



OPEN ACCESS

EDITED BY

Bing Yang,
Tianjin Medical University, China

REVIEWED BY

Shitang Ma,
West Anhui University, China
Shasha Shi,
University of Colorado Anschutz Medical
Campus, United States

*CORRESPONDENCE

Jing Dai
✉ Daijing@znhospital.cn
Yongchang Wei
✉ weiyongchang@whu.edu.cn

[†]These authors share first authorship

RECEIVED 19 May 2025

ACCEPTED 25 September 2025

PUBLISHED 17 October 2025

CITATION

Lv Y, Wang Q, Xu H, Dai J and Wei Y (2025)
Machine learning-based predictive model for
immune checkpoint inhibitors response in
gastrointestinal cancers.
Front. Med. 12:1631011.
doi: 10.3389/fmed.2025.1631011

COPYRIGHT

© 2025 Lv, Wang, Xu, Dai and Wei. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Machine learning-based predictive model for immune checkpoint inhibitors response in gastrointestinal cancers

Yufan Lv^{1,2†}, Qingbin Wang^{3†}, Huiting Xu⁴, Jing Dai^{1,2*} and
Yongchang Wei^{1,2*}

¹Department of Radiation and Medical Oncology, Zhongnan Hospital of Wuhan University, Wuhan University, Wuhan, China, ²Hubei Key Laboratory of Tumor Biological Behaviors, Zhongnan Hospital of Wuhan University, Wuhan University, Wuhan, China, ³Department of Hematology, Zhongnan Hospital of Wuhan University, Wuhan University, Wuhan, China, ⁴Hubei Cancer Hospital, Wuhan, China

Introduction: Gastrointestinal (GI) cancers present significant clinical challenges characterized by dismal survival outcomes and suboptimal prognoses. Currently, only partial indicators are available to predict the response of immunotherapy. A critical gap remains in the development of models capable of accurately predicting response rates to immunotherapy regimens. In this study, we developed a machine-learning (ML) model based on factorial, molecular, demographic, and clinical data to predict the response rate.

Methods: This multicentre retrospective study analyzed the clinical data of 506 patients, comprising 352 cases collected from Zhongnan Hospital of Wuhan University and Hubei Cancer Hospital, along with 154 cases obtained from the publicly available dataset of Memorial Sloan-Kettering Hospital. We used 14 features as input features, such as the patient's basic status, biochemical test results, and genetic test results. Eight ML methods were employed to build predictive models. Through rigorous validation using seven discriminative performance metrics (accuracy, precision, recall, F1-score, ROC-AUC, PR-AUC, and Brier score), the eXtreme Gradient Boosting (XGBoost) algorithm demonstrated superior predictive capability. Model interpretability was subsequently enhanced through Shapley Additive explanations (SHAP) analysis to elucidate feature contributions.

Results: We selected XGBoost with the best predictive performance to predict response (AUC: 0.829 [95% CI: 0.72–0.91], accuracy: 78.43%, sensitivity: 86.67%, specificity: 72.31%). The Delong test and calibration curve indicated that XGBoost significantly outperformed the other models in prediction. The SHAP values indicate that chemotherapy contributes the most to the model's predictive accuracy (contribution score = 0.28), while Ki-67 exhibits the lowest contribution rate (0.01). In addition, the study showed that chemotherapy, higher hemoglobin (HGB), body mass index (BMI), age, lower neutrophil-to-lymphocyte ratio (NLR), and tumor stage positively influenced the output of the model.

Conclusion: Interpretable XGBoost models have shown accuracy, efficiency, and robustness in determining the association between input features and response rates. Among the input features, chemotherapy and tumor stage played the most important role in the prediction model. Due to the varying efficacy of ICIs in gastrointestinal cancers, personalized predictive models can greatly assist

clinical decision-making. This model fills this gap in clinical practice and can provide more precise support for personalized treatment and risk avoidance.

KEYWORDS

predictive model, immune checkpoint inhibitors, treatment, gastrointestinal malignancies, machine learning

1 Introduction

Gastrointestinal (GI) cancers are a group of diseases that seriously endanger the health of human beings, including esophageal cancers (EC) (1), gastric cancers (GC) (2) and colorectal cancers (CRC) (1, 3, 4). In recent years, the global morbidity and mortality of GI cancers have gradually increased and shown a trend of rejuvenation (5, 6). GI cancers are characterized by inconspicuous symptoms, high malignancy degree and propensity for metastasis. These pathophysiological characteristics collectively pose substantial challenges for clinical management and therapeutic intervention (7).

In recent years, immunotherapy has emerged as a transformative therapeutic paradigm, revolutionizing the treatment landscape for GI cancers (8, 9). Immune checkpoint inhibitors (ICIs) have achieved revolutionized success in hematological malignancies, yet their clinical application in GI cancers has yielded paradoxically limited therapeutic efficacy (10, 11). It has been well documented that the rate of clinical benefit in patients with GI cancers is low when ICIs are used alone (9, 12). Therefore, ICIs are usually combined with chemotherapy, radiotherapy, and targeted therapy in the treatment regimen of GI cancers (13, 14). Currently, some indicators such as tumor mutational burden (TMB) (15–17), microsatellite instability (MSI) (18–20), and PD-L1 expression (21) can initially assess the efficacy of ICIs. However, the response to therapy varies widely among patients with GI cancers. A model to predict response to combination therapy is presently lacking.

Machine learning (ML) is an important branch of artificial intelligence that has already achieved significant results in the medical field (22, 23). Currently, many studies have used ML methods to predict the prognosis of malignant tumors. However, there are still few studies on prediction models constructed by ML in GI cancers. In this study, we constructed a prediction model by ML to predict patients with GI cancers who are undergoing treatment based on ICIs. The model has a total of 14 input features, most of which have been shown to correlate with response rates. The variables incorporated included hemoglobin (HGB) (24), neutrophil-to-lymphocyte ratio (NLR) (25), sex (26), age (27), body mass index (BMI) (28, 29), cancer type, tumor stage (30), treatment modalities, and genetic test results (16). Taking whether to respond as the output target. In this study, a total of patients ($n = 506$) diagnosed with GI cancers were used as basic data. We found that most of the treatments received for GI malignancies ($n = 352$) in China were all combination therapies, so we chose the patients at Memorial Sloan-Kettering ($n = 154$) who were treated with immunotherapy alone as a control (4).

In this study, we developed a predictive framework to evaluate treatment response to ICI-based combination regimens in GI cancers. Firstly, we used eight ML methods (XGBoost, LightGBM, CatBoost, RandomForest, LR, KNN, Naive Bayes, and QDA) to comprehensively analyze the patients' 14 input features before treatment. Subsequently, the model with the best predictive performance was selected and

validated. Finally, the implementation of Shapley Additive exPlanations (SHAP) to quantify feature contributions and visualize non-linear relationships through summary plots and dependence analysis.

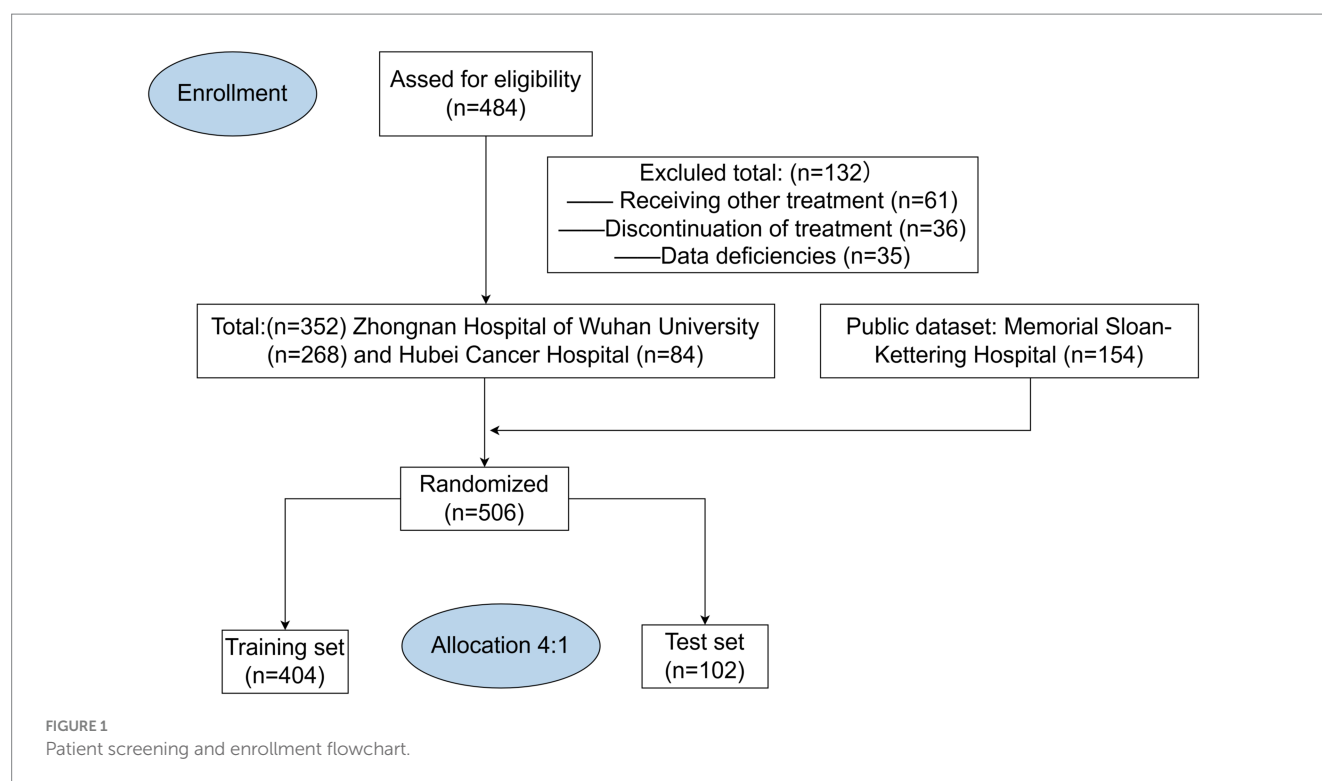
2 Methods

2.1 Patient data description

This multicentre retrospective study analyzed the clinical data of 506 patients, comprising 352 cases collected from Zhongnan Hospital of Wuhan University and Hubei Cancer Hospital, along with 154 cases obtained from the publicly available dataset of Memorial Sloan-Kettering Hospital (4). All MSK data are available online (<https://www.ioexplorer.org>). The inclusion criteria were as follows: (1) pathological diagnosis of gastrointestinal malignancy; (2) age ≥ 18 years; (3) having received at least four cycles of immunotherapy. The exclusion criteria were as follows: (1) having a primary or secondary history of cancer; (2) receiving traditional Chinese medicines, targeted therapies, or biologic therapies in the cycle of immunotherapy; (3) lack of follow-up information and clinical data. Patients initially selected for this study were those diagnosed with GI malignancies in 2021–2024 ($n = 484$), all of whom received at least four cycles of immunotherapy in the hospital. Subsequently, we retrospectively analyzed the clinical data of these patients. We excluded patients who had undergone targeted or biologic therapies during immunotherapy cycles ($n = 61$), and we further excluded patients who dropped out of treatment or died before completing four cycles of treatment ($n = 36$). At last, we excluded patients who were missing important basic clinical data ($n = 35$). After excluding all non-compliant patient data, we ultimately completed data collection from two Chinese hospitals ($n = 352$) (Figure 1).

2.2 Basic patient information and clinical data

We recorded basic health information by reviewing the nursing records before the first immunotherapy cycle, which included age, gender, and BMI. BMI was calculated as weight (KG) divided by the square of height (m^2). All clinical blood test results were within 3 days before the first immunization cycle. NLR was calculated as absolute neutrophil count (per nanoliter) divided by absolute lymphocyte count (per nanoliter). Hemoglobin (HGB) was expressed in units of $g L^{-1}$. We documented tumor type, ICB drug class, and other treatments during the ICB treatment cycle by looking at physician-recorded cases. Drug class: the patients' immunotherapy regimens were stratified into two cohorts: monotherapy with either PD-1/PD-L1



inhibitors or CTLA-4 inhibitors versus dual-agent immune checkpoint blockade combining both modalities. Cancers were staged according to the American Joint Committee on Cancer, 8th edition (31).

2.3 Genetic testing

Since numerous studies confirm that TMB is closely related to MSI (32, 33), we decided to choose MSI stability as an input feature (34). MSI: stable ($0 \leq \text{MSI score} < 3$), uncertain ($3 \leq \text{MSI score} < 10$), and unstable ($10 \leq \text{MSI score}$). In the ML model, we used two groups for MSI status: MSI unstable versus MSI stable/indeterminate. For patients with MMR deficiency, we further conduct genetic sequencing to confirm the MSI status. Gene mutations: it is well documented that HER-2 and K-RAS genes play an important role in GC and CRC and determine the prognosis of patients (35). Therefore, we incorporated the mutation status of these two genes as one of the input features in our predictive model. The mutation status of MSI, KRAS, and HER2 genes was determined using next-generation sequencing (NGS). To reduce patient costs and improve the accuracy of genetic testing, targeted sequencing panel approaches were employed for all analyses.

2.4 Ki-67 and CPS

Both CPS and Ki-67 scores were assessed through immunohistochemistry (IHC). Pathologists determined the scores by observing the percentage of Ki-67 and PD-L1 positive cells. In our study, the Ki-67 input score was based on the percentage of Ki-67 positive cells as documented in the pathology report. For PD-L1

expression (CPS score), a score greater than or equal to 1 was considered positive.

IHC Staining: tissue sections were dewaxed by immersing in xylene twice for 10 min each, followed by hydration in an alcohol gradient. Antigen retrieval was performed by placing the tissues in citrate sodium repair solution. The sections were incubated with the desired antibodies overnight at 4 °C. The next day, rapid color development was achieved using DAB, and expression levels were estimated using IHC scoring. Specific antibody catalog numbers and dilution ratios are provided in [Supplementary Table 3](#).

2.5 Response

We reviewed the doctor's case records to determine the patient's treatment outcome. Response was based on Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 (36). The primary outcome of the study was an assessment of overall treatment efficacy. Complete response (CR), partial response (PR), and stable disease (SD) were categorized as treatment effective, and progressive disease (PD) was categorized as treatment ineffective.

2.6 Model training

Data division: we divided the data of 506 patients into training (80%) and test (20%) sets using stratified random sampling, ensuring that both response rate and hospital distribution were balanced between the training set and test set.

Parameter selection: Hyper-parameter optimization was performed using the Optuna framework (37). For each model, we defined a search space (For XGBoost: We set the range of the

`n_estimators` parameter from 20 to 200, the `max_depth` parameter from 3 to 12, and the `learning_rate` parameter from 0.001 to 0.3.). The optimization objective was to maximize the mean cross-validated AUC under a five-fold stratified cross-validation scheme on the training set. Each Optuna trial was allowed to run for up to 200 iterations, and the trial with the best validation AUC was chosen. The final model was retrained on the entire training set using the best parameters. All eight ML models were trained following this procedure. Random seeds were fixed to ensure reproducibility (random seeds for python and numpy were set to 42).

2.7 ML methods and SHAP analysis

A total of eight ML methods were used in this study which are XGBoost, LightGBM, CatBoost, RandomForest, LR, KNN, Naivebayes, and QDA. We used hyperparameter optimization to optimize the performance of each ML model (38). Important metrics we used to evaluate the performance and generalization of ML models include area under the ROC curve (AUC), PR-AUC, accuracy, sensitivity, Specificity, and so on (39). From these, the best-performing model was selected and validated for analysis. SHAP is one of the most commonly used interpretability tools (40). In this study, we visualized the analysis by using the SHAP method to work out the contribution of each feature to the model output.

2.8 Handling of missing values

For the treatment of missing values, more than 35 % of the missing features were not included in our study. For models such as XGBoost, LightGBM, and CatBoost, the built-in mechanisms for

handling missing data values of these models eliminate the need for manual preprocessing. In contrast, for models including LR, KNN, Naivebayes, QDA, and Random Forest, we employed the Multiple Imputation by Chained Equations (MICE) method to impute missing values.

2.9 Statistical analysis

All analyses were performed using IBM SPSS software (version 26.0), R software (version 4.0.5), and the Python scikit-learn package (version 1.6.0). Response rates were compared by chi-square test and Fisher's exact test, we use the De-long test to compare the AUC of the different models. $p < 0.05$ was statistically significant. For full implementation details of this study, please refer to the source code repository: <https://github.com/wangqingbin/ML-Digestive-Cancer>.

3 Results

3.1 Baseline characteristics of the patient

Figure 2 illustrates the process of participant selection and study design. The basic characteristics of the 506 patients included in this study are shown in Table 1. The cohort was comprised of mostly males (65%), with a median age at diagnosis of 60 (IQR, 52–67) years. Of these patients, 44.5% had a history of surgery (patients with postoperative recurrence), median BMI was 22.65 (19.86–25.32). There were 127 (25.09%) patients diagnosed with EC, 228 (45.05%) with GC, and 151 (29.86%) with CRC. The total number of treatment responders was 300 (59.3%).

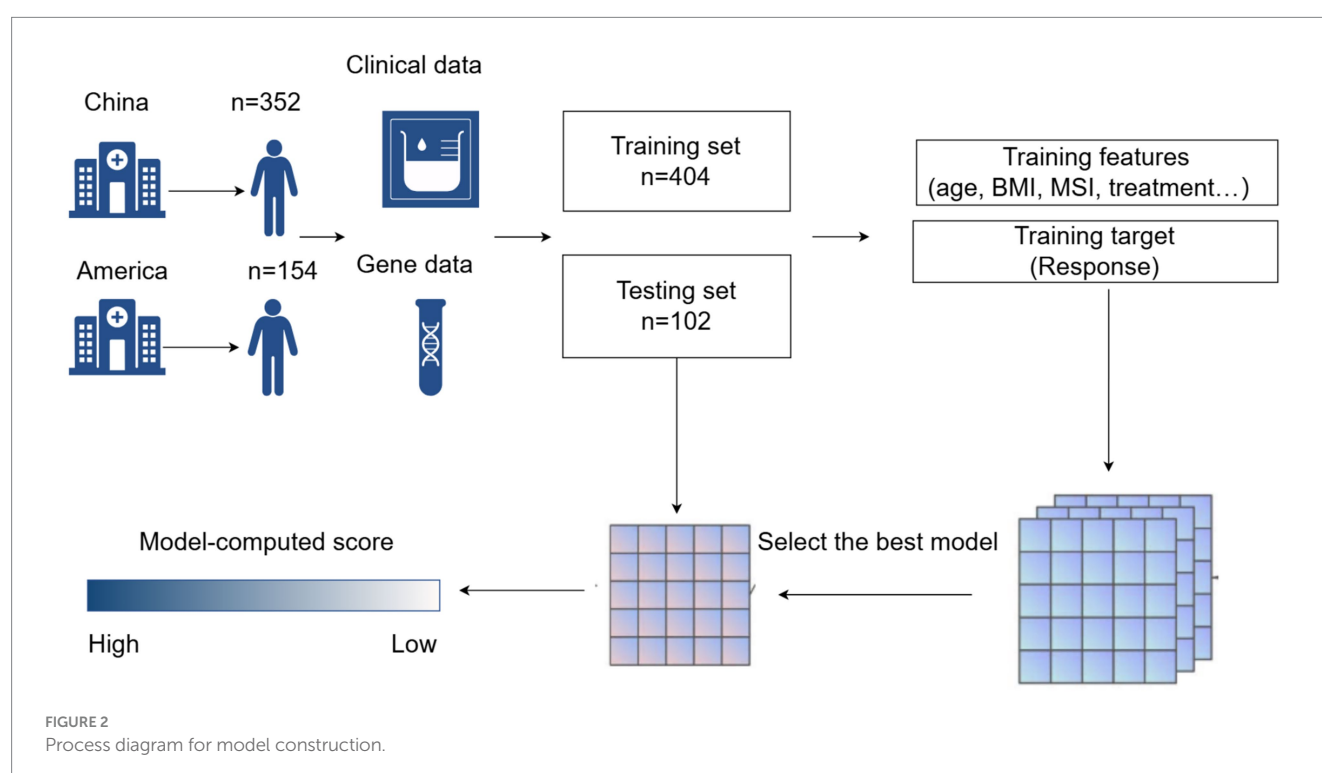


TABLE 1 Characteristics of patients in the study.

Characteristic	Total patients (<i>n</i> = 506)	Training set (<i>n</i> = 404)	Test set (<i>n</i> = 102)
Sex, <i>n</i> (%)			
Female	172	133	39
Male	334	271	63
Age, median, (IQR)	60 (52–67)	60 (52–67)	60 (53–68)
Cancer type, <i>n</i> (%)			
Esophageal	127 (25.09)	101 (25)	26 (25.49)
Gastric	228 (45.05)	182 (45.05)	46 (45.10)
Colorectal	151 (29.86)	121 (29.95)	30 (29.41)
Stage <i>n</i> (%)			
I–III	123 (24.4)	96 (23.76)	27 (26.47)
IV	383 (75.6)	308 (76.24)	75 (73.53)
Surgery history <i>n</i> (%)			
Yes	225 (44.5)	174 (43.07)	51 (50)
No	381 (55.5)	230 (56.93)	51 (50)
Response <i>n</i> (%)			
Yes	300 (59.3)	240 (59.41)	60 (58.82)
No	206 (40.7)	164 (40.59)	42 (41.18)

3.2 Machine prediction model

To predict the treatment response rate of patients with GI malignant tumors, we developed and trained eight ML models. The AUC curves of all of these models are shown in Figure 3A and the values of AUC are shown in Figure 3C. The decision curves of all models are shown in Figure 3B. The AUC value of XGBoost was 0.829 (95% CI: 0.73–0.91). The De-long test results suggested that the difference in the AUC between XGBoost and other ML models was statistically significant ($p < 0.05$). Given the imbalanced nature of our dataset, we incorporated the Precision-Recall AUC (PR-AUC) metric to comprehensively evaluate model performance beyond conventional ROC analysis, the Xgboost PR-AUC = 0.8723 (Figure 3D). Subsequently, we used metrics such as accuracy, sensitivity, and specificity to evaluate the accuracy of all models (Table 2). We showed the number of true positives, true negatives, false positives, and false negatives predicted by each model further demonstrated in the form of Figure 4. XGBoost model achieved the best performance among these methods.

3.3 SHAP analysis and importance of features

The feature importance analysis we performed on XGBoost by using an interpretable SHAP analysis approach (Figure 5). Chemotherapy scored highest in feature contribution, indicating the highest contribution to model accuracy. The lowest score was Ki67, indicating the lowest contribution to model accuracy.

3.4 Analysis of key risk factors

In the XGBoost-based feature importance analysis and SHAP analysis, treatment modality and tumor stage emerged as the two most

influential features. We performed a detailed analysis of the relationship between these features and response rate. We analyzed the effects of different treatment modalities and different tumor stages on response rates (Table 3).

4 Discussion

In recent years, ICIs have been widely used in the treatment of GI cancers (41). However, with the rise of immunotherapy, challenges have emerged. For example, the response rate remains relatively low and varies significantly among individuals in this field. How to enhance immune response rates and refine personalized immunotherapy strategies stands as a critical challenge in the field today. Therefore, we developed and trained eight ML models—XGBoost, LightGBM, CatBoost, RandomForest, LR, KNN, Naïve Bayes, and QDA—to analyze data from patients with GI cancers. Within our predictive framework, both the XGBoost and CatBoost classifiers demonstrated high predictive efficacy, achieving AUC values of 0.829 and 0.812, respectively. Further analysis revealed that the XGBoost classifier outperformed CatBoost in both accuracy and specificity metrics. Consequently, XGBoost proves to be a robust tool for accurately predicting the response of ICIs therapy. In short, these data indicate that our ML method can predict immunotherapy response rates in GI cancers with high accuracy prior to treatment.

From the baseline chart of patients, it can be seen that the incidence rate of GI cancers is much higher in men than in women, with the incidence rate reaching 65%, which may have a great relationship with factors such as smoking and drinking (42). In addition, the proportion of patients entering stage IV reaches 75.6%, which indicates that GI cancers are characterized by late detection. Most of the patients had already metastasized by the time they sought medical treatment.

We used 8 ML methods to construct the prediction model. XGboost, with an AUC value of 0.829 and a sensitivity of 0.8667, had the best

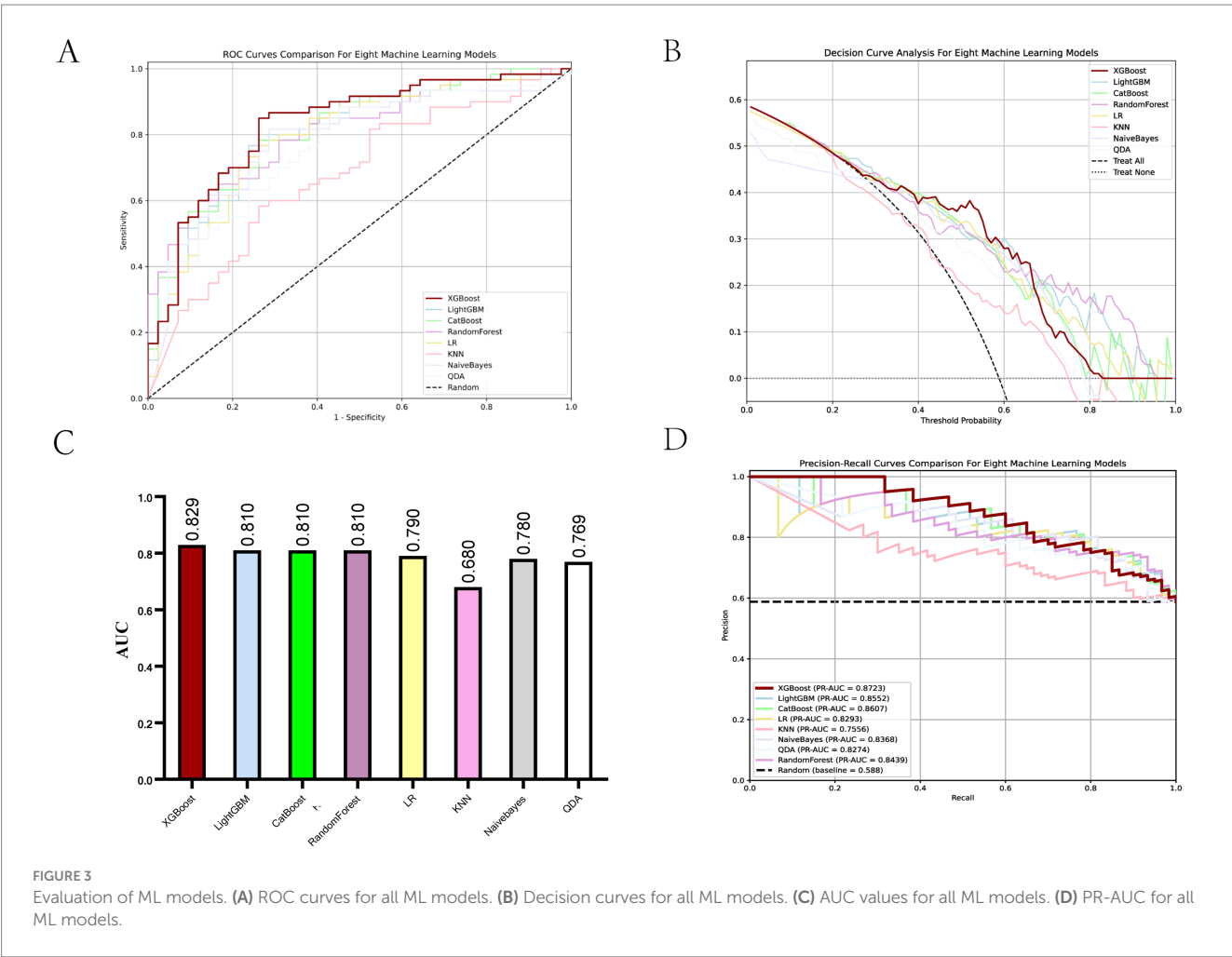


TABLE 2 Detailed parameters of each machine learning mode.

Model	Accuracy	Sensitivity	Specificity	PPV (%)	NPV (%)
XGBoost	78.43 (69.19–85.96)	86.67 (75.41–94.06)	66.67 (50.45–80.43)	78.79 (66.98–87.89)	77.78 (60.85–89.88)
LightGBM	72.55 (62.82–80.92)	86.67 (75.41–94.06)	52.38 (36.42–68.00)	72.22 (60.41–82.14)	73.33 (54.11–87.72)
CatBoost	74.51 (64.92–82.62)	88.33 (77.43–95.18)	54.76 (38.67–70.15)	73.61 (61.90–83.30)	76.67 (57.72–90.07)
RF	74.51 (64.92–82.62)	83.33 (71.48–91.71)	61.9 (45.64–76.43)	75.76 (63.64–85.46)	72.22 (54.81–85.80)
LR	74.51 (64.92–82.62)	76.67 (63.96–86.62)	71.43 (55.42–84.28)	79.31 (66.65–88.83)	68.18 (52.42–81.39)
KNN	61.76 (51.61–71.21)	68.33 (55.04–79.74)	52.38 (36.42–68.00)	67.21 (54.00–78.69)	53.66 (37.42–69.34)
Naivebayes	73.53 (63.87–81.78)	75 (62.14–85.28)	71.43 (55.42–84.28)	78.95 (66.11–88.62)	66.67 (51.05–80.00)
QDA	68.63 (58.69–77.45)	71.67 (58.56–82.55)	64.29 (48.03–78.45)	74.14 (60.96–84.74)	61.36 (45.50–75.64)

prediction performance among these models. The SHAP explanation indicates that chemotherapy is the most significant predictive feature (contribution score = 0.28), which aligns with the clinical practice of chemotherapy serving as the cornerstone of GI cancer treatment. Mechanistically, this process likely involves multiple factors. Firstly, chemotherapy enhances tumor antigen presentation and T-cell-mediated cytotoxicity, thereby potentiating immunotherapy through “sensitization” effects (43). Secondly, combination therapies significantly mitigate the risk of tumor cells developing resistance to single-treatment modalities, thereby enhancing therapeutic efficacy through synergistic effects (44). The study by Ningchen et al. investigated the association between

nutritional status and the efficacy of immune checkpoint inhibitor therapy in esophageal cancer. The research demonstrated that patients’ pretreatment HGB levels and BMI were significantly correlated with treatment effectiveness, and both served as independent prognostic indicators for survival outcomes (45). In our study, we found that a higher level of HGB and BMI significantly improved the therapeutic effect. In our predictive model, the feature importance of BMI and HGB was 0.14 and 0.15, respectively. Therefore, the patient’s baseline nutritional status positively influences the response rate to immunotherapy. In other studies, NLR is an important indicator of the degree of inflammation (25), and this was indirectly confirmed in our study. The higher the NLR ratio,

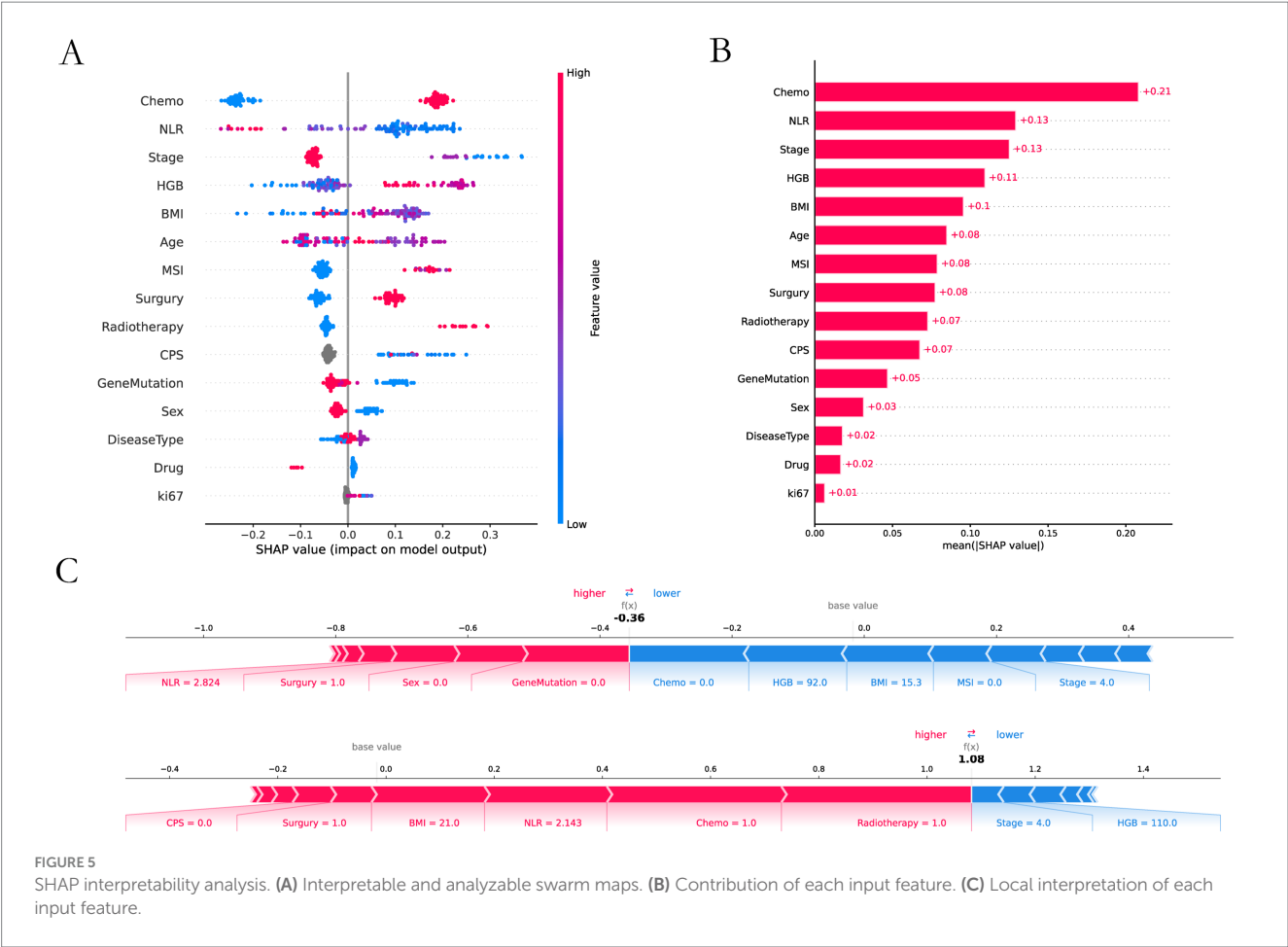
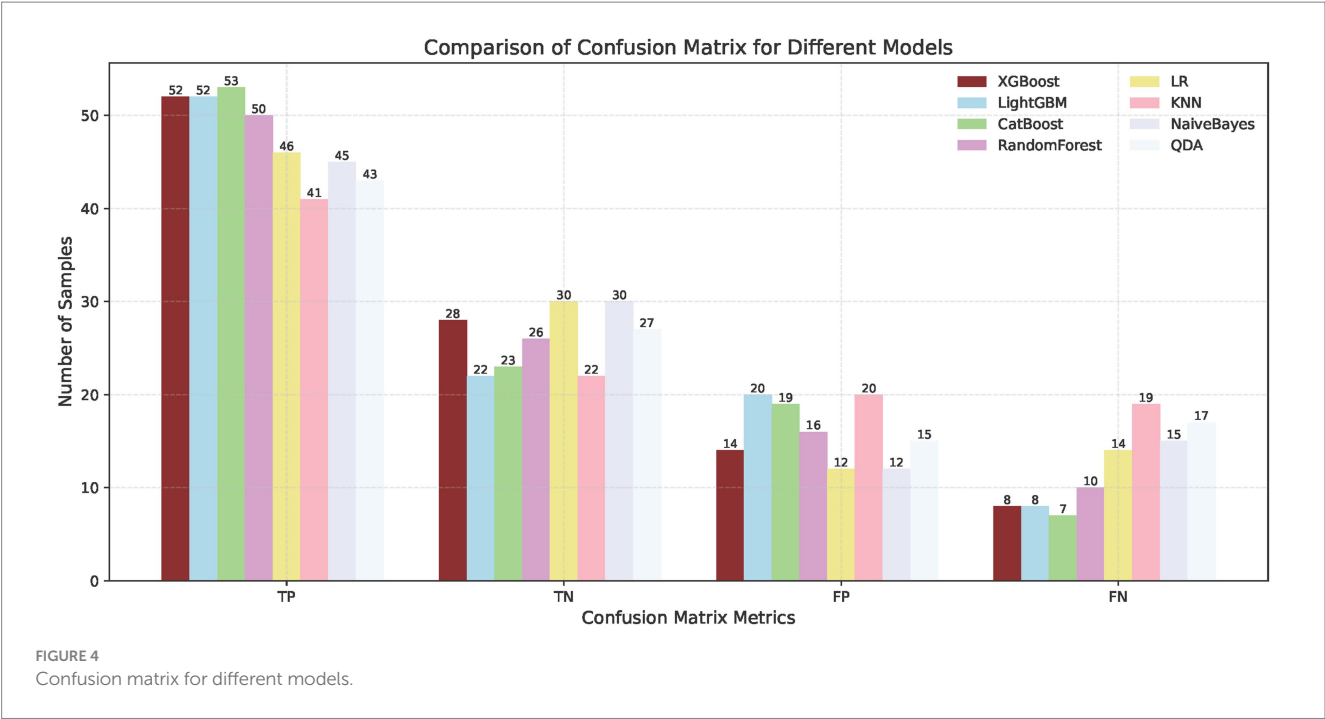


TABLE 3 Detailed analysis of important features.

Characteristic	Response, <i>n</i> (%)	No response, <i>n</i> (%)	χ^2	<i>p</i>
Treatment				
ICIs	118 (53.8)	102 (46.2)	16.96	<0.01
Chemo+ICIs	159 (72.6)	60 (27.4)		
Stage				
I–III	103 (83.7)	20 (16.7)	41.44	<0.01
IV	195 (50.90)	188 (49.10)		

the worse the outcome for the patients, which is probably related to the degree of inflammation in the patient's body. In tumor staging, once a patient enters stage IV and metastasis occurs, the response rate will be greatly reduced. Once tumor metastasis occurs, the therapeutic efficacy of immunotherapy is significantly diminished. The MSI and PD-L1 expression are very important features to measure the efficacy of immunotherapy (21, 46), but our prediction model is a combination therapy model based on immunotherapy, and the MSI and PD-L1 expression does not have absolute importance in terms of the model's contribution, and we speculate that in the combination therapy model. We speculate that in the combination therapy model, immunotherapy contribution is inherently low and assumes an adjunctive therapeutic role. Interestingly, age also plays an important role in the contribution of characteristics, and we found that the older the age, the higher the response rate, which we think may be related to the fact that young people have a fast basal metabolism, and tumors are more likely to progress and metastasize. In addition, gene mutations also contribute to treatment response rates, HER-2 positivity in GC and K-RAS mutations in CRC reduce response rates. Ki-67 is expressed in the nucleus. Once cells enter the quiescent G0 phase, Ki-67 undergoes rapid degradation, making its index value a reliable indicator of cellular proliferative activity (47). Paradoxically, while elevated Ki-67 levels correlate with accelerated tumor cell proliferation rates, this proliferation marker simultaneously demonstrates a strong positive association with chemosensitivity - tumors exhibiting high Ki-67 expression demonstrate enhanced responsiveness to chemotherapy and achieve superior treatment outcomes. This dual biological significance (pro-proliferative yet pro-chemosensitive) likely accounts for its low feature contribution rate (0.01) in our immunotherapy predictive model. In SHAP interpretability analysis, the treatment method and tumor stage are the two features with the highest contribution rates. Subsequently, we performed a deeper analysis of these two features. Table 3 shows that immunotherapy alone has a low response rate while combining immunotherapy with chemotherapy increases the response rate to 72.6%. Once the tumor reaches stage IV, the response rate drops dramatically, from 80.7 to 50.6%.

In recent studies, ML has shown significant potential in predicting the efficacy of immunotherapy. Hui Liu et al. developed a multimodal prediction model for immunotherapy of esophageal cancer, the study developed a predictive model for immunotherapy response in esophageal cancer by integrating pathology images, CT scans, and clinical data, achieving an AUC of 0.809 (48). Hong Wei Li et al. developed a predictive model for the efficacy of immunotherapy in gastric cancer, the study leveraged clinical data from 273 gastric cancer patients to construct predictive models for overall survival (OS) and progression-free survival (PFS) in response to immunotherapy, with a specific focus on patients' nutritional status. The XGBoost model achieved an AUC of 0.723 in predicting treatment outcomes (49). Current studies have focused

primarily on single cancer types rather than pan-GI malignancies. Our study addresses this gap by developing an interpretable ML framework to predict immunotherapy treatment responses across three major GI cancers: EC, GC, and CRC. Currently, clinical approaches for predicting immunotherapy responses still primarily rely on MSI status, TMB, or physicians' subjective clinical expertise. However, the tumor immune microenvironment is extremely complex, and relying solely on any single detection method cannot accurately predict immunotherapy response rates. Therefore, it is imperative to develop personalized immunotherapy strategies for patients and build predictive models for immunotherapy efficacy. Therefore, our study constructs a predictive model incorporating multiple dimensions—including common nutritional status indicators, blood biochemical markers, imaging findings, and genetic testing results. All metrics utilized are readily obtainable in routine clinical practice, enabling more effective tailoring of personalized treatment plans for individual patients.

Our study holds significant implications for clinical practice in cancer therapy. First, chemotherapy remains the cornerstone of comprehensive cancer treatment, and combination regimens can substantially enhance response rates to immunotherapy. Second, for gastrointestinal malignancies, once patients progress to stage IV, the efficacy of immunotherapy declines markedly. Hence, early screening, detection, and intervention are critically important in clinical management. Additionally, patients' systemic health status profoundly impacts immunotherapy outcomes—maintaining optimal nutritional status and controlling inflammatory responses are essential. Finally, traditional predictive biomarkers from genetic testing remain indispensable; notably, MSI status retains its irreplaceable role in forecasting immunotherapy responsiveness. In summary, the determinants of immunotherapy efficacy are multifaceted. To optimize therapeutic success, clinicians should adopt a holistic approach that integrates all relevant factors.

However, our study still has several limitations. While basic clinical characteristics including TNM staging, BMI, NLR, and HGB were assessed in 100% of patients, genetic testing was not performed in all cases. Specifically, out of a total of 506 patients with GI cancers, 381 underwent MSI testing; among 228 GC patients, 164 had HER-2 status evaluated; and among 151 CRC patients, 105 completed K-RAS testing. These missing data may have introduced bias that could potentially affect the accuracy of our predictive model. Furthermore, the lack of experimental validation remains a constraint, and additional experimental studies will be required to enhance the clinical applicability of our findings in future research. Furthermore, in our research, we split all the data into training and validation sets, but still lack an independent external validation set. To verify the accuracy of the model, we will need to use an additional independent external validation set for validation in the future.

5 Conclusion

XGBoost performed optimally with other ML methods in terms of modeling to predict response effects with clinical accuracy. Through comprehensive feature importance analysis, chemotherapy regimen and tumor staging parameters emerged as the most influential predictors, collectively accounting for 43% of the model's predictive capacity (Shapley value analysis). We will further conduct continuous tracking analysis and interpretation of the selected features to validate and apply the prediction model for the treatment effectiveness of patients with GI cancers.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by Zhongnan Hospital of Wuhan University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because Informed consent was waived due to the retrospective nature of the present study, and the data of the participants have been anonymized.

Author contributions

YL: Writing – original draft, Writing – review & editing. QW: Software, Writing – review & editing. HX: Formal analysis, Data curation, Writing – review & editing. JD: Validation, Writing – review & editing, Methodology. YW: Writing – review & editing, Funding acquisition.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The work was financially

supported by National Natural Science Foundation (22179100) and Youth Interdisciplinary Special Fund of Zhongnan Hospital of Wuhan University (ZNQNJ2023009).

Acknowledgments

The authors thank all the members of the group for the critical reading of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2025.1631011/full#supplementary-material>

References

- Morgan E, Soerjomataram I, Rungt H, Coleman HG, Thrift AP, Vignat J, et al. The global landscape of esophageal squamous cell carcinoma and esophageal adenocarcinoma incidence and mortality in 2020 and projections to 2040: new estimates from GLOBOCAN 2020. *Gastroenterology*. (2022) 163:649–658.e2. doi: 10.1053/j.gastro.2022.05.054
- Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet*. (2020) 396:635–48. doi: 10.1016/S0140-6736(20)31288-5
- Lu L, Mullins CS, Schafmayer C, Zeissig S, Linnebacher M. A global assessment of recent trends in gastrointestinal cancer and lifestyle-associated risk factors. *Cancer Commun*. (2021) 41:1137–51. doi: 10.1002/cac2.12220
- Chowell D, Yoo SK, Valero C, Pastore A, Krishna C, Lee M, et al. Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat Biotechnol*. (2022) 40:499–506. doi: 10.1038/s41587-021-01070-8
- Abnet CC, Corley DA, Freedman ND, Kamangar F. Diet and upper gastrointestinal malignancies. *Gastroenterology*. (2015) 148:1234–1243.e4. doi: 10.1053/j.gastro.2015.02.007
- Danpanichkul P, Suparan K, Tothananunroj P, Dejvajara D, Rakwong K, Pang Y, et al. Epidemiology of gastrointestinal cancers: a systematic analysis from the global burden of disease study 2021. *Gut*. (2024) 74:26–34. doi: 10.1136/gutjnl-2024-333227
- Haendchen Bento L, Kazuyoshi Minata M, Pires Batista C, Martins BDC, Lenz Tolentino LH, Scomparim RC, et al. Clinical and endoscopic aspects of metastases to the gastrointestinal tract. *Endoscopy*. (2019) 51:646–52. doi: 10.1055/a-0887-4401
- Subbiah V, Solit DB, Chan TA, Kurzrock R. The FDA approval of pembrolizumab for adult and pediatric patients with tumor mutational burden (TMB) ≥ 10 : a decision centered on empowering patients and their physicians. *Ann Oncol*. (2020) 31:1115–8. doi: 10.1016/j.annonc.2020.07.002

9. Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer*. (2016) 16:275–87. doi: 10.1038/nrc.2016.36
10. Lv Y, Luo X, Xie Z, Qiu J, Yang J, Deng Y, et al. Prospects and challenges of CAR-T cell therapy combined with ICIs. *Front Oncol*. (2024) 14:1368732. doi: 10.3389/fonc.2024.1368732
11. Luo X, Lv Y, Yang J, Long R, Qiu J, Deng Y, et al. Gamma delta T cells in cancer therapy: from tumor recognition to novel treatments. *Front Med*. (2024) 11:1480191. doi: 10.3389/fmed.2024.1480191
12. Koustas E, Trifylli EM, Sarantis P, Papadopoulos N, Karapedi E, Aloizos G, et al. Immunotherapy as a therapeutic strategy for gastrointestinal Cancer-current treatment options and future perspectives. *Int J Mol Sci*. (2022) 23:664. doi: 10.3390/ijms23126664
13. Malvicini M, Aquino JB, Mazzolini G. Combined therapy for gastrointestinal carcinomas: exploiting synergies between gene therapy and classical chemo-radiotherapy. *Curr Gene Ther*. (2015) 15:151–60. doi: 10.2174/1566523214666141224095757
14. Wang D, Lin J, Yang X, Long J, Bai Y, Yang X, et al. Combination regimens with PD-1/PD-L1 immune checkpoint inhibitors for gastrointestinal malignancies. *J Hematol Oncol*. (2019) 12:42. doi: 10.1186/s13045-019-0730-9
15. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. (2015) 348:124–8. doi: 10.1126/science.aaa1348
16. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther*. (2017) 16:2598–608. doi: 10.1158/1535-7163.MCT-17-0386
17. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet*. (2019) 51:202–6. doi: 10.1038/s41588-018-0312-8
18. Luksha M, Riaz N, Makarov V, Balachandran VP, Hellmann MD, Soloviyov A, et al. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature*. (2017) 551:517–20. doi: 10.1038/nature24473
19. Valero C, Lee M, Hoen D, Wang J, Nadeem Z, Patel N, et al. The association between tumor mutational burden and prognosis is dependent on treatment context. *Nat Genet*. (2021) 53:11–5. doi: 10.1038/s41588-020-00752-4
20. Mandal R, Samstein RM, Lee KW, Havel JJ, Wang H, Krishna C, et al. Genetic diversity of tumors with mismatch repair deficiency influences anti-PD-1 immunotherapy response. *Science*. (2019) 364:485–91. doi: 10.1126/science.aau0447
21. Holder AM, Dedeilia A, Sierra-Davidson K, Cohen S, Liu D, Parikh A, et al. Defining clinically useful biomarkers of immune checkpoint inhibitors in solid tumours. *Nat Rev Cancer*. (2024) 24:498–512. doi: 10.1038/s41568-024-00705-7
22. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. (2019) 25:44–56. doi: 10.1038/s41591-018-0300-7
23. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. (2019) 380:1347–58. doi: 10.1056/NEJMr1814259
24. Gupta D, Lis CG. Pretreatment serum albumin as a predictor of cancer survival: a systematic review of the epidemiological literature. *Nutr J*. (2010) 9:69. doi: 10.1186/1475-2891-9-69
25. Valero C, Lee M, Hoen D, Weiss K, Kelly DW, Adusumilli PS, et al. Pretreatment neutrophil-to-lymphocyte ratio and mutational burden as biomarkers of tumor response to immune checkpoint inhibitors. *Nat Commun*. (2021) 12:729. doi: 10.1038/s41467-021-20935-9
26. Conforti F, Pala L, Bagnardi V, De Pas T, Martinetti M, Viale G, et al. Cancer immunotherapy efficacy and patients' sex: a systematic review and meta-analysis. *Lancet Oncol*. (2018) 19:737–46. doi: 10.1016/S1470-2045(18)30261-4
27. Ikeguchi A, Machiorlatti M, Vesely SK. Disparity in outcomes of melanoma adjuvant immunotherapy by demographic profile. *Melanoma Manag*. (2020) 7:2–10. doi: 10.2217/mmt-2020-0002
28. Wang Z, Aguilar EG, Luna JJ, Dunai C, Khuat LT, Le CT, et al. Paradoxical effects of obesity on T cell function during tumor progression and PD-1 checkpoint blockade. *Nat Med*. (2019) 25:141–51. doi: 10.1038/s41591-018-0221-5
29. Sanchez A, Furberg H, Kuo F, Vuong L, Ged Y, Patil S, et al. Transcriptomic signatures related to the obesity paradox in patients with clear cell renal cell carcinoma: a cohort study. *Lancet Oncol*. (2020) 21:283–93. doi: 10.1016/S1470-2045(19)30797-1
30. Kuai J, Yang F, Li GJ, Fang XJ, Gao BQ. In vitro-activated tumor-specific T lymphocytes prolong the survival of patients with advanced gastric cancer: a retrospective cohort study. *Onco Targets Ther*. (2016) 9:3763–70. doi: 10.2147/OTT.S102909
31. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The eighth edition AJCC Cancer staging manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. (2017) 67:93–9. doi: 10.3322/caac.21388
32. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med*. (2017) 377:2500–1. doi: 10.1056/NEJMc1713444
33. Luchini C, Bibeau F, Ligtenberg MJL, Singh N, Nottegar A, Bosse T, et al. ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach. *Ann Oncol*. (2019) 30:1232–43. doi: 10.1093/annonc/mdz116
34. Salem ME, Bodor JN, Puccini A, Xiu J, Goldberg RM, Grothey A, et al. Relationship between MLH1, PMS2, MSH2 and MSH6 gene-specific alterations and tumor mutational burden in 1057 microsatellite instability-high solid tumors. *Int J Cancer*. (2020) 147:2948–56. doi: 10.1002/ijc.33115
35. Chen Z, Chen Y, Sun Y, Tang L, Zhang L, Hu Y, et al. Predicting gastric cancer response to anti-HER2 therapy or anti-HER2 combined immunotherapy based on multi-modal data. *Signal Transduct Target Ther*. (2024) 9:222. doi: 10.1038/s41392-024-01932-y
36. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. (2009) 45:228–47. doi: 10.1016/j.ejca.2008.10.026
37. Akiba T, Sano S, Yanase T, Ohta T, Koyama M, Optuna: A next-generation Hyperparameter optimization framework. Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining Anchorage, AK Association for Computing Machinery (2019) 2623–2631.
38. Mahardika TN, Fuadah YN, Jeong DU, Lim KM. PPG signals-based blood-pressure estimation using grid search in hyperparameter optimization of CNN-LSTM. *Diagnostics*. (2023) 13:2566. doi: 10.3390/diagnostics13152566
39. Goodswen SJ, Barratt JLN, Kennedy PJ, Kaufer A, Calarco L, Ellis JT. Machine learning and applications in microbiology. *FEMS Microbiol Rev*. (2021) 45:15. doi: 10.1093/femsre/ruab015
40. Sylvester S, Sagehorn M, Gruber T, Atzmueller M, Schone B. SHAP value-based ERP analysis (SHERPA): increasing the sensitivity of EEG signals with explainable AI methods. *Behav Res Methods*. (2024) 56:6067–81. doi: 10.3758/s13428-023-02335-7
41. 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) 17:65. doi: 10.1186/s13045-024-01578-x
42. Rumgay H, Shield K, Charvat H, Ferrari P, Sornpaisarn B, Obot I, et al. Global burden of cancer in 2020 attributable to alcohol consumption: a population-based study. *Lancet Oncol*. (2021) 22:1071–80. doi: 10.1016/S1470-2045(21)00279-5
43. Yuan SQ, Nie RC, Jin Y, Liang CC, Li YF, Jian R, et al. Perioperative toripalimab and chemotherapy in locally advanced gastric or gastro-esophageal junction cancer: a randomized phase 2 trial. *Nat Med*. (2024) 30:552–9. doi: 10.1038/s41591-023-02721-w
44. Birnboim-Perach R, Benhar I. Using combination therapy to overcome diverse challenges of immune checkpoint inhibitors treatment. *Int J Biol Sci*. (2024) 20:3911–22. doi: 10.7150/ijbs.93697
45. Chen N, Yu Y, Shen W, Xu X, Fan Y. Nutritional status as prognostic factor of advanced oesophageal cancer patients treated with immune checkpoint inhibitors. *Clin Nutr*. (2024) 43:142–53. doi: 10.1016/j.clnu.2023.11.030
46. Bao X, Zhang H, Wu W, Cheng S, Dai X, Zhu X, et al. Analysis of the molecular nature associated with microsatellite status in colon cancer identifies clinical implications for immunotherapy. *J Immunother Cancer*. (2020) 8:e001437. doi: 10.1136/jitc-2020-001437
47. Sun X, Kaufman PD. Ki-67: more than a proliferation marker. *Chromosoma*. (2018) 127:175–86. doi: 10.1007/s00412-018-0659-8
48. Liu H, Bai Y, Wang Z, Yin S, Gong C, Wang B. Multimodal deep learning for predicting PD-L1 biomarker and clinical immunotherapy outcomes of esophageal cancer. *Front Immunol*. (2025) 16:1540013. doi: 10.3389/fimmu.2025.1540013
49. Li HW, Zhu ZY, Sun YF, Yuan CY, Wang MH, Wang N, et al. Machine learning algorithms able to predict the prognosis of gastric cancer patients treated with immune checkpoint inhibitors. *World J Gastroenterol*. (2024) 30:4354–66. doi: 10.3748/wjg.v30.i40.4354