

89145\_Auto\_Edited.docx

**Name of Journal:** *World Journal of Gastrointestinal Surgery*

**Manuscript NO:** 89145

**Manuscript Type:** ORIGINAL ARTICLE

*Retrospective Cohort Study*

**Machine learning-based radiomics score improves prognostic prediction accuracy of stage II/III gastric cancer: A multi-cohort study**

Xiang YH *et al.* Novel biomarker for gastric cancer

Ying-Hao Xiang, Huan Mou, Bo Qu, Hui-Rong Sun

## **Abstract**

### **BACKGROUND**

Although accurately evaluating the overall survival (OS) of gastric cancer patients remains difficult, radiomics is considered an important option for studying prognosis.

### **AIM**

To develop a robust and unbiased biomarker for predicting OS using machine learning and computed tomography (CT) image radiomics.

### **METHODS**

This study included 181 stage II/III gastric cancer patients, 141 from Lichuan People's Hospital, and 40 from the Cancer Imaging Archive (TCIA). Primary tumors in the preoperative unenhanced CT images were outlined as regions of interest (ROI), and approximately 1700 radiomics features were extracted from each ROI. The skeletal muscle index (SMI) and skeletal muscle density (SMD) were measured using CT images from the lower margin of the third lumbar vertebra. Using the least absolute shrinkage and selection operator regression with 5-fold cross-validation, 36 radiomics features were identified as important predictors, and OS-associated CT image radiomics score (OACRS) was calculated for each patient using these important predictors.

### **RESULTS**

Patients with a high OACRS had a poorer prognosis than those with a low OACRS score ( $P < 0.05$ ) and those in the TCIA cohort. Univariate and multivariate analyses revealed that OACRS was a risk factor [RR = 3.023 (1.896-4.365),  $P < 0.001$ ] independent of SMI, SMD, and pathological features. Moreover, OACRS outperformed SMI and SMD and could improve OS prediction ( $P < 0.05$ ).

### **CONCLUSION**

A novel biomarker based on machine learning and radiomics was developed that exhibited exceptional OS discrimination potential.

**Key Words:** Radiomics; Machine learning; Gastric cancer; Skeletal muscle density; Skeletal muscle index

Xiang YH, Mou H, Qu B, Sun HR. Machine learning-based radiomics score improves prognostic prediction accuracy of stage II/III gastric cancer: A multi-cohort study. *World J Gastrointest Surg* 2024; In press

**Core Tip:** We investigated 141 patients with locally advanced gastric cancer and developed overall survival-associated computed tomography image radiomics score (OACRS) using machine learning and radiomics. The data revealed that OACRS was associated with overall survival (OS), as well as in the Cancer Imaging Archive cohort. Importantly, OACRS could improve the predicted accuracy of OS.

## **INTRODUCTION**

Gastric cancer is the fourth most common malignancy, and the application of multidisciplinary approaches in recent years has significantly improved its prognosis. However, the 5-year overall survival (OS) rate for locally advanced gastric cancer is less than 60%<sup>[1,2]</sup>. Tumor, node, and metastasis (TNM) staging is the cornerstone for guiding OS; however, the outcomes of gastric cancer patients who undergo radical resection with the same TNM stage can vary significantly<sup>[3]</sup>. Current research indicates that the prognosis of malignant tumors not only depends on immutable tumor-specific factors, such as histology and pathology but is also closely related to postoperative adjuvant therapy and the patient's own nutritional and physical conditions<sup>[4]</sup>. However, predicting surgical outcomes for locally advanced gastric cancer remains challenging, making developing new biomarkers related to its prognosis critical.

Previous studies have revealed an association between preoperative computed tomography (CT) image data and the prognosis of malignant tumors<sup>[5]</sup>, such as the skeletal muscle index (SMI) and skeletal muscle density (SMD)<sup>[6]</sup>. However, accurately evaluating OS remains challenging, as evidenced by the fact that approximately 50% of patients with locally advanced gastric cancer undergoing curative resection develop distant metastases in subsequent years. Radiomic progression has provided methods for extracting and analyzing thousands of image-related features to aid in diagnosis and treatment<sup>[7,8]</sup>. With applications beyond clinical decision-making, such as predicting peritoneal recurrence, disease-free survival, non-invasive tumor microenvironment evaluation, and treatment response, radiomics has shown great potential in personalized medicine, as it can improve OS prediction<sup>[9-11]</sup>. Recent studies have revealed that machine learning algorithms are highly flexible and powerful tools for modeling<sup>[12,13]</sup>. Radiomics provides an opportunity for machine learning because it is more suitable for data with multiple variables. However, the efficacy of a machine learning algorithm model using radiomics for evaluating OS and determining whether it outperforms manual indicators such as SMI and SMD in stage II/III gastric cancer remains unclear.

This study hypothesized that a machine learning model-based CT-derived radiomics could be a predictive biomarker to assess OS for stage II/III gastric cancer and outperform SMI and SMD.

## **MATERIALS AND METHODS**

### ***Patients and variables***

Clinical data from 186 patients who underwent radical resection for gastric cancer at Lichuan People's Hospital between 2013 and 2019 were collected, and all data were accessed between June 2023 and September 2023. The inclusion criteria for this study were as follows: Age of 18-80 years and primary stage II/III gastric adenocarcinoma. Forty-five patients were excluded due to tumor perforation and acute bleeding ( $n = 2$ ), R1 or R2 resection ( $n = 8$ ), missing preoperative CT images ( $n = 6$ ), number of harvested

lymph nodes < 12 ( $n = 10$ ), and missing follow up ( $n = 19$ ), and 141 patients were included. This study was approved by the medical ethics committee of Lichuan People's Hospital, and the authors had access to information that could identify individual participants during or after data collection. The clinical variables included in this study were as follows: gender, age, height, American Association of Anesthesiologists score (ASA), preoperative carcinoembryonic antigen (CEA), type of gastrectomy, tumor differentiation, tumor size, T, N, and TNM stages, nerve or vascular invasion, tumor deposition (TD), and postoperative chemotherapy. All patients were restaged according to the 8<sup>th</sup> AJCC staging system. Gastric cancer patients were followed up every three months for the first two years, every six months for postoperative 3-5 years, and every 12 months after five years at our center. The primary outcome of the current study was OS, defined as the time at which a patient died postoperatively from any cause. Another cohort of 40 patients was included in the Cancer Imaging Archive (TCIA cohort; <https://www.cancerimagingarchive.net/>).

#### ***Measurement of SMD and SMI***

Martin *et al*<sup>[14]</sup> proposed a classical method for calculating SMD. All CT examinations were performed within one week before surgery, and unenhanced images from the lower margin of the third lumbar vertebra were analyzed. First, soft tissue was visualized using Hounsfield Units (HU) [-150, 180]. Next, the skeletal muscle tissue area (SMA) was delineated using 29 to 150 HU, and the average value of HU was defined as SMD<sup>[14]</sup>. Moreover, SMI was calculated using  $SMA/height^{2[15,16]}$ . In this study, SMD and SMI were calculated using the Slice-O-Matic software (Tomovision, Montreal, Canada, version 5.0). A schematic diagram for calculating SMD and SMI is revealed in Figure 1.

#### ***Radiomics***

CT image radiomics was extracted using 3D-Slicer (5.4.0), a widely used freeware for medical image data. To evaluate gastric cancer, the 'Segmentation' module was used to delineate the primary tumor as a region of interest (ROI) from unenhanced CT images

by two surgeons with at least eight years of clinical experience (Sun HR and Qu B). Next, 'Radiomics' module was used to calculate the CT image radiomics of ROIs, such as 'shape', 'first-order', 'glcm', 'gldm', 'glrlm', 'glszm' and 'ngtdm' with their derived features. Finally, approximately 1700 CT image radiomics features were extracted. Figure 2 depicts the radiomics extraction procedure.

### *Feature selection and machine learning model*

Excessive features could increase the model complexity and make clinical applications inconvenient. A machine learning algorithm, the least absolute shrinkage and selection operator (LASSO) regression, is widely used for feature selection and modeling, which can filter unimportant features and discriminate significantly important features for predictions. Cox LASSO regression with 5-fold cross-validation was used to select the most useful predictors associated with OS to select radiomics features for modeling. The radiomics score calculated using the Cox LASSO regression model was defined as the OS-associated CT radiomics score (OACRS) and was used for subsequent machine learning modeling. Moreover, radiomics of the patients from TCIA cohort were extracted, and OACRSs were calculated using the 36 predictors. The predictive model was developed using random survival forest (RSF), a popular machine learning algorithm. Patients from our center were used as the training cohort and were classified as the discovery cohort. The high-performance model was trained using a 5-fold cross-validation and hyperparameter-adjustment approach.

### *Statistical analysis*

Categorical variables are represented as the number of cases (percentage), while continuous variables are represented as mean  $\pm$  SD. The chi-squared test was used to compare categorical variables, whereas the t-test was used for continuous variables. The 'Survminer', 'glmnet', and 'randomForstSRC' packages in R (version 0.4.9) were used to determine the best cut-off value, select features, and develop a predictive model. Independent risk factors associated with OS were analyzed using univariate and

multivariate proportional risk regression models (Cox). The Kaplan-Meier method and log-rank test were used to compare survival rates. All p-values represented the correlation by a two-tailed test, and  $P < 0.05$  indicated statistically significant differences. All analyses were performed using IBM SPSS 25.0 (SPSS for Windows, IBM Corporation, Armonk, NY, United States) or R (4.3.0).

## **RESULTS**

### ***Radiomics feature selection***

OS-related features were selected from the 1689 radiomics features initially extracted from CT images using LASSO regression with 5-fold cross-validation. A coefficient profile plot was produced according to the  $\log(\lambda)$  sequence  $(-10, 0)$  (Figure 3A). The binomial deviance curve indicated that when the  $\log(\lambda)$  value was  $4.33 \times 10^{-2}$ , there was a minimum mean-squared error, and 36 features were selected as important (Figure 3B).

### ***Patients***

In this study, 141<sup>4</sup> patients were included in the discovery cohort, and 40 were included in the TCIA cohort. To elucidate the relationship between radiomics features and OS, the OACRS of all patients was calculated using the LASSO regression model. The mean values of OACRS, SMI, and SMD were 0.48, 46.8, and 33.2 for discovery cohort patients and 0.53, 44.6, and 32.6 for TCIA cohort patients ( $P < 0.05$ ), respectively. Compared with the discovery cohort, males and elders were predominant in the TCIA cohort ( $P < 0.05$ ). Patients in the TCIA cohort exhibited later N and TNM stages ( $P < 0.05$ ). This may be due to the early screening for cancer in recent years (Table 1). In addition, patients in the discovery cohort received less proximal gastrectomy (12.8% vs 42.5%), which may be linked to the severe decline in the quality of life for proximal gastric surgery.

### ***OACRS was a biomarker for OS***

TNM staging is widely recognized as a useful prognostic indicator. Patients with stage II had a better OS than that of stage III in the discovery cohort ( $P < 0.001$ ; Figure 4A) but



not in the TCIA cohort ( $P = 0.56$ ) (Figure 4B), possibly due to insufficient number of cases. The receiver operating characteristic curve (ROC) is the most popular method for calculating the optimal cut-off value. However, ROC is applicable only to binary diagnostic tests and not to survival analysis. Consequently, ROC was unsuitable for this study. The 'surv\_cutpoint' function, developed based on the Log-Rank test, was used to calculate the optimal cut-off value of OACRS. A statistically significant difference was observed when OACRS was 0.54 (Figure 4C). Patients with high OACRS had poorer OS than those with low OACRS in the discovery and TCIA cohorts (Figure 4D and E), indicating that OACRS was associated with OS.

#### ***OACRS was independently linked to OS***

Collinearity may occur because SMI, SMD, and OACRS were calculated from unenhanced CT images. Consequently, univariate and multivariate Cox regression were used to analyze the role of OACRS in OS. Univariate analysis revealed that SMI, SMD, OACRS, CEA, nerve or vascular invasion, TD, N stage, TNM stage, and postoperative chemotherapy were significantly associated with OS ( $P < 0.05$ ). Furthermore, multivariate analysis indicated that SMI, SMD, CEA, OACRS, TNM stage, and postoperative chemotherapy were independent risk factors linked to OS ( $P < 0.05$ ; Table 2). These results indicate that OACRS is an OS predictor independent of SMI, SMD, and pathological features.

#### ***OACRS outperformed SMI and SMD for OS prediction***

Previous studies have demonstrated that SMI and SMD are independent risk factors for OS in gastric cancer, consistent with the present data. Moreover, the current study revealed that OACRS is an independent risk factor for OS in stage II/III gastric cancer, which has rarely been reported. OACRS might be a more robust predictor ( $P < 0.001$ ) than SMI ( $P = 0.009$ ) and SMD ( $P = 0.002$ ). To compare the OS prediction performance of SMI, SMD, and OACRS, the time-dependent area under the curve (TAUC) was plotted, a novel method that can calculate the area under ROC of multiple time points.

The TAUCs of SMI, SMD, and OACRS indicated that OACRS could predict OS more accurately ( $^aP < 0.05$ ) in the discovery and TCIA cohorts (Figure 5). These findings revealed that OACRS outperformed SMI and SMD for OS prediction.

#### *Development of machine learning model based on OACRS*

RSF, a prevalent, flexible, and capable algorithm, was applied to develop a prediction model using a discovery cohort to meet clinical practice requirements. Two RSF models were developed to determine whether OACRS could improve predictive accuracy: one with and one without OACRS. The AUCs of the RSF model demonstrating good discrimination of OS are displayed in Figure 6A. The model with OACRS outperformed that without OACRS ( $^aP < 0.05$ ). These findings revealed that OACRS could improve the predictive accuracy of OS. Furthermore, the C-index values for 3- and 5-year OS were 0.835 and 0.806, respectively, indicating favorable performance. The calibration curves of the 3- and 5-year OS also demonstrated good concordance between the predictions and the ground truth (Figure 6B and C). In addition, decision curve analysis indicated that the model, including OACRS, outperformed the models, including SMI or SMD (Supplementary Figure 1), suggesting that OACRS is a useful biomarker for OS prediction. Although the RSF model performed favorably, the precise contributions of each variable to the predictions remain unclear. To determine the contribution of each variable, the importance of each feature was calculated and digitized (Figure 6D). The results revealed that the five most important variables were tumor stage, OACRS, N stage, SMD, and SMI. OACRS was identified as an accurate predictor compared to SMD and SMI.

#### **DISCUSSION**

Radiomics assessment is a novel approach for evaluating tumor prognosis. SMI and SMD have been identified as biomarkers associated with OS in multiple cancers; however, the role of radiomics in gastric cancer remains poorly understood. A total of 141 patients with stage II/III gastric cancer from our center (discovery cohort) were

included in this study. OACRS, determined from 36 radiomics features selected from approximately 1700 radiomics features, was identified as a novel biomarker. Patients with high OACRS scores had poorer OS than those with low OACRS ( $P < 0.05$ ). A similar result was observed in the TCIA cohort of 40 stage II/III gastric cancer patients. Univariate and multivariate Cox regression were conducted to further elucidate the relationship between OACRS and OS; OACRS was identified as a risk factor for OS [RR = 3.023 (1.896-4.365),  $P < 0.001$ ], independent of pathological and manual image features. Moreover, TAUCs demonstrated that OACRS outperformed SMI and SMD in predicting OS. Furthermore, OACRS incorporated into the prediction model exhibited improved OS accuracy. Notably, OACRS was significantly associated with OS, providing useful complementary information regarding the prognosis of patients with gastric cancer. To meet the clinical practice requirements, an RSF model was developed based on the OACRS, with C-indices of 0.835 and 0.806 for 3- and 5-year OS, respectively. OS prediction can be significantly improved by including OACRS compared to manual image indicators (SMI and SMD).

The OS of locally advanced gastric cancer remains very poor and receives radical surgery and extended nodal dissection. Gastric cancer is heterogeneous, necessitating accurate prognostic prediction for the selection of appropriate treatment or long-term management. Compared to expensive or invasive assessments, noninvasive and inexpensive biomarkers are more easily accepted by patients and clinicians. Imaging data provides some opportunities to overcome these challenges, and accumulating studies have been reported to predict the outcome of gastric cancer using preoperative radiological imaging. One broad method has been explored. Many previous studies have reported the prognostic value of CT-associated indicators for malignant tumors, including visceral fat area, skeletal muscle area, SMI, and SMD<sup>[17-19]</sup>, and SMI was a widely reported prognostic indicator linked to tumor prognosis<sup>[19,20]</sup>. However, contradictory reports have revealed its relationship with nutritional status<sup>[21,22]</sup>. It is widely acknowledged that medical images contain information beyond manual quantitative features. The radiomics approach can automatically extract thousands of

features, including shape, texture, and wavelets, which may be closely linked to some tumor characteristics. As radiological phenotypes are determined by the underlying pathophysiology, there is a subtle association between radioactive changes and tumor pathophysiology.

Gastric cancer is a highly heterogeneous disease, and a reliable and comprehensive evaluation of gastric cancer may provide new insights into improving OS. Pathology and genomics are two typical approaches for predicting OS; however, the need for high-quality tissue and intratumoral spatial heterogeneity is limited in clinical practice. Radiomics provides some unique advantages that allow the evaluation of tumors in a wide-ranging and non-invasive manner, and growing evidence has revealed that radiomics is associated with the response to chemoradiotherapy, immunotherapy, and the heterogeneity of tumor cells<sup>[23-26]</sup>. Moreover, several previous studies have revealed the potential association between radiomics and the tumor microenvironment, such as the number of tumor-infiltrating lymphocytes, macrophages, and tumor stroma<sup>[27,28]</sup>. These findings suggest that radiomics provides a wealth of evidence associated with gastric cancer spatial heterogeneity and the tumor microenvironment related to prognosis. Accordingly, it is reasonable to guide OS using OACRS. <sup>2</sup> To the best of our knowledge, this is the first study to compare the prognostic value of radiomics and certain manual imaging indicators, as well as to evaluate the prognostic value of combined radiomics features and clinical variables in predicting OS for stage II /III gastric cancer. Considering the nature of retrospective studies, future prospective research is warranted regarding their ability to predict OS.

## **CONCLUSION**

In summary, this study demonstrated that the independent predictive value of radiomics for stage II /III gastric cancer outperformed SMI and SMD. Moreover, a prediction model incorporating OACRS was developed to meet the clinical application requirements, which exhibited exceptional discrimination potential. This study had several limitations. First, a selection bias was unavoidable because this was a

retrospective study. Second, the types and timing of anticancer drugs and radiotherapy applications were excluded, which may have increased the uncertainty in the results. Third, other factors associated with body composition, such as glucocorticoid usage and athlete status, were excluded, which may have challenged the accuracy of the results. Finally, the RSF model requires external validation. Consequently, prospective studies with large sample sizes are recommended to further validate the correlation between radiomics and stage II/III gastric cancer OS.

2%

SIMILARITY INDEX

PRIMARY SOURCES

1	Jiao Zhou, Sheng Wei, Xiumei Guo, Yanjun Huang, Yizheng Zhang, Yuming Hong, Xiaofang Chen, Ming Lu, Feng Zheng, Chaohui Zheng. "Correlation between preoperative peripheral blood NLR, PLR, LMR and prognosis of patients with head and neck squamous cell carcinoma", BMC Cancer, 2023	16 words — < 1%
2	www.dovepress.com	13 words — < 1%
3	profiles.stanford.edu	10 words — < 1%
4	worldwidescience.org	10 words — < 1%
5	www.science.gov	10 words — < 1%

EXCLUDE QUOTES ON  
EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES < 10 WORDS  
EXCLUDE MATCHES < 10 WORDS