategies in the management of colon lesions. *Gastroenterology* 2025;169:444-455. <https://doi.org/10.1053/j.gastro.2025.02.021>
4. Malik S, Naqvi A, Tenorio BG, Farrukh F, Tariq R, Adler DG. Machine learning for predicting colorectal cancer and advanced colorectal polyps: a systematic review and meta-analysis. *J Clin Gastroenterol* 2025. doi: 10.1097/MCG.0000000000002172. [Epub ahead of print]
5. Abuzuhri MA, Najjar S, Awad M, et al. Machine learning in colorectal cancer prediction and diagnosis: a systematic review of models' performance. *Int J Med Inform* 2026;206:106170. <https://doi.org/10.1016/j.ijmedinf.2025.106170>
6. Tokutake K, Morelos-Gomez A, Hoshi KI, Katouda M, Tejima S, Endo M. Artificial intelligence for the prevention and prediction of colorectal neoplasms. *J Transl Med* 2023;21:431. <https://doi.org/10.1186/s12967-023-04258-5>
7. Sharma A, Kumar R, Yadav G, Garg P. Artificial intelligence in intestinal polyp and colorectal cancer prediction. *Cancer Lett* 2023;565:216238. <https://doi.org/10.1016/j.canlet.2023.216238>
8. Hussan H, Zhao J, Badu-Tawiah AK, et al. Utility of machine learning in developing a predictive model for early-age-onset colorectal neoplasia using electronic health records. *PLoS One* 2022;17:e0265209. <https://doi.org/10.1371/journal.pone.0265209>
9. Jiang S, Robinson C, Anderson J, et al. Improving colorectal cancer screening and risk assessment through predictive modeling on medical images and records. *Am J Pathol* 2025. doi: 10.1016/j.ajpath.2025.09.016. [Epub ahead of print]
10. Eskelinen M, Selander T, Guimarães DP, Kaarniranta K, Syrjänen K, Eskelinen M. Four different artificial intelligence models versus logistic regression to enhance the diagnostic accuracy of fecal immunochemical test in the detection of colorectal carcinoma in a screening setting. *Anticancer Res* 2025;45:2477-2491. <https://doi.org/10.21873/anticancer.17619>
11. Merabet A, Saighi A, Saad H, et al. AI for colon cancer: a focus on classification, detection, and predictive modeling. *Int J Med Inform* 2026;206:106115. <https://doi.org/10.1016/j.ijmedinf.2025.106115>
12. Wong MC, Huang J, Lam TY, Lau LH, Chiu PW. The cost-effectiveness of AI-assisted colonoscopy as a primary or secondary screening test in a population-based colorectal cancer screening program: markov modeling-based cost effectiveness analysis. *J Med Internet Res* 2025;27:e67762. <https://doi.org/10.2196/67762>
13. Reitsam NG, Enke JS, Vu Trung K, Märkl B, Kather JN. Artificial intelligence in colorectal cancer: from patient screening over tailoring treatment decisions to identification of novel biomarkers. *Digestion* 2024;105:331-344. <https://doi.org/10.1159/000539678>
14. Salih AM, Raisi-Estabragh Z, Galazzo IB, et al. A perspective on explainable artificial intelligence methods: SHAP and LIME. *Adv Intell Syst* 2025;7:2400304. <https://doi.org/10.1002/aisy.202400304>