SYSTEMATIC REVIEW

Open Access



Narges Norouzkhani^{1†}, Hesam Mobaraki^{2†}, Shirin Varmazyar^{3†}, Hadis Zaboli⁴, Zhina Mohamadi⁵, Golnaz Nikeghbali⁶, Kamyar Bagheri⁷, Newsha Marivany⁸, Mirmehdi Najafi⁹, Mahdiyeh Nozad Varjovi¹⁰, Mohamed Abouzeid^{11*}, Hanieh Zeidi Baghrabad¹², Pooya Eini¹³, Aida Azhdarimoghaddam¹⁴, Farbod Khosravi⁶ and Mahsa Asadi Anar^{15*}

Abstract

Background and aim Artificial intelligence (AI) networks offer significant potential for predicting immunotherapy outcomes in gastrointestinal cancers by analyzing genetic mutation profiles. Their application in prognosis remains underexplored.

This systematic review and meta-analysis aim to evaluate the effectiveness of AI-based models, which refers to systems utilizing artificial intelligence to analyze data and make predictions, in predicting immunotherapy responses in gastrointestinal cancers using genetic mutation features.

Methods This study, adhering to PRISMA guidelines, aimed to evaluate AI networks for predicting gastrointestinal cancer prognosis in response to immunotherapy using genetic mutation features. A search in PubMed, WOS, and Scopus identified relevant studies. Data extraction and quality assessment were conducted, and statistical analysis included pooled estimates for sensitivity, specificity, accuracy, and AUC. Regression models and imputation methods addressed missing values, ensuring accurate and robust results. STATA version 18 was used to analyze the data.

Result A total of 45 studies, all published in 2024, involving 14,047 participants in training sets and 10,885 participants in test sets, were included. The pooled results of AI model performance for gastrointestinal cancers based on genetic mutation features were: AUC = 0.86 (95% CI: 0.86–0.87), Sensitivity = 83% (95% CI: 83%-84%), Specificity = 72% (95% CI: 72%-73%), and Accuracy = 82% (95% CI: 82%-83%). Heterogeneity was low to moderate, and no publication bias was detected. Subgroup analysis showed higher AUC for gastric cancer models (AUC: 0.87) and lower for pancreatic cancer models (AUC: 0.52).

[†]Narges Norouzkhani, Hesam Mobaraki and Shirin Varmazyar contributed equally to this work.

*Correspondence: Mohamed Abouzeid muhamedabozeid@gmail.com Mahsa Asadi Anar Mahsa.boz@gmail.com Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Conclusion Al networks demonstrate promising potential in predicting immunotherapy outcomes for gastrointestinal cancers based on genetic mutation features. This systematic review highlights their effectiveness in stratifying patients and optimizing treatment decisions. However, further large-scale studies are needed to validate Al models and integrate them into clinical practice for improved precision in cancer immunotherapy.

Keywords ANN, Gastrointestinal cancer, Genetic mutation, Immunotherapy, Artificial Intelligence Networks

Introduction

Gastrointestinal cancers (GICs) are a major global health concern, accounting for 26% of all cancer diagnoses and contributing to 35% of cancer-related deaths [1], making them among the most common and deadly malignancies worldwide [1].

According to the World Cancer Observatory estimates in 2022, East Asia reported 1,469,225 new cases of gastrointestinal cancer, accounting for 43.1% of the global incidence. Notably, gastrointestinal cancers were responsible for 837,360 deaths in the region, constituting 41.7% of all cancer-related mortality [2]. Immunotherapy, a new therapeutic approach based on the principles of cancer immunoediting, which highlights the significant impact of immune evasion on tumor progression and development, has initiated a transformative change in the landscape of cancer treatment [3]. Immunotherapy is known as a therapeutic approach aimed at re-establishing the normal immune response against tumors, thereby reactivating the interaction between the immune system and tumors and ultimately facilitating the eradication of cancer cells [4]. This category includes a variety of methods, including immune checkpoint inhibitors (ICIs), cancer vaccines, cell therapy, and oncolytic viruses (OVs) [2]. ICIs have been recognized as a promising therapeutic approach for various types of cancer, especially gastrointestinal cancers. However, the efficacy of ICIs is limited, with response rates ranging from 10 to 20% depending on the specific tumor type. Therefore, the development of biomarkers that can reliably identify patients most likely to respond positively to ICI therapy is critical [5]. Currently, the most reliable prognostic biomarkers for evaluating the efficacy of ICIs are the degree of microsatellite instability (MSI) and programmed death ligand 1 (PD-L1) expression levels [6]. Tumors with high MSI have response rates over 50%, but they represent only 4% of GICs, prompting increased interest in PD-L1 expression. The choice of antibodies in immunohistochemistry affects the accuracy of tumor evaluations and therapy eligibility. PD-L1 expression has a significant negative predictive value; its absence correlates with a 2–6% response rate for ICI monotherapy. In contrast, a PD-L1 combined positive score (CPS) of ≥ 1 is associated with a 15–16% response rate, while a CPS of ≥ 10 correlates with a 24-25% response rate. Additionally, TMB shows strong potential as a biomarker for ICI responses [7]. TMB has not yet been validated as a reliable biomarker for GECs. The immunological impact of mutations varies, with specific mutations in proteins such as PBRM1, KEAP1, and STK11. potentially affecting the efficacy of ICI therapy in both beneficial and deleterious ways. Furthermore, the predictive ability of TMB scoring systems for ICIs appears to be limited, as they cannot account for the unique consequences of these mutations. To deal with this limitation, recent researches have proposed modifying the TMB calculation method or creating gene mutation-based signatures to improve the accuracy of predicting ICI treatment outcomes [5]. Machine learning (ML) and deep learning (DL) represent significant advancements in addressing intricate challenges within the medical field, particularly through the utilization of extensive clinical data sets [8]. These methodologies have proven their effectiveness and success in various predictive and clustering applications [9]. The implementation of these innovative technologies enables a comprehensive investigation into the mechanisms underlying therapy resistance across multiple dimensions, including transcriptional, epigenetic, and translational aspects, thereby providing valuable insights to enhance the effectiveness of ICIs [1]. The application of DL techniques in forecasting responses to immunotherapy has not yet reached its full potential, despite their increasing prevalence. This is particularly pertinent for patients with GICs who are receiving immune checkpoint blockade (ICB) therapy, where the demand for effective predictive models is intensifying. Central to these models are Artificial Neural Networks (ANNs), which provide the foundational framework for predictions. By leveraging next-generation sequencing data, our objective is to accurately anticipate individual responses to therapeutic interventions [10]. In this topic, conflicting reports have emerged regarding the performance of ANNs. The connector segment of stratification through ANNs can exert a significant influence. Given that the overall performance of ANNs has not yet been comprehensively reported, it is essential to address the inconsistencies and contradictions where some studies deem them effective while others consider them ineffective. Consequently, we conducted a systematic investigation to arrive at a conclusive understanding of this matter.

Table 1 Curated search strategies and result of the searching procedure

PubMed (((((((("Artificial intelligence network"[Title/Abstract]) OR ("Artificial intelligence networks"[Title/Abstract])) OR ("Artificial 171 intelligence"[Title/Abstract])) OR ("Machine Intelligence"[Title/Abstract])) OR ("AI"[Title/Abstract])) OR ("Machine learning"[Title/ Abstract])) OR ("ML"[Title/Abstract])) OR ("Deep learning"[Title/Abstract])) OR ("DL"[Title/Abstract])) OR ("convolutional neural cancer"[Title/Abstract]) OR ("Gastrointestinal cancers"[Title/Abstract])) OR ("Gastrointestinal neoplasm"[Title/Abstract])) OR ("Gastrointestinal neoplasm"[Title/Abstract])) OR ("Gastrointestinal neoplasm"] intestinal tumor"[Title/Abstract])) OR ("Gl cancer"[Title/Abstract])) OR ("Gl neoplasm"[Title/Abstract])) OR ("Gastrointestinal stromal tumor"[Title/Abstract])) OR ("GIST"[Title/Abstract])) OR ("Digestive system neoplasm"[Title/Abstract])) OR ("Stomach Cancer"[Title/ Abstract])) OR ("Intestine Cancer"[Title/Abstract])) OR ("esophagus Cancer"[Title/Abstract])) OR ("colon cancer"[Title/Abstract])) OR ("rectum Cancer"[Title/Abstract])) OR ("anus Cancer"[Title/Abstract])) OR ("colorectal cancer"[Title/Abstract])) OR ("pancreas cancer"[Title/Abstract])) OR ("liver cancer"[Title/Abstract])) OR ("ileum cancer"[Title/Abstract])) OR ("Jejunal adenocarcinoma"[Title/ Abstract])) OR ("JA"[Title/Abstract])) OR ("Stomach neoplasm"[Title/Abstract])) OR ("Intestine neoplasm"[Title/Abstract])) OR ("esophagus neoplasm"[Title/Abstract])) OR ("colon neoplasm"[Title/Abstract])) OR ("rectum neoplasm"[Title/Abstract])) OR ("anus neoplasm"[Title/Abstract])) OR ("colorectal neoplasm"[Title/Abstract])) OR ("pancreas neoplasm"[Title/Abstract])) OR ("liver neoplasm"[Title/Abstract]))) OR ("Ileal Neoplasm"[Title/Abstract])) OR ("Jejunal Neoplasm"[Title/Abstract])) OR ("Gastrointestinal Neoplasm"[Title/Abstract]) AND (((((("Immunotherapy"[Title/Abstract]) OR ("immuno-therapeutic"[Title/Abstract])) OR ("immune system therapy"[Title/Abstract])) OR ("immune-based therapy"[Title/Abstract])) OR ("immune-targeted therapy"[Title/Abstract])) OR ("immune-based treatment"[Title/Abstract])) OR ("sipuleucel-t"[Title/Abstract])) OR ("Immunotherapy"[Mesh]) WOS ((((((((TS = ("Artificial intelligence network")) ORTS = ("Artificial intelligence networks")) ORTS = ("Artificial intelligence")) ORTS 294 = ("Machine Intelligence")) ORTS = ("AI")) ORTS = ("Machine learning")) ORTS = ("ML")) ORTS = ("Deep learning")) ORTS = ("DL")) trointestinal cancers")) ORTS = ("Gastrointestinal neoplasm")) ORTS = ("Gastrointestinal tumor")) ORTS = ("GI cancer")) ORTS = ("GI neoplasm")) OR TS = ("Gastrointestinal stromal tumor")) OR TS = ("GIST")) OR TS = ("Digestive system neoplasm")) OR TS = ("Stomach Cancer")) ORTS = ("Intestine Cancer")) ORTS = ("small bowel cancer")) ORTS = ("esophagus Cancer")) ORTS = ("colon cancer")) ORTS = ("rectum Cancer")) ORTS = ("anus Cancer")) ORTS = ("colorectal cancer")) ORTS = ("pancreas cancer")) ORTS = ("liver cancer")) ORTS = ("ileum cancer")) OR TS = ("Jeiunal adenocarcinoma")) OR TS = ("JA")) OR TS = ("Stomach neoplasm")) OR TS = ("Intestine neoplasm")) ORTS = ("small bowel neoplasm")) ORTS = ("esophagus neoplasm")) ORTS = ("colon neoplasm")) ORTS = ("rectum neoplasm")) ORTS = ("anus neoplasm")) OR TS = ("colorectal neoplasm")) OR TS = ("pancreas neoplasm")) OR TS = ("liver neoplasm")) O plasm")) OR TS = ("Jejunal Neoplasm") AND (((((((TS = ("Immunotherapy")) OR TS = ("immuno-therapeutic")) OR TS = ("immune system therapy")) OR TS = ("immune-based therapy")) OR TS = ("immune-targeted therapy")) OR TS = ("immune-based treatment")) OR TS = ("sipuleucel-t")) OR TS = ("immune response therapy")) OR TS = ("immune cell therapy") Scopus (TITLE-ABS-KEY ("Artificial intelligence network") OR TITLE-ABS-KEY ("Artificial intelligence networks") OR TITLE-ABS-KEY ("Artificial intel-1003 ligence") OR TITLE-ABS-KEY ("Machine Intelligence") OR TITLE-ABS-KEY ("AI") OR TITLE-ABS-KEY ("Machine learning") OR TITLE-ABS-KEY ("ML") OR TITLE-ABS-KEY ("Deep learning") OR TITLE-ABS-KEY ("DL") OR TITLE-ABS-KEY ("convolutional neural network") OR TITLE-ABS-

("ML") OR TITLE-ABS-KEY ("Deep learning") OR TITLE-ABS-KEY ("DL") OR TITLE-ABS-KEY ("convolutional neural network") OR TITLE-ABS-KEY ("CNN")) AND (TITLE-ABS-KEY ("Gastrointestinal cancer") OR TITLE-ABS-KEY ("Gastrointestinal cancers") OR TITLE-ABS-KEY ("Gastrointestinal neoplasm") OR TITLE-ABS-KEY ("Gastrointestinal tumor") OR TITLE-ABS-KEY ("Gastrointestinal cancers") OR TITLE-ABS-KEY ("Gi neoplasm") OR TITLE-ABS-KEY ("Gastrointestinal stromal tumor") OR TITLE-ABS-KEY ("GI cancer") OR TITLE-ABS-KEY ("Gi neoplasm") OR TITLE-ABS-KEY ("Stomach Cancer") OR TITLE-ABS-KEY ("Intestine Cancer") OR TITLE-ABS-KEY ("Digestive system neoplasm") OR TITLE-ABS-KEY ("Stomach Cancer") OR TITLE-ABS-KEY ("Intestine Cancer") OR TITLE-ABS-KEY ("small bowel cancer") OR TITLE-ABS-KEY ("esophagus Cancer") OR TITLE-ABS-KEY ("colon cancer") OR TITLE-ABS-KEY ("rectum Cancer") OR TITLE-ABS-KEY ("anus Cancer") OR TITLE-ABS-KEY ("colorectal cancer") OR TITLE-ABS-KEY ("pancreas cancer") OR TITLE-ABS-KEY ("anus neoplasm") OR TITLE-ABS-KEY ("colon neoplasm") OR TITLE-ABS-KEY ("small bowel neoplasm") OR TITLE-ABS-KEY ("sophagus neoplasm") OR TITLE-ABS-KEY ("colon neoplasm") OR TITLE-ABS-KEY ("rectum neoplasm") OR TITLE-ABS-KEY ("anus neoplasm") OR TITLE-ABS-KEY ("colorectal neoplasm") OR TITLE-ABS-KEY ("pancreas neoplasm") OR TITLE-ABS-KEY ("anus neoplasm") OR TITLE-ABS-KEY ("colorectal neoplasm") OR TITLE-ABS-KEY ("pancreas neoplasm") OR TITLE-ABS-KEY ("anus neoplasm") OR TITLE-ABS-KEY ("colorectal neoplasm") OR TITLE-ABS-KEY ("pancreas neoplasm") OR TITLE-ABS-KEY ("anus neoplasm") OR TITLE-ABS-KEY ("colorectal neoplasm") OR TITLE-ABS-KEY ("pancreas neoplasm") OR TITLE-ABS-KEY ("integun Neoplasm") OR TITLE-ABS-KEY ("jejunal Neoplasm")) AND (TITLE-ABS-KEY ("integunse-based therapy") OR TITLE-ABS-KEY ("immune-targeted therapy") OR TITLE-ABS-KEY ("immune based treatment") OR TITLE-ABS-KEY ("immune-targeted therapy") OR TITLE-ABS-KEY ("immune cell therapy"))

Methods

This systematic review and meta-analysis study adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [11]. The study protocol was registered in the Open Science Framework (OSF) and is accessible at the following https://doi.org/10.17605/OSF.IO/CBU8Y..

Search strategy

A literature search was conducted in PubMed, Web of Science (WOS), and Scopus up to August 31, 2024, by an independent reviewer (NM). The search strategy employed three primary conceptual groups: artificial intelligence networks, gastrointestinal cancers, and immunotherapy. Keywords within each group were combined using'OR,'and then these combined terms were further combined using'AND'to form the final search strategy. The search strategy was adjusted for query options specific to each database. Additionally, reference lists of relevant systematic reviews were manually searched to identify potential studies. The search strategy is detailed in Table 1.

Eligibility criteria

This study aimed to utilize artificial intelligence networks to assess the prognosis of gastrointestinal cancers in response to immunotherapy based on genetic mutation features. No restrictions were placed on publication date or language. To enhance search specificity, keyword searches were limited to article titles. Studies involving animal models, unrelated to the research objective, or classified as review articles, abstracts, case reports, case series, or other similar types were excluded during the initial screening process. Duplicate studies were also eliminated from the dataset.

Data extraction and study quality assessment

One independent reviewer (GN) utilized the RAYYAN intelligent tool for systematic reviews to conduct a blinded analysis and screening of titles and abstracts to identify relevant studies. In case of discrepancies, a second reviewer (MAA) was consulted to resolve the issue through discussion. Additionally, two independent reviewers (HZ and ZHM) extracted the data from the included studies. Furthermore, one independent reviewer (MAA) employed the critical appraisal tools developed by the Joanna Briggs Institute (JBI) to assess the quality and risk of bias of the included studies. The results of this assessment are presented in Supplementary Table 2.

Statistical analysis

A meta-analysis was undertaken to derive pooled estimates of sensitivity, specificity, accuracy, and the area under the curve (AUC) for evaluating the performance of artificial intelligence networks in predicting the prognosis of gastrointestinal cancer immunotherapy based on genetic mutation features. Heterogeneity among studies was examined using the Chi-square test and quantified using the I² statistic, which measures the proportion of variability across studies due to heterogeneity rather than random variation. The I² statistic was computed using the formula $100\% \times (Q - df)/Q$. Study weights were determined through the inverse variance method. A randomeffects model was applied to integrate data across studies, minimizing the effects of heterogeneity. Statistical significance was defined as P < 0.05. Data extracted from graphical figures in the studies were digitized using WebPlot Digitizer (Automeris LLC, Frisco, Texas).

In order to estimate the missing values of sensitivity, specificity, and accuracy in this meta-analysis, we employed regression models and imputation methods. These methods were used to handle the missing data in a statistically robust manner, allowing for a more accurate synthesis of study outcomes.

Regression models were utilized to predict the missing values based on the relationships between observed data points, while imputation methods (e.g., multiple imputation) were applied to replace missing values with plausible estimates, accounting for the uncertainty in the imputation process. Specifically, multiple linear regression was used as the predictive model for estimating missing performance values, followed by multiple imputation using chained equations (MICE) to account for uncertainty in the imputed data. These approaches were chosen because they minimize bias and maximize the accuracy of pooled estimates, ensuring that the results of the meta-analysis reflect the most reliable representation of the underlying data. Additionally, these methods help to preserve the statistical power of the analysis by making use of all available data, rather than excluding studies with missing values, which could introduce bias or reduce generalizability. Further, methods like maximum likelihood estimation and Bayesian approaches could also be considered for robust handling of missing data in future analyses. Although regression models and multiple imputation techniques were employed to manage missing values, it is important to recognize that these methods rely on certain assumptions, such as data being missing at random. Violation of these assumptions could potentially introduce bias or lead to inaccuracies in the pooled estimates, particularly in the context of clinical data with high variability.

Result

Our search strategy yielded 168 articles. After removing 608 duplicate records the remaining papers underwent screening by their title and abstract to be chosen according to our eligibility criteria. After removing irrelevant records that did not meet our criteria and papers that did not have a full text available online finally 45 studies were included in the synthesis of our analysis and investigation (Fig. 1).

Summary characteristics of included studies are brought in Supplementary Table 2.

Meta-analysis

A total of 45 studies published between 2020 and 2024 were included in this meta-analysis, comprising 14,047 participants in training sets and 10,885 participants in test sets. Among the test set participants, 4,156 were female (38.2%) and 5,079 were male (46.7%), with the remaining 15.1% not reporting gender. In the training sets, 3,608 participants were female (25.7%) and 3,853 were male (27.4%), while 46.9% of participants had unreported or missing gender information. This missing data reflects limitations in the original studies, many of which did not comprehensively report gender breakdowns.The mean age of all participants was 64.7 years (standard deviation [SD]: 1.35 years), with a 95% confidence interval (CI) of 62.06 to 67.5 years.

The included studies focused on 7 distinct types of gastrointestinal cancers and analyzed the impact of



Fig. 1 PRISMA flowchart of study selection procedure

mutations in 211 genes. The AI models utilized genetic mutation features to predict prognosis and immunotherapy response in these cancers. The performance metrics of the models, including the area under the curve (AUC), sensitivity, specificity, and accuracy, were pooled using a random-effects model to account for variability among studies (Figs. 2, 3, 4, 5).

Pooled performance metrics

The pooled estimates for the performance of the AI network models were as follows:

- Area Under the Curve (AUC): 0.86 (95% CI: 0.86– 0.87).
- Sensitivity: 83% (95% CI: 83%-84%).
- Specificity: 72% (95% CI: 72%-73%).
- Accuracy: 82% (95% CI: 82%-83%).

These results demonstrate the high potential of AIbased networks in assessing prognosis and predicting response to immunotherapy for gastrointestinal cancers based on genetic mutation features.

Heterogeneity analysis

The Galbraith plot indicated low to moderate heterogeneity across studies. Subgroup analyses were conducted to explore sources of heterogeneity, which revealed variations in model performance based on cancer type, genetic mutation features, and sample size (Figs. 6 and 7).

Publication bias

The funnel plot for publication bias appeared symmetrical, suggesting a lack of publication bias. Egger's test further confirmed this with a statistically insignificant result (*p*-value: 1.00), indicating that studies with smaller



Fig. 2 Forest plot of pooled AUC estimates of ML models

sample sizes or less favorable outcomes were not underrepresented in the meta-analysis (Fig. 8).

Subgroup analyses

Subgroup analyses based on cancer type revealed differences in model performance, with the highest AUC observed in studies focusing on Gastric cancers (AUC: 0.87, 95% CI: 0.79–0.96) and lower AUC in studies on pancreatic cancer (AUC: 0.52 95% CI: 0.19–0.86). Additionally, models incorporating comprehensive gene panels tended to outperform those using fewer genetic features (Fig. 6).

Discussion

The performance of AI-based prognostic models in this meta-analysis demonstrated promising but variable results. The pooled AUC of 0.86 (95% CI: 0.86–0.87) indicates a strong discriminative ability of these models in predicting immunotherapy outcomes in gastrointes-tinal cancers. Furthermore, the sensitivity (83%; 95% CI: 83%–84%) suggests that the models are generally capable of correctly identifying a substantial proportion of true responders to immunotherapy. However, the specificity

(72%; 95% CI: 72%–73%) indicates that approximately 28% of non-responders may be misclassified as likely responders, which could have implications for clinical decision-making. The accuracy of 82% further supports the overall robustness of these models. Despite these promising figures, it is important to consider that the variability across studies, as well as differences in AI architectures and input features, may impact model performance when applied to new or independent datasets. Therefore, while AI models show clear potential, further optimization and external validation are needed before broad clinical implementation.

The recent development of AI has brought a wave of optimism in the oncology field, significantly improving diagnosis, prognosis, and treatment strategy in colorectal cancer. Various studies have proposed AI and ML-based prediction models in this arena and promise a bright future for patient care and clinical decision-making.

Machine learning algorithms were applied by Suzhen Bi et al. to study the involvement of the TAS2R gene family in CRC [12]. Their finding has established that high expression of TAS2R correlates with poor survival and low immune cell infiltration, the process by which



Fig. 3 Forest plot of pooled accuracy estimates of ML models

different immune cells migrate from the bloodstream into a tumor [13], making it a potential biomarker for prognosis. They, therefore, developed a TAS2R expression-based gradient boosting machine model that would provide clinicians with a tool to guide them in patient management.

Rui Cao's team focused on predicting MSI, one of the most important CRC prognostic factors, using an Ensemble Patch Likelihood Aggregation model (EPLA) [14]. The AUC values were 0.8848 and 0.8504 for the TCGA-COAD and Asian-CRC datasets, respectively. EPLA is a practical alternative to traditional MSI testing methods, which are usually invasive and expensive and can be integrated into routine clinical workflows.

This integration of multi-omics data, which involves combining various biological datasets such as genomics, transcriptomics, and proteomics [15], with AI promises an improved prognosis for CRC and is a major step toward personalized therapy in CRC patients. Jiamin Chen and colleagues identified a pyroptosisrelated long non-coding RNA signature that could be used to personalize therapy for CRC patients, putting the patient at the heart of these advancements [16].



Fig. 4 Forest plot of pooled sensitivity estimates of ML models

Sen Lin's single-cell multi-omics study has uncovered how immune dysfunction drives liver metastasis in CRC patients [17]. Further, this study indicates that tumorimmune interaction will contribute significantly to forecasting metastasis and prognosis. On the other hand, Zaoqu Liu found distinct tumor stemness clusters in CRC, which may provide potential directions for targeted therapy based on tumor characteristics [17].

Furthermore, Hou et al. built a gene signature using the concept of the cancer-immunity cycle that might improve the stratification of prognosis and predict immunotherapy response in CRC patients [18]. These studies demonstrate how AI, multi-omics data, and immune profiling

are increasingly integrated to further improve CRC treatment strategies and form a basis for personalized therapeutic approaches.

Hepatocellular carcinoma (HCC)

AI has also shown transformative potential in hepatocellular carcinoma, one of the significant types of liver cancer. Chen et al. developed a new classifier based on cancer stem cell features using RNA sequencing data from TCGA and ICGC datasets [19]. Their classifier utilized a stemness index mRNA approach that stratified patients into high and low-stemness subtypes. Accordingly, the high-stemness tumors had lower immune



Fig. 5 Forest plot of pooled specificity estimates of ML models

infiltrations and resisted immunotherapies, proving that stemness has an important role in immune escape and therapeutic resistance. The developed classifier showed a high AUC of 0.953 in the training set, reflecting strong predictive capability. AUC values range from 0 to 1, with one being perfect and 0.5 representing a random prediction; an AUC of 0.953 is considered high and thus indicates a strong predictive model.

Further refining the approach, in a successive study, Chen et al. widened their scope to include immunerelated gene signatures in developing a nine-gene signature that predicted one-year survival in HCC patients, with an AUC of 0.8. This study indicated that high mRNA scores reflect a more immunosuppressive tumor microenvironment and predict poor responses to ICIs, marking a significant shift in predicting cancer prognosis [20].

Cheng et al. integrated ERSRGs into the prognosis of HCC, establishing an ANN model with an AUC of 0.979. Their work not only demonstrated the predictive power of AI in this setting but also described how ER stress contributes to the advancement and immune regulation of the tumor. With the incorporation of molecular data, this model can give a more accurate prognosis and prediction of therapy compared to the traditional approach based on imaging [21].

Focusing on NK cell-related genes through singlecell RNA sequencing, Feng et al. developed an 11-gene

Study		Effect size with 95% CI	Weight
Colon cancer		mar 3576 Of	(/0)
Bi S. et al.		0.90 [0.89] 0.93	3.38
Zhong X. et al.		0.71 [0.63 0.80]	3.16
Heterogeneity: $\tau^2 = 0.02$, $I^2 = 94.99\%$. $H^2 = 19.96$		0.81 [0.63. 1.00]	0.10
Test of $\theta_i = \theta_i$: Q(1) = 19.96, p = 0.00		,	
Test of $\theta = 0$: z = 8.66, p = 0.00			
Colorectal cancer			
Dong Z, et al.	_ =	0.85 [0.82, 0.88]	3.34
Yang M, et al.		0.72 [0.71, 0.74]	3.38
Jiang M, et al.		0.68 [0.38, 0.97]	1.71
Han X, et al.		0.96 [0.94, 0.99]	3.37
Zhou L, et al.		0.71 [0.69, 0.74]	3.36
Heterogeneity: $\tau^2 = 0.01$, $I^2 = 98.82\%$, $H^2 = 84.85$		0.80 [0.69, 0.91]	
Test of $\theta_i = \theta_j$: Q(4) = 439.33, p = 0.00			
Test of θ = 0: z = 14.51, p = 0.00			
Esophageal cancer			
Wang J, et al.		0.70 [0.68, 0.72]	3.37
Heterogeneity: $\tau^2 = 0.00$, $I^2 = .\%$, $H^2 = .$	♦	0.70 [0.68, 0.72]	
Test of $\theta_i = \theta_j$: Q(0) = 0.00, p = .			
Test of θ = 0: z = 64.09, p = 0.00			
Gastric canser			
Jiang D, et al.	_	0.69 [0.66. 0.72]	3.35
Wang W, et al.		0.94 [0.91. 0.96]	3.37
He W, et al.		0.98 [0.95, 1.00]	3.37
Wan R, et al.		0.97 [0.93, 1.02]	3.31
Wang J, et al.	· · · · · · · · · · · · · · · · · · ·	0.80 [0.75, 0.85]	3.30
Wang X, et al.		0.75 [0.59, 0.91]	2.66
Wei Z, et al.		0.94 [0.93, 0.95]	3.38
Heterogeneity: τ ² = 0.01, I ² = 98.47%, H ² = 65.48		0.87 [0.79, 0.96]	
Test of $\theta_i = \theta_j$: Q(6) = 263.73, p = 0.00	· · · · · · · · · · · · · · · · · · ·		
Test of θ = 0: z = 19.76, p = 0.00			
Heneteeslluler earsing ma			
	_	0.9510.92 0.981	3 35
Chen E et al		0.55[0.52, 0.56]	3.24
Wang T et al			3.38
Zheng J, et al.		0.75[0.73, 0.77]	3.37
Wu L. et al.		0.79 [0.63, 0.95]	2.66
Chen X. et al.		0.73 [0.67. 0.79]	3.28
Cai X, et al.		0.68 [0.62, 0.75]	3.24
He Y, et al.		0.69 [0.64, 0.73]	3.32
Xu X, et al.	-	0.75 [0.70, 0.80]	3.30
Shi Y, et al.		0.95 [0.93, 0.97]	3.37
Song F, et I.	-	0.85 [0.79, 0.91]	3.26
Heterogeneity: τ ² = 0.01, I ² = 98.29%, H ² = 58.50	•	0.81 [0.74, 0.88]	
Test of $\theta_i = \theta_j$: Q(10) = 731.77, p = 0.00			
Test of θ = 0: z = 23.59, p = 0.00			
Pancraatic cancer			
Jiang L, et al.		0,70 [0,64. 0.76]	3.25
Su H, et al.	• -	0.19 [0.15. 0.23]	3.33
Yu X, et al.		0.69 [0.62, 0.76]	3.19
Heterogeneity: τ^2 = 0.08, I ² = 98.98%, H ² = 97.81		0.52 [0.19, 0.86]	
Test of $\theta_i = \theta_j$: Q(2) = 250.60, p = 0.00			
Test of θ = 0: z = 3.11, p = 0.00			
Stomach adenocarcinoma	_	0.731.0.00 0.70	2.24
Dan MSC Z, et al.	-	0.73 [0.69, 0.76]	3.34
MOD Q, et al.		0.77 [0.73, 0.81]	3.33
Therefore the end of $P_{1} = 0.00$, $f^{-} = 64.58\%$, $H^{-} = 2.82$ Therefore $P_{2} = 0.0(1) = 2.82$, $r_{2} = 0.00$		0.75[0.71, 0.79]	
Test of $\theta = 0$; $u(1) = 2.02$, $p = 0.09$ Test of $\theta = 0$; $z = 34.03$, $p = 0.00$			
rest 01 0 = 0. 2 = 34.03, p = 0.00			
Overall	•	0.79 [0.73, 0.84]	
Heterogeneity: τ^2 = 0.02, I^2 = 99.28%, H^2 = 138.52	Ţ		
Test of $\theta_i = \theta_j$: Q(30) = 3420.51, p = 0.00			
Test of θ = 0: z = 27.67, p = 0.00			
Test of group differences: $Q_b(6) = 26.44$, p = 0.00		-	
	.2 .4 .6 .8	1	

Random-effects REML model





Fig. 7 Galbraith plot for heterogeneity assessment

prognostic signature for HCC [22]. This signature illustrates that low-risk groups benefit more from immune therapies such as PD1 blockers and have better overall survival. The work further elucidates the integrated immune-related genetic signature with AI in personalizing HCC therapy.

In another important study, Li et al. constructed lysosome-related gene-based prognostic models for HCC [23]. Their research evidenced the role of AI-based models in immune response and drug sensitivity estimation, with a major emphasis on genomic data regarding cancer prognosis and treatment prediction.

Gastric cancers

Several contemporary publications on gastric cancer have demonstrated its potency in improving prognosis by personalizing the treatment approach. Chen et al. investigated immune-related genomic alterations within gastric cancer using deep learning algorithms and singlecell sequencing [24]. Indeed, specific genomic changes were associated with poorer survival, underlining the relevance of immune escape mechanisms within gastric cancer prognosis.

Deng et al. identified an immune cell infiltration- and immune-related biomarker-centric degradome-based prognostic signature [25]. Using ten machine learning models, their model showed the AUC values as equally fantastic: 0.976, 0.900, and 0.976 for mRNA expression, CNV, and DNA methylation, respectively. Moreover, high-risk scores associated with lower chemotherapy sensitivity highlighted the possibility that immune markers might guide more effective therapy approaches. Jiang et al. integrated radiological imaging with deep learning to predict the TME status and responses to immunotherapies in gastric cancer [26]. Their model was superior to the traditional clinicopathologic variables for noninvasive prediction of patient outcomes and personalized treatment approaches. This study, therefore, points to the increasing potential of radionics in integrating imaging data with genomic features.

Li et al. developed a machine learning-based programmed cell death-related signature; the AUC values were 0.771, 0.751, and 0.827 for 1-year, 3-year, and 5-year survival predictions, respectively. Their findings underlined the impact of immune responses on treatment outcomes, opening a new direction for personalized treatment approaches [27].

Liu et al. proposed a biologically informed graph neural network model, PGLCN, for predicting tumor mutation burden and immunotherapy response in gastric cancer. This approach integrates multi-omics data of mRNA expression, CNV, and DNA methylation with AUC values of 0.948, 0.910, and 0.791 for STAD, COAD, and UCEC, respectively. Compared with other traditional machine learning models, their model performed much better and found some important biomarkers of immune cell infiltrations that would provide further scope in the personalized treatment regime [27].

Other cancers

Besides colorectal, hepatocellular, and gastric cancers, an increasing number of studies have investigated the use of AI in both esophageal and pancreatic malignancies. For instance, in esophageal cancer, Liu et al. developed the



Fig. 8 Funnel plot demonstrating a symmetric view confirms the lack of publication bias

MRCRTR score to predict patients'prognosis and resistance to chemoradiotherapy based on the expression of mitochondria-related genes. MRCRTR score correlated with immune escape and biological processes, including angiogenesis [28].

AI-based models have integrated data in various other types of cancer, too. Itgenomic, metabolic, and immune data in pancreatic cancer have been integrated to predict patient outcomes. Chen et al. found 425 super-enhancerassociated genes critical for the development of cancer. The SEMet classifier developed could predict responses to immunotherapy [29]. Guo et al. developed a metabolic biomarker signature (MBS) related to improved survival outcomes in patients with high immune cell infiltration [30]. These studies epitomize the personalized treatment approach through AI integrated with multi-omics data and, therefore, show the broad application of AI in oncology.

The notably lower AUC observed for pancreatic cancer models (AUC =0.52) compared to gastric cancer (AUC =0.87) may be attributed to several well-known biological factors. Pancreatic cancer is widely recognized as an immunologically"cold"tumor, characterized by a profoundly immunosuppressive tumor microenvironment, including abundant regulatory T cells, myeloid-derived suppressor cells, and tumor-associated fibroblasts, which hinder effective immune activation. Additionally, pancreatic tumors typically exhibit a low tumor mutation burden (TMB) and minimal PD-L1 expression, both of which limit the ability of immune checkpoint inhibitors to produce favorable responses [31]. These tumor-intrinsic features likely contributed to the reduced responsiveness to immunotherapy reported in the included studies

and may explain the poor performance of AI models trained on pancreatic cancer data. Furthermore, the limited availability of high-quality, large-scale datasets for pancreatic cancer may also have impaired model training and generalization [32].

While TMB has emerged as a potential biomarker for predicting immunotherapy responses, its clinical applicability in gastrointestinal cancers (GICs) is not yet fully validated. The predictive value of TMB varies considerably across GIC subtypes, and the inconsistency of its performance has been noted in several included studies. Additionally, methodological differences in calculating TMB and the absence of standardized cut-off points further complicate its utility in clinical practice. Therefore, despite its promise, TMB should be interpreted cautiously when considered for immunotherapy decisionmaking in GICs.

To enhance the generalizability of AI models for predicting immunotherapy outcomes in gastrointestinal cancers, several methodological strategies should be considered. First, employing rigorous cross-validation techniques, such as k-fold cross-validation or nested cross-validation, can provide more reliable estimates of model performance and mitigate overfitting during model development. Second, external validation using independent datasets from different institutions or populations is crucial for assessing model robustness in diverse clinical settings. Third, transfer learning, which allows models pre-trained on large datasets to be fine-tuned on smaller, domain-specific datasets, could improve performance, especially in cancers with limited sample sizes. Finally, the integration of multi-center and multi-omics datasets, when available, could further

enhance the model's ability to generalize across heterogeneous patient populations. Implementing these strategies will be critical for the development of AI systems capable of supporting clinical decision-making in real-world settings.

Based on the synthesis of current evidence, several recommendations can be made to enhance the development and implementation of AI models in gastrointestinal cancer immunotherapy. First, future AI models should prioritize multi-omics data integration (genomics, transcriptomics, imaging, and immune profiling) to capture the complexity of tumor biology more comprehensively. Second, AI systems should be developed and validated using multicenter datasets to ensure robustness and generalizability across diverse populations. Third, incorporating external validation and prospective studies will be essential to confirm model reliability and prevent overfitting. From an ethical and practical perspective, enhancing model transparency through explainable AI approaches will be critical to ensure clinician trust and patient safety. Finally, collaborations between AI researchers, oncologists, geneticists, and ethicists should be strengthened to create clinically viable AI-driven decision-support systems tailored to immunotherapy in gastrointestinal cancers.

Attributed to variations in cancer types, AI methodologies, and genetic input features. Gastric cancer models performed relatively better than other cancer types, possibly due to the availability of richer datasets or more homogeneous patient populations. Conversely, pancreatic cancer models showed lower performance, reflecting the known challenges of imaging and genetic profiling in pancreatic tumors. Additionally, models incorporating larger gene panels generally outperformed those with fewer genetic features.

Notably, one of the challenges identified in this metaanalysis is the inconsistency in the reported performance of Artificial Neural Networks (ANNs) across different studies. While several studies demonstrate high predictive accuracy for ANN-based models, others report moderate to poor performance. These discrepancies may stem from variations in sample sizes, cancer subtypes, gene panel compositions, preprocessing strategies, or ANN architectures. Furthermore, differences in reporting standards and the absence of external validation in some studies may further contribute to these conflicting results. Therefore, although the pooled estimates suggest that ANNs have potential in predicting immunotherapy outcomes, these results should be interpreted with caution.

One of the limitations observed in this review is the limited use of external validation across the included studies. The majority of AI-based models were evaluated using internal validation techniques such as cross-validation or random data splits. However, without external validation on independent cohorts or datasets, the generalizability of these models to other populations, healthcare settings, and imaging protocols remains uncertain. This gap highlights the need for future research to prioritize external validation to enhance the robustness and clinical applicability of AI models.

Limitations

A key limitation of this meta-analysis is the inconsistency of findings regarding ANN performance across included studies. The variability in model performance may undermine the generalizability of our results. Heterogeneity in AI model design, training data, patient characteristics, and genetic features used could explain these contradictions. Future studies with standardized methodologies and external validation are essential to resolve this inconsistency and enhance the reliability of ANN-based prognostic models.

Another limitation of this review is the uncertain predictive value of TMB as a biomarker for immunotherapy in GICs. Although widely studied, the lack of consensus regarding its clinical validity and methodological inconsistencies limits its utility. Further research with standardized definitions and large-scale validation is essential to determine its role in guiding immunotherapy in GIC patients.

A further limitation concerns the handling of missing data. Although we used regression-based imputation methods to reduce bias and preserve statistical power, these techniques inherently assume that missingness is random or can be predicted based on available data. However, clinical datasets often exhibit complex patterns of missingness, and unmeasured confounders could have influenced the imputed values. Therefore, the results should be interpreted with caution, acknowledging the potential for bias introduced by the imputation process.

Despite the subgroup analyses, a portion of heterogeneity remained unexplained, likely due to variations in data quality, patient populations, or AI model implementation details not fully reported in the included studies [33]. Further studies with standardized reporting and more detailed methodological descriptions are needed to clarify these factors.

Data variability affects AI algorithms

AI algorithms for assessing the prognosis of gastrointestinal cancers and their response to immunotherapy can be influenced by data variability across diverse populations. Differences in genetic mutations, clinical features, and tumor biology between populations can limit the ability of AI models to generalize. For example, certain mutations may be more prevalent in one ethnic group, and clinical characteristics like age, sex, and comorbidities may vary across populations. If AI models are trained primarily on data from a specific population, they may not perform well on others, leading to reduced accuracy and biased predictions.

This variability can result in AI models that overestimate or underestimate the efficacy of treatments for certain patient groups, particularly those from underrepresented populations. The lack of diversity in training datasets can also amplify existing healthcare disparities, where minority or rural populations may not benefit from AI-driven predictions, despite potentially needing them the most.

To address this, AI models should be trained on diverse datasets that include a broad range of demographic and clinical features. Stratifying models based on patient characteristics, such as ethnicity or cancer subtype, can also improve prediction accuracy for specific groups. Additionally, continuous validation on multiple datasets from diverse populations and incorporating multi-omics data can help ensure that AI models generalize better and provide more personalized treatment recommendations. By addressing these challenges, AI can more reliably predict responses to immunotherapy and enhance treatment outcomes across all patient populations.

Several drawbacks associated with using AI models in gastrointestinal cancer are related to their broader applicability in a clinical setting.

First, the quality and uniformity of the datasets used in such studies pose serious challenges. Many AI models rely on open-source datasets, like The Cancer Genome Atlas and the International Cancer Genome Consortium, which are usually incomplete and lack uniformity. This will eventually limit their prediction accuracy. Besides, multi-hospital or multi-population data bears inherent variability, making generalization to diverse patient cohorts difficult.

Another important area that needs improvement is the imbalance in training data. Many AI algorithms perform well for the most frequent patient groups but fail to predict the outcomes of rarer subpopulations, such as patients with rare genetic mutations or those from underrepresented demographics. This may lead to biased predictions, especially in infrequent cancers with unique patient attributes.

Moreover, multi-omics data integration is still a technically challenging task, primarily due to the heterogeneity of these datasets. Variability in scale, missing values, and differing quality levels impede the development of robust AI models. Overfitting is also a considerable concern. While AI models may perform impressively within their training environments, their generalization to clinical applications is often better due to the significant variability inherent in-patient data. In this sense, the ability for such models to generalize into more practical applications is often insufficient, barring further validation and refinement of the AI models.

A major challenge is the limited number of studies focusing on this specific intersection. Many existing studies use small, institution-specific datasets, which limits generalizability. The data is often heterogeneous, with variations in patient demographics, cancer subtypes, genetic panels, and treatment approaches. Furthermore, proprietary datasets are commonly used in AI model development, making external validation difficult. These factors contribute to a fragmented body of evidence, complicating systematic reviews or meta-analyses.

The variability and complexity of genetic data present significant hurdles. Different studies may rely on distinct sets of genetic mutations or genomic features, making it hard to standardize the input data for AI algorithms. Additionally, the quality and completeness of genetic data vary widely across studies, and the multifactorial nature of genetic interactions with immunotherapy outcomes is often not fully captured in AI models. Key factors such as the tumor microenvironment and immune system markers are frequently underrepresented.

Validation and Benchmarking External validation of AI models is often lacking. Models are rarely tested across diverse datasets or clinical settings, which reduces their reliability in real-world scenarios [33, 34]. Additionally, performance metrics such as sensitivity, specificity, and accuracy are inconsistently reported, making it challenging to compare models or integrate findings. The absence of longitudinal studies to assess long-term outcomes also limits the robustness of existing research.

Clinical Integration and Applicability Integrating AI-based prognostic tools into clinical practice poses significant challenges. These include the need for substantial infrastructure, training for healthcare providers, and adjustments to existing workflows. Regulatory barriers further complicate the adoption of AI models, as approval processes for AI tools, particularly those incorporating genetic data, are still evolving. Ethical considerations, such as data privacy and the potential misuse of genetic information, also need to be addressed [34, 35].

The rapid advancements in both AI and genomics add another layer of complexity. As new methodologies and findings emerge, synthesizing the latest data becomes challenging. This evolving nature of the field makes it difficult to establish a stable foundation for systematic reviews or meta-analyses.

Addressing these limitations requires collaborative efforts, larger and more diverse datasets, standardized methodologies, and robust validation processes. Only through such advancements can the potential of AI in personalizing GI cancer immunotherapy be fully realized.

Future aspects

Despite these challenges, artificial intelligence possesses tremendous potential to improve the diagnosis and treatment of cancers. One advantage is the high predictive accuracy of AI models, especially those that incorporate multi-omics data, which encompasses genetic, transcriptomic, and immune profiling information. Such models can enable more accurate predictions of survival rates and therapeutic responses, laying the foundation for personalized medicine.

More interestingly, AI will be instrumental in early detection and risk stratification of gastrointestinal cancer. The models will aid in the early identification of high-risk patients by analyzing both imaging data and biomarker profiles, hence allowing timely interventions that will improve clinical outcomes.

Another promising aspect of AI involves its use in devising personalized treatment. AI algorithms can synthesize multiple data sources into tumor or immune profiles that correspond with drug responses or resistance, enabling clinicians to offer therapy tailored to those characteristics.

Future efforts should address the ethical and regulatory challenges of using AI in this context. This includes Developing clear frameworks for the approval and oversight of AI models. Ensuring equitable access to AI technologies across different healthcare systems. Protecting patient privacy when handling sensitive genetic and clinical data.

The use of AI models in predicting immunotherapy outcomes based on genetic data raises significant ethical considerations. Patient privacy and data security are paramount, as genetic information is uniquely identifiable and potentially sensitive. Ensuring that data collection, sharing, and AI model training comply with ethical and legal standards, such as obtaining informed consent and adhering to data protection regulations (e.g., GDPR, HIPAA), is essential. Additionally, algorithmic bias poses a risk when AI models are trained on datasets that underrepresent certain populations, leading to unequal predictive performance across demographic groups. This could exacerbate existing healthcare disparities. Furthermore, the"black-box"nature of many AI models, particularly deep learning architectures, limits their transparency and interpretability, which may hinder clinical acceptance. Improving the explainability of AI models and incorporating fairness assessments should be prioritized in future research to ensure that AI-driven clinical decision-support systems are both trustworthy and ethically sound.

Finally, AI-driven decision-support systems could enhance clinical decision-making through evidencebased recommendations. These systems support clinicians by navigating through complex datasets, identifying potential complications, and predicting the response to treatment, hence increasing the quality and efficiency of the clinical decisions made within cancer treatment.

Future advancements will rely on strong collaboration among AI researchers, oncologists, geneticists, immunologists, and bioinformaticians. Multidisciplinary teams can bridge the gap between technological innovation and clinical application, ensuring that AI systems are both scientifically rigorous and practical for patient care. By addressing these future directions, AI networks can become transformative tools in predicting GI cancer prognosis and guiding immunotherapy decisions, ultimately improving outcomes for patients and advancing the field of precision oncology.

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s12876-025-03884-1.

Supplementary Material 1 Supplementary Material 2

Informed consent Not applicable.

Clinical trial number

Not applicable

Human ethics and consent to participate declarations Not applicable.

Authors' contributions

Study concept and design: MAA,MA. Acquisition of the data: NM,PE,AA. Analysis and interpretation of the data: MAA. Drafting of the manuscript: ZM,KB,SV,HM,NN,HZ,MN,HZB, MNV,GN,KB,FK. Critical revision of the manuscript for important intellectual content:MAA, MA,FK. Administrative, technical, and material support: MAA, M.A. Study supervision: MAA.

Funding

This study was not supported or funded by any institution.

Data availability

The data is available upon reasonable request from the corresponding author.

Declarations

Ethics approval and consent to participate Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran. ²Faculty of Medicine, İstanbul Yeniyuzyil

University, Istanbul, Turkey. ³Department of Radiology, University of Michigan, Michigan, USA. ⁴Student Research Committee, Babol University Medical Science, Babol, Iran. ⁵Faculty of Medicine, Kermanshah University of Medical Sciences, Kermanshah, Iran. ⁶School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ⁷Student Research Committee, Abadan University of Medical Sciences, Abadan, Iran. ⁸Student Research Committee, Zanjan University of Medical Sciences, Zanjan, Iran.⁹Universal Scientific Education and Research Network (USERN), Tehran, Iran. ¹⁰Student Research Committee, Faculty of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran. ¹¹Mersey & West Lancashire Teaching Hospitals NHS Trust, Southport Hospital, Lancashire, UK. ¹²Tabriz University of Medical Sciences, Gastroenterology and Hepatology Research Center, Tabriz, Iran. ¹³Toxicological Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ¹⁴Student Research Committee, School of Medicine, Zahedan University of Medical Sciences, Zahedan, Iran. ¹⁵College of Medicine, University of Arizona, Tucson, USA

Received: 18 January 2025 Accepted: 10 April 2025 Published online: 29 April 2025

References

- 1. Ye B, Li Z, Wang Q. A novel artificial intelligence network to assess the prognosis of gastrointestinal cancer to immunotherapy based on genetic mutation features. Front Immunol. 2024;15:1428529.
- Chong X, Madeti Y, Cai J, Li W, Cong L, Lu J, et al. Recent developments in immunotherapy for gastrointestinal tract cancers. J Hematol Oncol. 2024;17(1):65.
- Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoediting. Annu Rev Immunol. 2004;22:329–60.
- Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. Immunity. 2013;39(1):1–10.
- Yang B, Cheng C, Zhou J, Ni H, Liu H, Fu Y, et al. Al-powered genomic mutation signature for predicting immune checkpoint inhibitor therapy outcomes in gastroesophageal cancer: a multi-cohort analysis. Discov Oncol. 2024;15(1):507.
- 6. Fuchs CS, Doi T, Jang RW, Muro K, Satoh T, Machado M, et al. Safety and efficacy of pembrolizumab monotherapy in patients with previously treated advanced gastric and gastroesophageal junction cancer: phase 2 clinical KEYNOTE-059 trial. JAMA Oncol. 2018;4(5):e180013-e.
- Moeckel C, Bakhl K, Georgakopoulos-Soares I, Zaravinos A. The efficacy of tumor mutation burden as a biomarker of response to immune checkpoint inhibitors. Int J Mol Sci. 2023;24(7):6710.
- Qiu T, Shi X, Wang J, Li Y, Qu S, Cheng Q, et al. Deep learning: a rapid and efficient route to automatic metasurface design. Adv Sci (Weinh). 2019;6(12):1900128.
- Zhao S, Wang L, Ding W, Ye B, Cheng C, Shao J, et al. Crosstalk of disulfidptosis-related subtypes, establishment of a prognostic signature and immune infiltration characteristics in bladder cancer based on a machine learning survival framework. Front Endocrinol (Lausanne). 2023;14:1180404.
- Sunkara P. Predicting melanoma immunotherapy efficacy: neural network models with gene expression and clinical data. 2024. https:// doi.org/10.20944/preprints202405.1160.v1.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Rev Esp Cardiol (Engl Ed). 2021;74(9):790–9.
- Bi S, Zhu J, Huang L, Feng W, Peng L, Leng L, et al. Comprehensive analysis of the function and prognostic value of TAS2Rs family-related genes in colon cancer. Int J Mol Sci. 2024;25(13): 6849.
- Melssen MM, Sheybani ND, Leick KM, Slingluff CL. Barriers to immune cell infiltration in tumors. J Immunother Cancer. 2023;11(4): e006401.
- Cao R, Yang F, Ma S-C, Liu L, Zhao Y, Li Y, et al. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in colorectal cancer. Theranostics. 2020;10(24):11080–91.
- Krassowski M, Das V, Sahu SK, Misra BB. State of the field in multi-omics research: from computational needs to data mining and sharing. Front Genet. 2020;11:610798.
- Chen J, Jin D, Shao L, Wang L, Zhou L, Cai J. Machine learning-derived multi-omics prognostic signature of pyroptosis-related IncRNA with

regard to ZKSCAN2-DT and tumor immune infiltration in colorectal cancer. Comb Chem High Throughput Screen. 2024;27(8):1161–74.

- Lin S, Ma L, Mo J, Zhao R, Li J, Yu M, et al. Immune cell senescence and exhaustion promote the occurrence of liver metastasis in colorectal cancer by regulating epithelial-mesenchymal transition. Aging (Albany NY). 2024;16(9):7704–32.
- Hou Y, Zhang R, Zong J, Wang W, Zhou M, Yan Z, et al. Comprehensive analysis of a cancer-immunity cycle-based signature for predicting prognosis and immunotherapy response in patients with colorectal cancer. Front Immunol. 2022;13: 892512.
- Chen D, Liu J, Zang L, Xiao T, Zhang X, Li Z, et al. Integrated machine learning and bioinformatic analyses constructed a novel stemnessrelated classifier to predict prognosis and immunotherapy responses for hepatocellular carcinoma patients. Int J Biol Sci. 2022;18(1):360–73.
- Chen E, Zou Z, Wang R, Liu J, Peng Z, Gan Z, et al. Predictive value of a stemness-based classifier for prognosis and immunotherapy response of hepatocellular carcinoma based on bioinformatics and machinelearning strategies. Front Immunol. 2024;15:1244392.
- Cheng Z, Li S, Yang S, Long H, Wu H, Chen X, et al. Endoplasmic reticulum stress promotes hepatocellular carcinoma by modulating immunity: a study based on artificial neural networks and single-cell sequencing. J Transl Med. 2024;22(1):658.
- Feng Q, Huang Z, Song L, Wang L, Lu H, Wu L. Combining bulk and single-cell RNA-sequencing data to develop an NK cell-related prognostic signature for hepatocellular carcinoma based on an integrated machine learning framework. Eur J Med Res. 2023;28(1):306.
- Li X, Gu M, Hu Q, Weng Y, Cai X. Development and validation of metabolic models for predicting survival and immune status of hepatocellular carcinoma patients. Adv Clin Exp Med. 2023;32(12):1423–39.
- 24. Chen W, Liu X, Wang H, Dai J, Li C, Hao Y, et al. Exploring the immune escape mechanisms in gastric cancer patients based on the deep Al algorithms and single-cell sequencing analysis. J Cell Mol Med. 2024;28(10):e18379.
- Kikuchi Y, Kunita A, Iwata C, Komura D, Nishiyama T, Shimazu K, et al. The niche component periostin is produced by cancer-associated fibroblasts, supporting growth of gastric cancer through ERK activation. Am J Pathol. 2014;184(3):859–70.
- 26. Li F, Feng Q, Tao R. Machine learning-based cell death signature for predicting the prognosis and immunotherapy benefit in stomach adenocarcinoma. Medicine (Baltimore). 2024;103(10):e37314.
- 27. Liu C, Wan AH, Liang H, Sun L, Li J, Yang R, et al. Biological informed graph neural network for tumor mutation burden prediction and immunotherapy-related pathway analysis in gastric cancer. Comput Struct Biotechnol J. 2023;21:4540–51.
- Liu Z, Zeinalzadeh Z, Huang T, Han Y, Peng L, Wang D, et al. Mitochondria-related chemoradiotherapy resistance genes-based machine learning model associated with immune cell infiltration on the prognosis of esophageal cancer and its value in pan-cancer. Transl Oncol. 2024;42:101896.
- Chen D, Cao Y, Tang H, Zang L, Yao N, Zhu Y, et al. Comprehensive machine learning-generated classifier identifies pro-metastatic characteristics and predicts individual treatment in pancreatic cancer: a multicenter cohort study based on super-enhancer profiling. Theranostics. 2023;13(10):3290–309.
- Guo Y, Wang R, Shi J, Yang C, Ma P, Min J, et al. Machine learning-based integration develops a metabolism-derived consensus model for improving immunotherapy in pancreatic cancer. J Immunother Cancer. 2023;11(9):e007466.
- Liu X, Yao L, Qu J, Liu L, Lu N, Wang J, et al. Cancer-associated fibroblast infiltration in gastric cancer: the discrepancy in subtypes pathways and immunosuppression. J Transl Med. 2021;19(1):325.
- Wu Y, Han W, Dong H, Liu X, Su X. The rising roles of exosomes in the tumor microenvironment reprogramming and cancer immunotherapy. MedComm (2020). 2024;5(4):e541.
- Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ. 2020;368:m689.
- Lotter W, Hassett MJ, Schultz N, Kehl KL, Van Allen EM, Cerami E. Artificial intelligence in oncology: current landscape, challenges, and future directions. Cancer Discov. 2024;14(5):711–26.

35. Javanmard Z, Shahraki SZ, Safari K, Omidi A, Raoufi S, Rajabi M, et al. Artificial intelligence in breast cancer survival prediction: a comprehensive systematic review and meta-analysis. Front Oncol. 2025;14:1420328.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.