Journal of Radiology Medicine in

Original Article

Doi: 10.51271/JRM-0037

Integration of CT radiomics and clinical factors to predict pathological response to FLOT chemotherapy in gastric cancer

Görkem Özdemir[1], Çağatay Neftali Tulu[2], Ömer Işık[2], Tolga Ölmez[1], Alper Sözütek[1]

[1]Department of Gastroenterological Surgery, Adana City Training and Research Hospital, Adana, Turkiye
[2]Department of Software Engineering, Adana Alparslan Türkeş Science and Technology University, Adana, Turkiye

Received: 04.08.2025 ◆ **Accepted**: 29.09.2025 ◆ **Published**: 15.10.2025

Cite this article: Özdemir G, Tulu ÇN, Işık Ö, Ölmez T, Sözütek A. Integration of CT radiomics and clinical factors to predict pathological response to FLOT chemotherapy in gastric cancer. *J Radiol Med*. 2025;2(4):64-70.

Corresponding Author: Görkem Özdemir, dr.gorkemozdemir@yahoo.com.tr

ABSTRACT

Aims: Neoadjuvant chemotherapy (NAC) has become a standard of care for locally advanced gastric cancer (LAGC), but a significant portion of patients do not respond to treatment, leading to unnecessary toxicity and delayed definitive surgery. Non-invasive, predictive tools are urgently needed to identify these non-responders prior to treatment. This study aimed to develop and validate a predictive model using an integrated approach of clinical factors and radiomic features extracted from pretreatment computed tomography (CT) scans to classify patients with LAGC as responders or non-responders to the neoadjuvant FLOT protocol.

Methods: This retrospective study utilized 163 patients. A comprehensive feature set, including 11 patient biomarkers and 101 radiomic features, was used to train various machine learning algorithms via the pyCaret library. The Light Gradient Boosting Machine (LightGBM) was identified as the top-performing model.

Results: On the internal validation, the LightGBM model achieved an accuracy of 80% with an AUC of 0.7926. Key features contributing to the model's prediction included albumin, CA 19-9, age, and specific radiomic features such as Robust Mean Absolute Deviation. However, the model's performance decreased in the independent external validation, with an accuracy of 54.59% and an AUC of 0.5409, highlighting challenges in generalizability.

Conclusion: This study demonstrates the potential of an integrated clinical and radiomic approach using a LightGBM model to non-invasively predict treatment response to FLOT in LAGC. While the model showed strong internal performance, the decreased performance on an external validation set underscores the critical need for larger, multi-institutional, prospective studies with standardized data collection to establish clinical generalizability and widespread utility.

Keywords: Gastric cancer, neoadjuvant chemotherapy, FLOT protocol, radiomics, predictive model, machine learning

INTRODUCTION

Gastric cancer is a serious global health challenge, ranking among the leading causes of cancer-related mortality worldwide, with its incidence being particularly high in Asia and Eastern Europe.[1] Locally advanced gastric cancer (LAGC) presents a substantial clinical challenge, characterized by a heterogeneous prognosis and a significant risk of recurrence even after aggressive therapeutic interventions.[2] The standard of care for resectable LAGC has evolved to a multimodal approach that incorporates neoadjuvant chemotherapy (NAC) as a pivotal component. The primary goals of NAC include improving surgical outcomes, achieving tumor downstaging, eradicating micrometastases, and ultimately enhancing resection rates and long-term patient survival.[2-4]

Among various regimens, the FLOT (Fluorouracil, Leucovorin, Oxaliplatin, and Docetaxel) protocol has emerged as a preferred NAC regimen, consistently demonstrating superior survival outcomes compared to other approaches in randomized clinical trials.[3,4] However, a significant proportion of patients exhibit a poor or non-existent response to FLOT, rendering them susceptible to unnecessary toxicities and detrimental delays in definitive

This work is licensed under a Creative Commons Attribution 4.0 International License.

surgical intervention.[5] Identifying these non-responders before treatment initiation is of paramount importance for optimizing patient care. Such a predictive tool would enable clinicians to personalize therapeutic strategies, potentially directing non-responsive patients to alternative therapies or appropriate clinical trials, thereby optimizing outcomes and reducing treatment-related morbidity.[6,7]

Traditional methods for assessing treatment response are broadly categorized into radiological and pathological approaches. Radiological assessments, performed non-invasively, have historically relied on methods such as the response evaluation criteria in solid tumors (RECIST). However, RECIST, which primarily focuses on macroscopic tumor size changes, has demonstrated significant limitations in accurately predicting patient outcomes. For instance, it has been shown to fail in identifying a substantial portion of patients who progress to metastatic disease and cannot reliably stratify patients by long-term survival.[8,9] A more effective radiological approach involves TNM re-staging, which compares the baseline CT stage to the post-NAC pathological stage (downstaging). This method has been demonstrated to successfully stratify patients into distinct groups with different long-term survival rates and is recommended for routine clinical use.[10] Pathological evaluation of the surgical specimen provides the most definitive assessment of treatment response. Key pathological indicators include pathological complete response (pCR) and graded histological regression (GHR). Achieving pCR after NAC is a promising prognostic factor, strongly correlated with significantly improved overall survival and disease-free survival compared to non-pCR patients.[3,11] The survival benefit is particularly pronounced in patients who also achieve ypN0 status (no residual nodal disease).[5] Furthermore, a GHR of ≥50% is considered a reliable criterion for predicting a favorable prognosis and can inform postoperative chemotherapy decisions.[3] The post-NAC pathological stage (ypTNM) and Lauren classification also serve as independent prognostic factors.[3] Indeed, response to NAC is widely considered the strongest independent predictor of survival after curative resection for gastric cancer.[12]

Given the inherent limitations of conventional radiological downsizing methods and the pressing need for more precise, non-invasive predictive tools, there is growing interest in integrating advanced computational approaches. Radiomics, a promising field, extracts high-dimensional quantitative features from standard medical images such as CT scans, which can capture subtle phenotypic characteristics of tumors not discernible by the naked eye.[13,14] When combined with sophisticated machine learning algorithms, radiomics-based models have shown considerable promise in non-invasively predicting NAC response and survival, often with improved accuracy when integrated with clinical data.[14,15] Furthermore, the rise of advanced machine learning algorithms, such as the Light Gradient Boosting Machine (LightGBM), provides a powerful framework for integrating and analyzing these complex, high-dimensional, multi-modal datasets. These models can uncover intricate, non-linear relationships between a multitude of features and the final treatment outcome, surpassing the limitations of traditional statistical methods.[16-18]

Therefore, this study aims to develop and thoroughly validate a promising predictive model using the LightGBM algorithm to accurately classify LAGC patients as responders or non-responders to the neoadjuvant FLOT protocol. By comprehensively integrating patient clinical, demographic, and laboratory data with quantitative features derived from pre-treatment CT scans, we seek to provide a novel, data-driven tool. This tool is intended to improve personalized treatment decision-making, enhance risk stratification, and ultimately optimize the clinical management of LAGC.

METHODS

Ethics

The study was approved by the Adana City Training and Research Hospital Clinical Researches Ethics Committee (Date: 07.11.2024, Decision No: 220) and was conducted in accordance with the Declaration of Helsinki.

Study Design and Patient Cohorts

This retrospective analysis included a total of 163 consecutive patients with LAGC who received neoadjuvant FLOT chemotherapy followed by total or subtotal gastrectomy with D2 lymph node dissection.

The inclusion criteria for the study were as follows: a diagnosis of LAGC, the availability of a pre-treatment (baseline) contrast-enhanced CT examination, and the receipt of the neoadjuvant FLOT protocol followed by surgical resection.

Patients were excluded if they did not complete the FLOT regimen (n=12), did not undergo surgical resection (n=10), or had CT examinations with insufficient image quality (n=20).

The final study population of 163 eligible patients was divided into two distinct cohorts:

Internal cohort: A total of 126 patients were used for model training and internal validation. The training cohort consisted of 46 responders and 54 non-responders, while the internal validation test cohort consisted of 10 non-responders and 16 responders.

External cohort: A total of 37 patients were used for independent external validation. This cohort consisted of 18 responders and 19 non-responders. These patients met the inclusion criteria but their pre-treatment CT scans were performed at external imaging centers, providing a source of independent data for testing the model's generalizability.

Data Collection and Preprocessing

Baseline assessments were conducted for all patients before the initiation of neoadjuvant therapy. Data were collected on patient age, gender, tumor grade, carcinoembryonic antigen (CEA) level, CA 19-9 level, albumin and hemoglobin levels, tumor location, dysphagia, vomiting, and weight loss.

The final pathological response was determined by a pathologist's examination of the surgical specimen using the Ryan Tumor Regression Grade system (RRR). Patients were categorized into two groups for the purpose of this study: responders (partial or complete response, RRR=1) and non-responders (no response, RRR=0).

Radiomics Feature Extraction

Radiomics features were extracted from the pre-treatment contrast-enhanced CT scans. Tumor regions of interest (ROIs) were manually delineated by a single experienced gastroenterological surgeon on all relevant axial slices. It is important to note that a formal assessment of inter-observer and intra-observer variability was not conducted. To facilitate the feature extraction process, these delineated ROIs were used to create a binary mask for each image.

Prior to feature extraction, the CT images and corresponding masks were processed using the pyRadiomics open-source package (version 3.0). No image-specific preprocessing steps, such as intensity normalization or resampling, were applied.

A total of 101 radiomics features were extracted using the default settings of the pyRadiomics package. These features included:

- First-order statistics (e.g., Mean, Skewness, Kurtosis)
- Shape features
- Gray-Level Co-occurrence Matrix (GLCM) features
- Gray-Level Run-Length Matrix (GLRLM) features
- Gray-Level Size Zone Matrix (GLSZM) features
- Gray-Level Dependence Matrix (GLDM) features

These 101 radiomics features were then combined with the 11 patient-specific clinical and laboratory biomarkers, creating a comprehensive feature set of 112 variables for model training and evaluation.

Feature Selection

To mitigate the risk of overfitting due to high dimensionality, we applied a feature selection step prior to model development. A classical recursive feature elimination (RFE) strategy was used, employing LightGBM as the base estimator to iteratively remove features with the lowest importance weights. The selection process was guided by cross-validation performance, and the procedure ultimately identified 21 features (a combination of clinical biomarkers and radiomic descriptors) as the optimal subset for model training. These selected features were then used to construct the predictive models.

Predictive Model Development and Evaluation

A LightGBM model was trained on the internal cohort to classify patients as responders or non-responders. The model's performance was evaluated using standard metrics, including:

Accuracy: The proportion of correctly classified patients.

AUC: The area under the receiver operating characteristic curve.

Recall: The ability of the model to correctly identify responders.

Precision: The proportion of positive identifications that were actually correct.

F1 score: The harmonic mean of precision and recall.

Feature importance was calculated to determine the relative influence of each variable on the model's predictions. The final trained model was then applied to the independent external validation cohort to assess its generalizability.

RESULTS

Performance of the Predictive Models

After training, various machine learning models were evaluated on a dedicated test dataset. The LightGBM model demonstrated the highest overall performance, achieving an accuracy of 0.80 and an AUC of 0.7926. The model's recall was 0.825, its precision was 0.8461, and its F1-score was 0.8354. Performance metrics for all models are summarized in **Table 1**.

Table 1. Performance metrics of the models on the test dataset after training

Model	Accuracy	AUC	Recall	Prec.	F1
Light gradient boosting machine	0.8	0.7926	0.825	0.8461	0.8354
Gradient boosting classifier	0.7538	0.7852	0.7875	0.8077	0.7975
Ada boost classifier	0.7308	0.736	0.7	0.8358	0.7619
Extreme gradient boosting	0.6538	0.7492	0.6625	0.7465	0.702
Extra trees classifier	0.5692	0.6652	0.5125	0.7069	0.5942
Decision tree classifier	0.5462	0.5562	0.5125	0.6721	0.5816
SVM-Linear Kernel	0.6308	0.5875	0.775	0.6739	0.7209
Logistic regression	0.6692	0.6812	0.775	0.7126	0.7425
K neighbors classifier	0.5692	0.6347	0.4375	0.7609	0.5556
Linear discriminant analysis	0.6462	0.6842	0.7	0.7179	0.7089
Quadratic discriminant analysis	0.6615	0.662	0.8	0.6957	0.7442
Dummy classifier	0.3846	0.5	0	0	0
Naive bayes	0.6615	0.7547	0.9	0.6667	0.766

AUC: Area under the curve

Feature Importance Analysis

Following training with the LightGBM method, a feature importance analysis was conducted to identify the most influential variables in the model's decision-making process. The top five features, ranked by their importance scores, were albumin, CA 19-9, age, hemoglobin, and CEA. These are often clinically relevant markers, which add credibility to the model's findings. Other features, including weight loss, grade, and several image diagnostics features, also showed some level of influence. The detailed feature importance scores are presented in **Table 2**.

External Validation Results

The pre-trained LightGBM model was then tested on the independent external validation dataset (n=37). The model achieved an accuracy of 0.5459, an AUC of 0.5409, a recall of 0.5312, and a precision of 0.5666. The F1-score was 0.5483.

Table 2. Feature importances for the LightGBM model after training

Feature	Importance score
Albumin	0.1484
CA 19-9	0.1391
Age	0.1251
Hemoglobin	0.1158
CEA	0.1033
Weight loss	0.0420
Grade	0.0420
Robust mean absolute deviation (first-order)	0.0404
Mean intensity (image diagnostics)	0.0311
Zone variance (GLSZM)	0.0256
Skewness (first-order)	0.0241
Gray level non-uniformity normalized (GLRLM)	0.0218
Dependence non-uniformity normalized (GLDM)	0.0202
Gray level non-uniformity (GLRLM)	0.0186
Gray level non-uniformity (GLSZM)	0.0186
Tumor location	0.0186
Dependence variance (GLDM)	0.0171
Minimum intensity (first-order)	0.0163
Gender	0.0140
Kurtosis (first-order)	0.0093
Energy (first-order)	0.0085

LightGBM: Light Gradient Boosting Machine, CEA: Carcinoembryonic antigen

DISCUSSION

The accurate prediction of pathological response to NAC is a critical step towards personalized medicine in LAGC. Our study aimed to address this challenge by developing a machine learning model, specifically a LightGBM classifier, which integrates both clinical and quantitative radiomic features from pre-treatment CT scans to predict treatment response to the neoadjuvant FLOT regimen. The model demonstrated promising predictive performance in internal validation, suggesting potential value of a multi-modal approach, though this was not replicated externally in identifying patients most likely to benefit from NAC.The prognostic significance of pathological response, particularly pCR and good histologic response, is well-established in gastric cancer literature, correlating strongly with improved long-term outcomes.[19] However, conventional radiological response assessments, such as RECIST, often show poor consistency with pathological findings.[20,21] Furthermore, their predictive value for overall survival can be limited.[21] Our study bypasses this limitation by directly predicting pathological response, which is a more reliable prognostic marker. By identifying patients who will not respond to treatment preoperatively, our model could potentially guide clinicians to alternative strategies, such as upfront surgery or different neoadjuvant regimens, thereby avoiding unnecessary treatment toxicity and delaying definitive surgery.

The integration of radiomic features into our model proved to be highly effective. Radiomics allows for the extraction of high-dimensional, quantitative data from medical images that are imperceptible to the human eye, providing a "non-invasive biopsy" of the tumor's phenotype and microenvironment.[22] The fact that features such as mean intensity and zone variance were identified as important by our model highlights their potential as imaging biomarkers. Several studies have similarly demonstrated the power of radiomics in predicting treatment response and survival in gastric cancer, often outperforming traditional methods.[20,23] Moreover, our use of LightGBM, a powerful and efficient gradient-boosting framework, is also noteworthy. It is well-suited for handling complex, high-dimensional data and has shown superior performance compared to other machine learning algorithms in similar predictive tasks.[24]

Our findings are consistent with previous reports suggesting that multi-modal, AI-driven approaches may have potential to may enhance precision medicine in oncology, although our failed external validation emphasizes that their clinical utility remains unproven.

For instance, recent studies have shown that machine learning models combining clinical data, radiomics, and even genomic biomarkers can accurately predict pathological response to both chemotherapy and immunotherapy in gastric cancer.[24] This multi-omics approach represents a promising future direction for a more comprehensive understanding of tumor biology and patient-specific response. While our study focused on a single institution and a specific treatment regimen, the methodology provides a strong foundation for future research, including prospective, multicenter trials needed to validate the model's clinical utility and establish its role in routine clinical practice.

A key methodological consideration in our study is the manual delineation of the tumor regions, which was performed by a single gastroenterological surgeon. The rationale for this approach was to develop a practical predictive tool specifically tailored for the primary treating physician in high-volume clinical settings, where the opportunity for extensive multidisciplinary council discussions on every patient may be limited. While this method benefits from expert clinical knowledge, it inherently introduces the potential for inter-observer and intra-observer variability. We acknowledge that a formal assessment of this variability was not conducted, which represents a limitation on the reproducibility and generalizability of our findings. This potential for inconsistency in radiomics feature extraction may have contributed to the performance discrepancy between our internal and external cohorts. This limitation highlights a critical area for future research, where the implementation of standardized, multi-reader delineation protocols or the development of robust automated segmentation algorithms could significantly improve reproducibility. For the present study, the manual delineation represents a potential source of bias, reinforcing the need for further validation with data from diverse institutions and standardized imaging workflows.

The notable decrease in predictive performance observed during the independent external validation, as evidenced by the drop in accuracy from 0.80 to 0.5459 and AUC from 0.7926 to 0.5409, represents a critical challenge for the broader applicability of our model. This significant drop in predictive performance is a well-recognized issue in machine learning and is often attributed to several factors. The most

likely reason is that the external dataset of 37 patients is fundamentally different from the training cohort.

This could be due to differences in CT image acquisition protocols. Moreover, demographics, disease stage distribution, or comorbidities may affect the results. Radiomic features are highly sensitive to variations in image acquisition. Even subtle differences in the CT scanner manufacturer, model, slice thickness, reconstruction algorithm, or contrast injection protocol between the original institution and the source of the external dataset could lead to different feature values. This "batch effect" can severely degrade a model's performance. Our model, trained on data from a single institution, may have learned patterns specific to that population and failed to generalize to the new, unseen cohort. Furthermore, the model may have overfit to the training data. While it performed well on the internal validation set, it might have learned noise or specific characteristics of the training data rather than true underlying biological signals. The external dataset then exposes this weakness, revealing that the model is not robust enough for a different population. Additionally, with only 37 patients in the external validation set, the performance metrics (accuracy, AUC, etc.) are highly susceptible to individual patient outcomes. This small sample size makes it difficult to draw a definitive conclusion about the model's true performance on a larger, more representative external population. The reported values might not be a stable or reliable measure of the model's capability.

A critical consideration in evaluating the generalizability of our predictive model is the relatively small size of the independent external validation cohort. While the inclusion of an external validation set is a significant strength of this study, a cohort of this size is inherently susceptible to individual patient variations, which can disproportionately influence performance metrics such as accuracy and AUC. This limited sample size complicates drawing definitive conclusions about the model's true generalizability and robust performance across diverse populations. Consequently, these external validation results, while indicative, underscore the crucial need for further validation in significantly larger, multi-institutional, and prospectively collected cohorts to establish the model's widespread clinical utility and broader applicability.

While acknowledged contributing factors include inherent dataset shift, variations in clinical practices, and potential overfitting to the internal cohort, this finding underscores the paramount importance of data harmonization. Moving forward, standardization of imaging protocols (e.g., scanner type, acquisition parameters, contrast injection timing) and detailed clinical definitions across different institutions will be crucial to mitigate these discrepancies. Furthermore, this finding strongly reinforces the necessity for larger, multi-institutional, prospective studies, which are essential to thoroughly validate the model's generalizability, minimize the impact of small sample size variability, and establish its consistent clinical utility across diverse patient populations and healthcare settings. Such collaborative efforts will be fundamental in transitioning this promising predictive tool into routine clinical practice. While our model demonstrated robust performance within the internal cohort, its failure to replicate this success on the external dataset can be attributed to several factors. These likely include subtle but impactful differences in CT imaging protocols across institutions, variations in clinical practices, and inherent dataset shift.

The feature importance analysis revealed that clinical variables such as albumin, CA 19-9, age, hemoglobin, and CEA were consistently ranked higher than most radiomic features. This pattern suggests that the predictive power of the model may be primarily driven by easily accessible clinical markers, with radiomics contributing minimally. While this does not reduce the potential role of radiomics, it highlights the need to systematically evaluate whether their integration provides significant incremental value over clinical features alone. Future work should include direct comparisons between models based solely on clinical data and those that integrate radiomic features, to determine the true added value of radiomics in predictive modeling for gastric cancer.

Limitations

This study has several key limitations. Its retrospective, single-center design may introduce selection bias and limit generalizability, and the relatively small sample size,

increases the risk of overfitting. Manual segmentation by a single observer also represents a potential source of bias, as radiomics features are sensitive to variations in both delineation and imaging protocols. Future multicenter studies with larger cohorts, standardized imaging, and automated or multi-reader segmentation approaches are needed to improve reproducibility and validation.

Another limitation of our study is the absence of a formal intra-observer and inter-observer reliability analysis of manual tumor segmentation. The manual delineations were performed by a single experienced gastroenterological surgeon, which provided clinical expertise but potentially introduces the possibility of intra-observer variability. Since repeated segmentations by the same observer or a radiologist were not available, an intraclass correlation coefficient (ICC) analysis could not be performed. This methodological limitation underscores the need for future research employing standardized multi-reader protocols or automated segmentation tools, which could enhance reproducibility and minimize observer-related variability in radiomics studies.

Another limitation of our study is the absence of image preprocessing or harmonization across datasets. The external validation cohort included scans from multiple centers with varying scanners, acquisition parameters (slice thickness, kVp, mAs), and contrast protocols, but detailed parameter-level information was not available for all patients. This heterogeneity likely contributed to the performance decline in external validation. While such variability reflects real-world clinical conditions, future studies should adopt standardized acquisition protocols or apply harmonization techniques (e.g., ComBat, intensity normalization) to improve reproducibility and generalizability of radiomics-based models.

A further methodological challenge concerns the generalizability of the selected features. Although we applied recursive feature elimination (RFE) with a LightGBM estimator and reduced the initial 112 features to 21, the external validation still showed a notable drop in accuracy

and AUC. This indicates that while feature selection reduced overfitting within the internal cohort, the chosen features may not have been sufficiently reliable across different institutions and imaging protocols. Future studies with larger, multi-institutional datasets and standardized imaging workflows are needed to validate the stability of selected features and improve reproducibility.

Furthermore, the predictive model is specific to the neoadjuvant FLOT regimen, and its applicability to other treatment protocols is not established. Lastly, while the LightGBM model is effective, its "black box" nature offers limited biological insight, suggesting a need for future studies that integrate radiomic findings with molecular data and confirm clinical utility through prospective trials.

CONCLUSION

This study explored the development of a LightGBM-based predictive model that showed potential in internal validation, but failed to generalize to an external cohort in classifying LAGC patients as responders or non-responders to neoadjuvant FLOT chemotherapy. These findings highlight both the opportunities and challenges of radiomics-based predictive modeling, and underscore the necessity of larger multicenter prospective validation before clinical application.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was approved by the Adana City Training and Research Hospital Clinical Researches Ethics Committee (Date: 07.11.2024, Decision No: 220).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660
2. Ajani JA, D'Amico TA, Bentrem DJ, et al. Gastric cancer, version 2.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2022; 20(2):167-192. doi:10.6004/jnccn.2022.0008
3. Al-Batran SE, Homann N, Pauligk C, et al. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. *Lancet*. 2019; 393(10184):1948-1957. doi:10.1016/S0140-6736(18)32557-1
4. van Laarhoven H. Management of locally advanced gastric cancer. *Lancet Oncol*. 2025;26(3):267-269. doi:10.1016/S1470-2045(24)00727-7
5. Paszt A, Simonka Z, Budai K, et al. Impact of neoadjuvant FLOT treatment of advanced gastric and gastroesophageal junction cancer following surgical therapy. *Front Surg*. 2023;10:1148984. doi:10.3389/fsurg.2023.1148984
6. Sun F, Gao X, Wang W, Zhao X, Zhang J, Zhu Y. Predictive biomarkers in the era of immunotherapy for gastric cancer: current achievements and future perspectives. *Front Immunol*. 2025;16:1599908. doi:10.3389/fimmu.2025.1599908
7. Zhou H, Gu Q, Bao R, et al. Machine learning based models for predicting presentation delay risk among gastric cancer patients. *Front Oncol*. 2025;14:1503047. doi:10.3389/fonc.2024.1503047
8. Zubarayev M, Min EK, Son T. Clinical and molecular prognostic markers of survival after surgery for gastric cancer: tumor-node-metastasis staging system and beyond. *Transl Gastroenterol Hepatol*. 2019;4:59. doi:10.21037/tgh.2019.08.05
9. Kuhl CK, Alparslan Y, Schmoee J, et al. Validity of RECIST version 1.1 for response assessment in metastatic cancer: a prospective, multireader study. *Radiology*. 2019;290(2):349-356. doi:10.1148/radiol.2018180648
10. Chan AK, Wong A, Jenken D, Heine J, Buie D, Johnson D. Posttreatment TNM staging is a prognostic indicator of survival and recurrence in tethered or fixed rectal carcinoma after preoperative chemotherapy and radiotherapy. *Int J Radiat Oncol Biol Phys*. 2005;61(3):665-677. doi:10.1016/j.ijrobp.2004.06.206
11. Lowy AM, Mansfield PF, Leach SD, Pazdur R, Dumas P, Ajani JA. Response to neoadjuvant chemotherapy best predicts survival after curative resection of gastric cancer. *Ann Surg*. 1999;229(3):303-308. doi:10.1097/00000658-199903000-00001
12. Chen YH, Xiao J, Chen XJ, et al. Nomogram for predicting pathological complete response to neoadjuvant chemotherapy in patients with advanced gastric cancer. *World J Gastroenterol*. 2020;26(19):2427-2439. doi:10.3748/wjg.v26.i19.2427
13. Sun KY, Hu HT, Chen SL, et al. CT-based radiomics scores predict response to neoadjuvant chemotherapy and survival in patients with gastric cancer. *BMC Cancer*. 2020;20(1):468. doi:10.1186/s12885-020-06970-7
14. Adili D, Mohetaer A, Zhang W. Diagnostic accuracy of radiomics-based machine learning for neoadjuvant chemotherapy response and survival prediction in gastric cancer patients: a systematic review and meta-analysis. *Eur J Radiol*. 2024;173:111249. doi:10.1016/j.ejrad.2023.111249
15. Song R, Cui Y, Ren J, et al. CT-based radiomics analysis in the prediction of response to neoadjuvant chemotherapy in locally advanced gastric cancer: a dual-center study. *Radiother Oncol*. 2022;171: 155-163. doi:10.1016/j.radonc.2022.04.023
16. Wang HN, An JH, Wang FQ, Hu WQ, Zong L. Predicting gastric cancer survival using machine learning: a systematic review. *World J Gastrointest Oncol*. 2025;17(5):103804. doi:10.4251/wjgo.v17.i5.103804
17. Akcay M, Etiz D, Celik O. Prediction of survival and recurrence patterns by machine learning in gastric cancer cases undergoing radiation therapy and chemotherapy. *Adv Radiat Oncol*. 2020;5(6):1179-1187. doi:10.1016/j.adro.2020.07.007
18. Wang W, Sheng R, Liao S, et al. LightGBM is an effective predictive model for postoperative complications in gastric cancer: a study integrating radiomics with ensemble learning. *J Imaging Inform Med*. 2024;37(6):3034-3048. doi:10.1007/s10278-024-01172-0
19. Wang S, Zheng G, Cheng B, et al. Circulating tumor cells (CTCs) detected by RT-PCR and its prognostic role in gastric cancer: a meta-analysis of published literature. *PLoS One*. 2014;9(6):e99259. doi:10.1371/journal.pone.0099259
20. Sun C, Niu P, Zhang X, et al. Concurrent clinical and pathological response predicts favorable prognosis of patients with gastric cancer after neoadjuvant therapy: a real-world study. *BMC Cancer*. 2023;23(1):996. doi:10.1186/s12885-023-11508-8
21. D'Alessio A, Lombardi P, Stefanini B, et al. Use of pathological response versus RECIST assessment in predicting relapse-free survival following neoadjuvant immunotherapy for hepatocellular carcinoma. Meeting Abstract: 2025 ASCO Gastrointestinal Cancers Symposium. *J Clin Oncol*. 2025;43(Suppl 4). doi:10.1200/JCO.2025.43.4_suppl.633

22. Wang Y, Jin ZY. Radiomics approaches in gastric cancer: a frontier in clinical decision making. *Chin Med J (Engl)*. 2019;132(16):1983-1989. doi:10.1097/CM9.0000000000000360
23. Yang L, Chu W, Li M, et al. Radiomics in gastric cancer: first clinical investigation to predict lymph vascular invasion and survival outcome using 18F-FDG PET/CT images. *Front Oncol*. 2022;12:836098. doi:10.3389/fonc.2022.836098
24. Huang ZN, Zhang HX, Sun YQ, et al. Multi-cohort study in gastric cancer to develop CT-based radiomic models to predict pathological response to neoadjuvant immunotherapy. *J Transl Med*. 2025;23(1):362. doi:10.1186/s12967-025-06363-z