

<https://doi.org/10.1038/s41746-025-02183-z>

# A transformer-based prognostic signature integrating tumor and body composition CT images predicts postoperative recurrence in gastric cancer



Qiuying Chen<sup>1,5</sup>, Hua Xiao<sup>2,5</sup>, Yueyue Li<sup>1,5</sup>, Lian Jian<sup>1,3,5</sup>, Lu Zhang<sup>1,5</sup>, Bo Lai<sup>1</sup>, Xuewei Wu<sup>1,4</sup>, Jingjing You<sup>1</sup>, Zhe Jin<sup>1</sup>, Hui Shen<sup>1</sup>, Jie Sun<sup>1</sup>, Wenle He<sup>1</sup>, Shuixing Zhang<sup>1</sup> ✉ & Bin Zhang<sup>1</sup> ✉

Accurate preoperative prognosis prediction is crucial for gastric cancer (GC) treatment planning, yet existing models overlook body composition integration. This study demonstrates the potential of integrating multimodal data, including skeletal muscle (SM), adipose tissue (AT), and primary tumor computed tomography images, to improve prognostic stratification in GC patients using an entire cohort of 1862 patients. By leveraging a Vision Transformer-based deep learning approach, we developed and validated a SM-AT-Tumor-Clinical (SMAT-TC) integrated score to predict recurrence-free survival (RFS) in GC patients. The SMAT-TC score achieved a C-index of 0.966 (95% CI: 0.937–0.990), 0.890 (95% CI: 0.866–0.915), and 0.855 (95% CI: 0.829–0.881) in the training, internal validation, and external validation cohorts, respectively, outperforming the Clinical, SM, AT, Tumor, Tumor-Clinical (TC), and SM-Tumor-Clinical (SM-TC) models. The net reclassification improvement and integrated discrimination improvement confirmed the incremental value of body composition. The SMAT-TC score was an independent risk factor for recurrence. The SMAT-TC model could stratify patients into high-, medium-, and low-risk groups with distinct 3- (99.6% vs. 67.0% vs. 10.9%) and 5-year RFS rates (98.8% vs. 61.7% vs. 2.4%). Collectively, the SMAT-TC score may serve as a novel imaging biomarker for GC patients, enhancing risk stratification and guiding individualized treatment strategies.

Gastric cancer (GC) ranks as the fifth most common cancer and the fifth leading cause of cancer-related mortality worldwide<sup>1</sup>. Although the prognosis of GC patients is gradually improving, more than half of patients remain suffering recurrence, even after curative resection<sup>2,3</sup>. Reliable preoperative assessment is crucial for guiding treatment strategies and improving patient outcomes. Current prognostic stratification relies primarily on the established TNM staging system<sup>4</sup>. However, significant survival disparities remain among patients with identical tumor stages undergoing similar therapeutic regimens<sup>5</sup>. Therefore, there is an urgent need to identify more powerful biomarkers that can provide prognostic information, thereby optimizing therapeutic planning and monitoring.

Malnutrition represents a prevalent and serious concern among GC patients. The progression of GC correlates with a drastic decline in nutritional and functional status, leading to poor treatment response and unfavorable prognosis<sup>6</sup>. Numerous studies have identified malnutrition as a crucial risk factor for postoperative complications and survival in GC patients<sup>7,8</sup>. Currently, body composition analysis is gaining prominence in nutritional research. Skeletal muscle (SM) and adipose tissue (AT), as key body composition parameters, serve as indicators of nutritional status<sup>8,9</sup>. The gold standard for evaluating body composition quality is computed tomography (CT)<sup>10</sup>. As CT is routinely used for GC diagnosis and staging, preoperative scans contain not only tumor images that reflect biological

<sup>1</sup>Department of Radiology, the First Affiliated Hospital, Jinan University, Guangzhou, Guangdong, China. <sup>2</sup>Department of Hepatobiliary and Intestinal Surgery, Hunan Cancer Hospital and the Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, Hunan, China. <sup>3</sup>Department of Radiology, Hunan Cancer Hospital, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, Hunan, China. <sup>4</sup>Medical Oncology, National Center for Tumor Diseases (NCT), University Hospital Heidelberg, Heidelberg, Germany. <sup>5</sup>These authors contributed equally: Qiuying Chen, Hua Xiao, Yueyue Li, Lian Jian, Lu Zhang. ✉e-mail: [shui7515@126.com](mailto:shui7515@126.com); [xld\\_Jane\\_Eyre@126.com](mailto:xld_Jane_Eyre@126.com)

heterogeneity but also SM and AT images that indicate nutritional status. Integrating these multiple imaging data sources may enhance outcome prediction for GC patients. However, to date, they have been utilized separately for evaluating disease status and prognosis<sup>6,11–16</sup>.

Deep learning (DL), an emerging machine learning technique, offers significant potential for discovering novel imaging biomarkers using cross-sectional imaging data. DL has achieved satisfactory performance in

predicting histopathological characteristics, treatment response, and clinical outcomes of GC<sup>11,12,16,17</sup>. Transformers are a promising DL technology for medical imaging applications, having shown remarkable success in natural language processing and computer vision<sup>18</sup>. Its core self-attention mechanism overcomes convolutional operation limitations, enabling superior capture of long-range dependencies<sup>19</sup>. Additional advantages include greater scalability, enhanced robustness against common corruptions and perturbations, and powerful parallel computing capabilities<sup>20</sup>. Furthermore, its weak inductive bias allows Transformer to outperform convolutional neural networks (CNNs) when leveraging large-scale models and datasets<sup>20</sup>. Transformer-based models have achieved promising results across multiple medical imaging applications<sup>21–23</sup>, suggesting potential for GC image analysis.

In this study, we developed and validated Transformer-based DL models to predict postoperative recurrence in GC patients using baseline CT images of both tumors and body composition. We aimed to examine whether integrating the different imaging sources could enhance the models' predictive performance and risk stratification. The robustness and clinical applicability of this model were further validated using independent cohorts, ensuring its reliability across diverse clinical settings.

## Results

### Baseline characteristics

The study comprised 1862 eligible patients (mean age,  $57.5 \pm 11.1$  years; 1207 [64.8%] male). Baseline characteristics of patients in the training ( $n = 1242$ ), internal validation ( $n = 311$ ), and external validation ( $n = 309$ ) cohorts are summarized in Table 1. Most patients (96.3%, 1793/1862) were diagnosed with pathological stage II or III disease, and 1322 (70.9%) received postoperative adjuvant chemotherapy (AC). Over a median follow-up of 23.0 months (interquartile range [IQR], 11.9–47.0), 745 (40.0%) patients experienced recurrence. The recurrence rates were 38.3% (476/1242) in the training, 38.3% (119/311) in the internal validation, and 48.5% (150/309) in the external validation cohorts. Locoregional recurrence occurred in 3.6% (45), 3.9% (12), and 15.2% (47) of patients across the three cohorts, respectively, while distant recurrence was observed in 34.7% (431), 34.4% (107), and 33.3% (103), respectively.

Satisfactory inter-reader segmentation consistency was achieved, with a mean Dice Similarity Coefficient (DSC) of  $0.951 \pm 0.034$  and a mean Average Surface Distance (ASD) of  $0.104 \text{ mm} \pm 0.112$ . The 95th Percentile Hausdorff Distance (HD95) was 0 mm in the majority of cases (60%), confirming high boundary agreement. These results collectively indicate that inter-reader variability was minimal and unlikely to introduce significant bias.

### Comparison of the performance of predictive models

After multivariable Cox regression analysis of the clinicopathological data, carcinoembryonic antigen (CEA) level, tumor location, T stage, N stage, positive lymph node ratio (LNR), and AC were identified as independent risk factors associated with tumor recurrence (all  $P < 0.05$ ). The Clinical model based on the independent risk factors achieved a Harrell's concordance index (C-index) of 0.782 (95% Confidence Interval [CI]: 0.757–0.808), 0.737 (95% CI: 0.713–0.761), and 0.603 (95% CI: 0.587–0.621) in forecasting recurrence-free survival (RFS) in the training, internal validation, and external validation cohorts, respectively (Table 2).

Table 2 shows the progressively improved performance of the models in the prediction of RFS. Overall, the combined models (Tumor-Clinical [TC], SM-Tumor-Clinical [SM-TC], and SM-AT-Tumor-Clinical [SMAT-TC] models) outperformed the single models (Clinical, Tumor, SM, and AT models). In both internal and external validation cohorts, the SMAT-TC model yielded higher performance than other models, achieving C-indexes of 0.890 (95% CI: 0.866–0.915) and 0.855 (95% CI: 0.829–0.881), respectively. When stratifying by recurrence pattern, the model achieved C-indexes of 0.937, 0.859, and 0.827 for locoregional recurrence and 0.982,

**Table 1 | Clinicopathologic features of patients in the training, internal, and external validation cohorts**

| Clinicopathologic features | Training cohort ( $n = 1242$ ) | Internal validation cohort ( $n = 311$ ) | External validation cohort ( $n = 309$ ) |
|----------------------------|--------------------------------|------------------------------------------|------------------------------------------|
| Age, years                 | $57.2 \pm 11.0$                | $56.4 \pm 10.6$                          | $59.7 \pm 11.8$                          |
| Sex                        |                                |                                          |                                          |
| Male                       | 814 (65.5)                     | 201 (64.6)                               | 192 (62.1)                               |
| Female                     | 428 (34.5)                     | 110 (35.4)                               | 117 (37.9)                               |
| CEA level                  |                                |                                          |                                          |
| Normal                     | 965 (77.7)                     | 249 (80.1)                               | 254 (82.2)                               |
| Elevated                   | 152 (12.2)                     | 35 (11.3)                                | 52 (16.8)                                |
| Unknown                    | 125 (10.1)                     | 27 (8.7)                                 | 3 (1.0)                                  |
| T stage                    |                                |                                          |                                          |
| T1                         | 70 (5.6)                       | 15 (4.8)                                 | 15 (4.9)                                 |
| T2                         | 116 (9.3)                      | 33 (10.6)                                | 19 (6.1)                                 |
| T3                         | 228 (18.4)                     | 51 (16.4)                                | 165 (53.4)                               |
| T4                         | 828 (66.7)                     | 212 (68.2)                               | 110 (35.6)                               |
| N stage                    |                                |                                          |                                          |
| N0                         | 243 (19.6)                     | 54 (17.4)                                | 78 (25.2)                                |
| N1                         | 269 (21.7)                     | 72 (23.2)                                | 47 (15.2)                                |
| N2                         | 308 (24.8)                     | 69 (22.2)                                | 68 (22.0)                                |
| N3a                        | 308 (24.8)                     | 88 (28.3)                                | 83 (26.9)                                |
| N3b                        | 114 (9.2)                      | 28 (9.0)                                 | 33 (10.7)                                |
| TNM stage                  |                                |                                          |                                          |
| I                          | 37 (3.0)                       | 11 (3.5)                                 | 21 (6.8)                                 |
| II                         | 389 (31.3)                     | 84 (27.0)                                | 95 (30.7)                                |
| III                        | 816 (65.7)                     | 216 (69.5)                               | 193 (62.5)                               |
| Differentiation            |                                |                                          |                                          |
| Low                        | 1112 (89.5)                    | 268 (86.2)                               | 226 (73.1)                               |
| Middle/high                | 129 (10.5)                     | 43 (13.8)                                | 83 (26.9)                                |
| Tumor location             |                                |                                          |                                          |
| Fundus                     | 110 (8.9)                      | 35 (11.3)                                | 42 (13.6)                                |
| Body                       | 310 (25.0)                     | 85 (27.3)                                | 46 (14.9)                                |
| Antrum                     | 782 (63.0)                     | 184 (59.2)                               | 201 (65.0)                               |
| $\geq 2$                   | 40 (3.2)                       | 7 (2.3)                                  | 20 (6.5)                                 |
| Lymph node ratio           | $0.3 \pm 0.3$                  | $0.3 \pm 0.3$                            | $0.3 \pm 0.3$                            |
| Adjuvant chemotherapy      |                                |                                          |                                          |
| No                         | 298 (24.0)                     | 93 (29.9)                                | 149 (48.2)                               |
| Yes                        | 944 (76.0)                     | 218 (70.1)                               | 160 (51.8)                               |
| Recurrence                 |                                |                                          |                                          |
| No                         | 766 (61.7)                     | 192 (61.7)                               | 159 (51.5)                               |
| Locoregional               | 45 (3.6)                       | 12 (3.9)                                 | 47 (15.2)                                |
| Distant                    | 431 (34.7)                     | 107 (34.4)                               | 103 (33.3)                               |

Continuous variable was expressed as mean  $\pm$  standard deviation, and categorical variable was expressed as count (percent).

0.902, and 0.861 for distant recurrence across the training, internal validation, and external validation cohorts, respectively.

In the internal validation cohort, the net reclassification index (NRI) and integrated discrimination improvement (IDI) for the SM-TC model compared to the TC model were 0.074 and 0.009, respectively; in the external validation cohort, these values were 0.185 and 0.019 (all  $P < 0.05$ ), indicating improved discrimination ability when adding the SM features (Table 3). Similarly, the SMAT-TC model also yielded positive and statistically significant NRI and IDI values versus the SM-TC model, suggesting that the addition of AT features further improved model performance (Table 3). These findings show the added value of body composition biomarkers in predicting GC prognosis.

### Predictive ability of the SMAT-TC model

Figure 1a–c shows the time-dependent receiver operating characteristic (ROC) curves of the SMAT-TC model for predicting RFS. This combined model accurately predicted 1-, 3-, and 5-year RFS, with area under the

curves (AUCs) of 0.964 (95% CI: 0.950–0.986), 0.960 (95% CI: 0.945–0.981), and 0.960 (95% CI: 0.946–0.981) in the training cohort, 0.881 (95% CI: 0.868–0.901), 0.901 (95% CI: 0.887–0.920), and 0.897 (95% CI: 0.883–0.917) in the internal validation cohort, and 0.855 (95% CI: 0.842–0.874), 0.860 (95% CI: 0.847–0.879), and 0.853 (95% CI: 0.841–0.872) in the external validation cohort, respectively. The calibration curves of the SMAT-TC model indicated good agreement between the predicted and observed probabilities for 1-, 3-, and 5-year RFS in GC patients across all cohorts (Fig. 1d–f). Decision curve analysis (DCA) revealed that across clinically relevant probability thresholds, the SMAT-TC model provided greater clinical net benefit than “treat-all” or “treat-none” strategies (Fig. 1g–o).

Patients were stratified into high-, medium-, and low-risk groups using the cutoff values of 0.283 and 0.589 derived from the SMAT-TC model. Figure 2 compares the RFS rates of high-, medium-, and low-risk groups across all cohorts. Risk distribution across cohorts was as follows: training, 27.0% high-risk (335/1242), 46.0% medium-risk (571/1242), 27.1% low-risk (336/1242); internal validation, 23.2% (72/311), 54.3% (169/311), 22.5% (70/311); external validation, 19.1% (59/309), 61.5% (190/309), 19.4% (60/309). The SMAT-TC model discriminated risk groups with progressively declining 3-year (99.6% vs. 67.0% vs. 10.9%) and 5-year RFS rates (98.8% vs. 61.7% vs. 2.4%). Cohort-specific stratification consistently revealed significant RFS differences: for 3-year RFS rates, training (100.0% vs. 72.3% vs. 8.6%), internal validation (100% vs. 68.2% vs. 15.0%), external validation (97.1% vs. 50.0% vs. 19.7%); for 5-year RFS rates, training (100% vs. 68.1% vs. 1.5%), internal validation (100% vs. 59.7% vs. 3.7%), external validation (91.7% vs. 45.9% vs. 5.9%). High-risk groups showed significantly reduced RFS versus medium-risk (training: Hazard Ratio [HR] = 6.2, 95% CI: 5.1–7.6; internal validation: HR = 4.5, 95% CI: 2.9–7.0; external validation: HR = 2.5, 95% CI: 1.6–3.8; all  $P < 0.001$ ) and low-risk groups (training: 154.8, 95% CI: 123.5–194.1; internal validation: infinitely large HR; external validation: HR = 23.9, 95% CI: 14.0–40.8; all  $P < 0.001$ ). Medium-risk patients also showed shorter RFS than low-risk patients (training: 25.5, 95% CI: 18.4–35.3; internal validation: infinitely large HR; external validation: HR = 9.0, 95% CI: 5.8–14.1; all  $P < 0.001$ ). Subgroup analyses further established the model-derived score as a robust predictor across all subgroups (Fig. 3).

Univariate and multivariable Cox regression analyses were performed to identify independent prognostic factors for RFS. As shown in Table 4, the SMAT-TC score was a strong independent predictor of RFS in both the internal (HR = 255.2, 95% CI: 70.1–929.1,  $P < 0.001$ ) and external (HR = 125.7, 95% CI: 44.5–355.2,  $P < 0.001$ ) validation cohorts, after adjusting for clinicopathologic variables. Table 5 shows

that Tumor, SM, and AT scores were independent predictors of RFS in all cohorts (all  $P < 0.001$ ), supporting the complementary prognostic information provided by tumor, SM, and AT features.

### Discussion

In this prognostic study, we developed a CT-based fusion model to improve postoperative recurrence risk prediction in GC patients. By integrating body composition with tumor characteristics, we showed that the fusion model's predictive performance significantly surpassed models utilizing either a

**Table 2 | Predictive performance of models in the training, internal, and external validation cohorts**

| Dataset                    | Model                | RFS (C-Index)              |
|----------------------------|----------------------|----------------------------|
| Training cohort            | Clinical model       | 0.782 (0.757–0.808)        |
|                            | Tumor model          | 0.946 (0.919–0.971)        |
|                            | SM model             | 0.939 (0.907–0.969)        |
|                            | AT model             | 0.914 (0.881–0.959)        |
|                            | TC model             | 0.965 (0.934–0.985)        |
|                            | SM-TC model          | 0.975 (0.944–0.994)        |
|                            | <b>SMAT-TC model</b> | <b>0.966 (0.937–0.990)</b> |
| Internal validation cohort | Clinical model       | 0.737 (0.713–0.761)        |
|                            | Tumor model          | 0.866 (0.844–0.891)        |
|                            | SM model             | 0.858 (0.833–0.888)        |
|                            | AT model             | 0.835 (0.802–0.871)        |
|                            | TC model             | 0.871 (0.842–0.894)        |
|                            | SM-TC model          | 0.888 (0.861–0.917)        |
|                            | <b>SMAT-TC model</b> | <b>0.890 (0.866–0.915)</b> |
| External validation cohort | Clinical model       | 0.603 (0.587–0.621)        |
|                            | Tumor model          | 0.820 (0.796–0.846)        |
|                            | SM model             | 0.808 (0.785–0.830)        |
|                            | AT model             | 0.791 (0.767–0.828)        |
|                            | TC model             | 0.825 (0.797–0.850)        |
|                            | SM-TC model          | 0.846 (0.824–0.875)        |
|                            | <b>SMAT-TC model</b> | <b>0.855 (0.829–0.881)</b> |

SM skeletal muscle, AT adipose tissue, TC tumor-clinical, SM-TC SM-tumor-clinical, SMAT-TC SM-AT-tumor-clinical.

**Table 3 | The incremental prognostic value of SM and AT in predicting recurrence-free survival of GC**

| Models                  | Internal validation cohort | P-value | External validation cohort | P-value |
|-------------------------|----------------------------|---------|----------------------------|---------|
| SM-TC vs. TC model      |                            |         |                            |         |
| NRI                     | 0.074 (0.046–0.125)        | 0.037   | 0.185 (0.109–0.299)        | 0.016   |
| IDI                     | 0.009 (0.006–0.024)        | 0.048   | 0.019 (0.010–0.031)        | 0.029   |
| SMAT-TC vs. SM-TC model |                            |         |                            |         |
| NRI                     | 0.004 (0.003–0.012)        | 0.020   | 0.073 (0.045–0.114)        | 0.010   |
| IDI                     | 0.018 (0.011–0.027)        | 0.051   | 0.033 (0.022–0.052)        | 0.024   |

NRI net reclassification improvement, IDI integrated discrimination improvement, SM skeletal muscle, AT adipose tissue, RFS recurrence-free survival, GC gastric cancer.

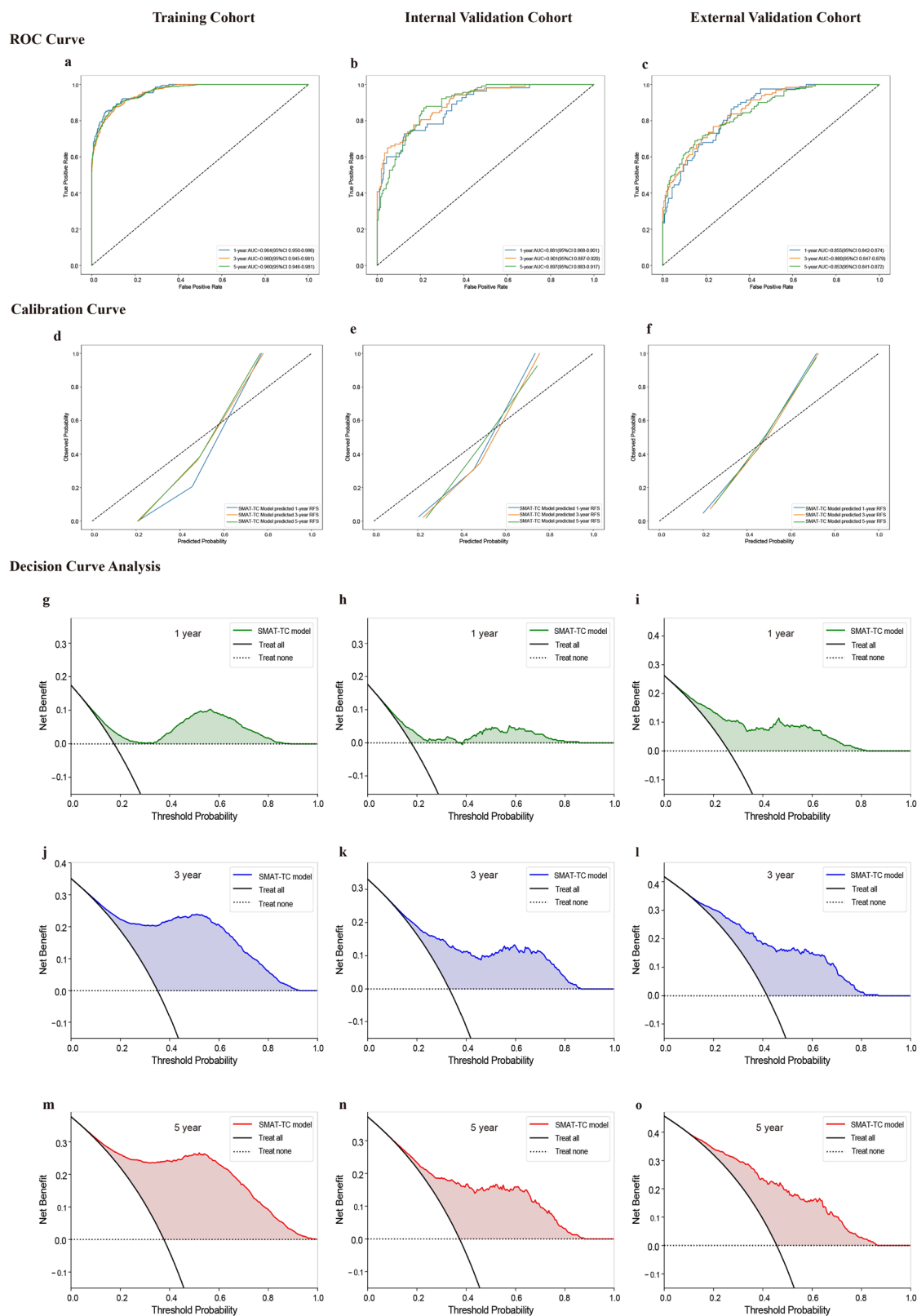


**Fig. 1 | Performance of the SMAT-TC model for predicting 1-, 3-, and 5-year RFS of GC patients in the training, internal validation, and external validation cohorts, respectively. a–c** Time-dependent receiver operating characteristic curves; **d–f** Calibration curves; **g–o** Decision curve analysis. RFS recurrence-free survival, GC gastric cancer.

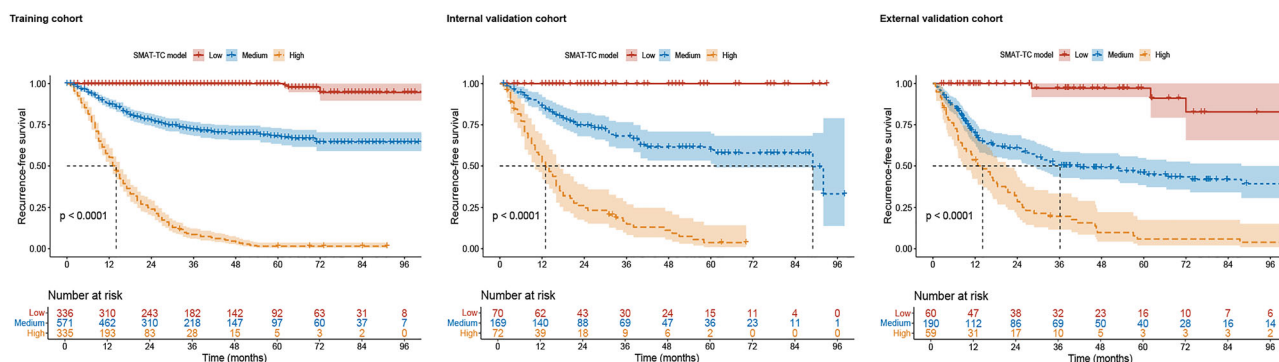


**Fig. 2 | Survival analysis of patient risk groups.** Kaplan-Meier survival curves for patients categorized into high-, medium-, and low-risk groups in the training, internal validation, and external validation cohorts. Statistical significance between groups was determined by the log-rank test.

single data source. This approach illustrates how to maximize existing imaging data without requiring additional invasive testing. The SMAT-TC model, integrating primary tumor, SM, AT, and clinical features, showed superior performance in predicting individual RFS. The results indicated that the SMAT-TC model-predicted risk score was an independent risk factor associated with tumor recurrence risk. This score stratified patients into three prognostic subgroups with distinct RFS rates while identifying high-risk patients who may benefit from intensified therapies such as AC.

The third lumbar vertebra (L3-level) SM area on CT images is well correlated with whole-body muscle mass<sup>24</sup>. Due to its automatic and reproducible measurement, it has been widely used to diagnose sarcopenia, which is closely associated with cancer-related cachexia and decreased survival rates. Numerous studies have confirmed that sarcopenia adversely affects postoperative complications, chemotherapy efficacy, and long-term survival in GC patients<sup>14,25–30</sup>. However, only several studies have integrated SM status alongside tumor features in their predictive models. Lan et al.<sup>31</sup> developed a nomogram for predicting postoperative complications in GC patients, which includes the NRS2002 score, sarcopenia, blood loss, and tumor radiomic signature. This model showed superior predictive accuracy after adding sarcopenia to the radiomic signature and conventional clinical features. Similarly, Anconina et al.<sup>32</sup> found that combining tumor radiomic features with imaging-derived sarcopenia on pretreatment<sup>18</sup>. F-FDG PET/CT further improved survival prediction in surgically treated adenocarcinoma esophagogastric cancer patients. These studies collectively confirmed the incremental value of SM for predicting GC prognosis.

Compared to conventional morphological metrics, DL provides a more accurate assessment of body composition. In our previous work<sup>33</sup>, we constructed a DL-based SM model to predict survivals in GC patients undergoing radical gastrectomy, showing higher accuracy compared to conventional sarcopenia metrics. In this current study, our integrated SM-TC and SMAT-TC models outperformed models based solely on SM, AT, or TC, underscoring the diagnostic superiority of combined biomarkers. Unlike previous studies that utilized only the SM or AT area for modeling, we extracted high-dimensional DL features that go beyond human perception to enhance predictive accuracy. Furthermore, by processing volumetric spatial information rather than single-slice tumor data as DL network input, we enabled sufficient extraction of tumor heterogeneity. Beyond target regions, we incorporated surrounding tissue influences by inputting entire image slices into the network. This allows the DL architecture to autonomously learn global context and establish long-range dependencies, thereby achieving enhanced predictive accuracy.

Despite the widespread success of CNNs in medical imaging over the past decade, inherent architectural limitations constrain their performance potential. Most current CNNs employ small convolution kernels, introducing bias toward local spatial structures that hampers the modeling of long-range dependencies essential for comprehensive contextual understanding<sup>34</sup>. Transformer has large theoretical and effective receptive fields, thus possibly bring better understanding of contextual information

and long-range dependency than CNNs. This advantage is particularly evident in the 3D Transformer for volumetric data, where anatomical context is critical for medical image analysis. As a result, researchers have leveraged these properties to develop Transformers that outperform CNNs across various medical imaging tasks. To validate the superiority of the Transformer over CNN, we compared their predictive performance in our previous study<sup>33</sup>. As a result, the Transformed-based DL model exhibited better discrimination ability both in the internal and external validation cohorts, achieving C-indexes of 0.888 and 0.846 for predicting RFS, respectively. These results highlight the superior predictive accuracy of the Transformer as a network framework.

Our study also has some limitations. First, as our cohort exclusively comprised Chinese patients, the generalizability of our findings to other populations remains uncertain. Pathological subtypes, clinical characteristics, and disease manifestations may differ in Western populations; thus, external validation across diverse ethnic groups is essential. Second, body composition analysis was confined to SM and AT at the L3 level. Future studies may consider whole-abdominal body composition to better capture regional heterogeneity. Additionally, the model relies on manually segmented tumor regions, which may introduce interreader variability despite the satisfactory segmentation consistency in our study. Integrating automated segmentation tools into the pipeline would improve reproducibility and clinical applicability. Third, the retrospective nature of the study limited our ability to analyze chemotherapy compliance in relation to multimodal features such as SM composition. Due to the lack of prospectively collected detailed data on dose modifications, treatment interruptions, and their underlying causes (e.g., toxicities, patient-related factors, and socioeconomic barriers), we were unable to develop robust predictive models for chemotherapy compliance. This gap should be addressed in future prospective studies designed to systematically collect such information. Finally, the inherent “black box” nature of DL models remains a challenge. Developing more intuitive and clinically transparent interpretation methods is critical to fostering trust and facilitating the integration of DL-based tools into routine practice.

In this current study, we developed and validated a prognostic and predictive biomarker, a multisource imaging risk score, which allows GC risk stratification by mining features from primary tumor, SM, and AT on routinely acquired abdominal CT scans. The addition of L3-level body composition images to the tumor images could enhance the accuracy of recurrence risk prediction. The findings highlight the score’s potential to refine personalized therapeutic strategies, reduce overtreatment, and improve clinical decision-making in GC patients after surgery.

## Methods

### Study design and patient cohorts

The flowchart of the study population is presented in Fig. 4. We collected two independent patient cohorts. First, a primary cohort from Center 1 was used for model development and hyperparameter optimization, comprising

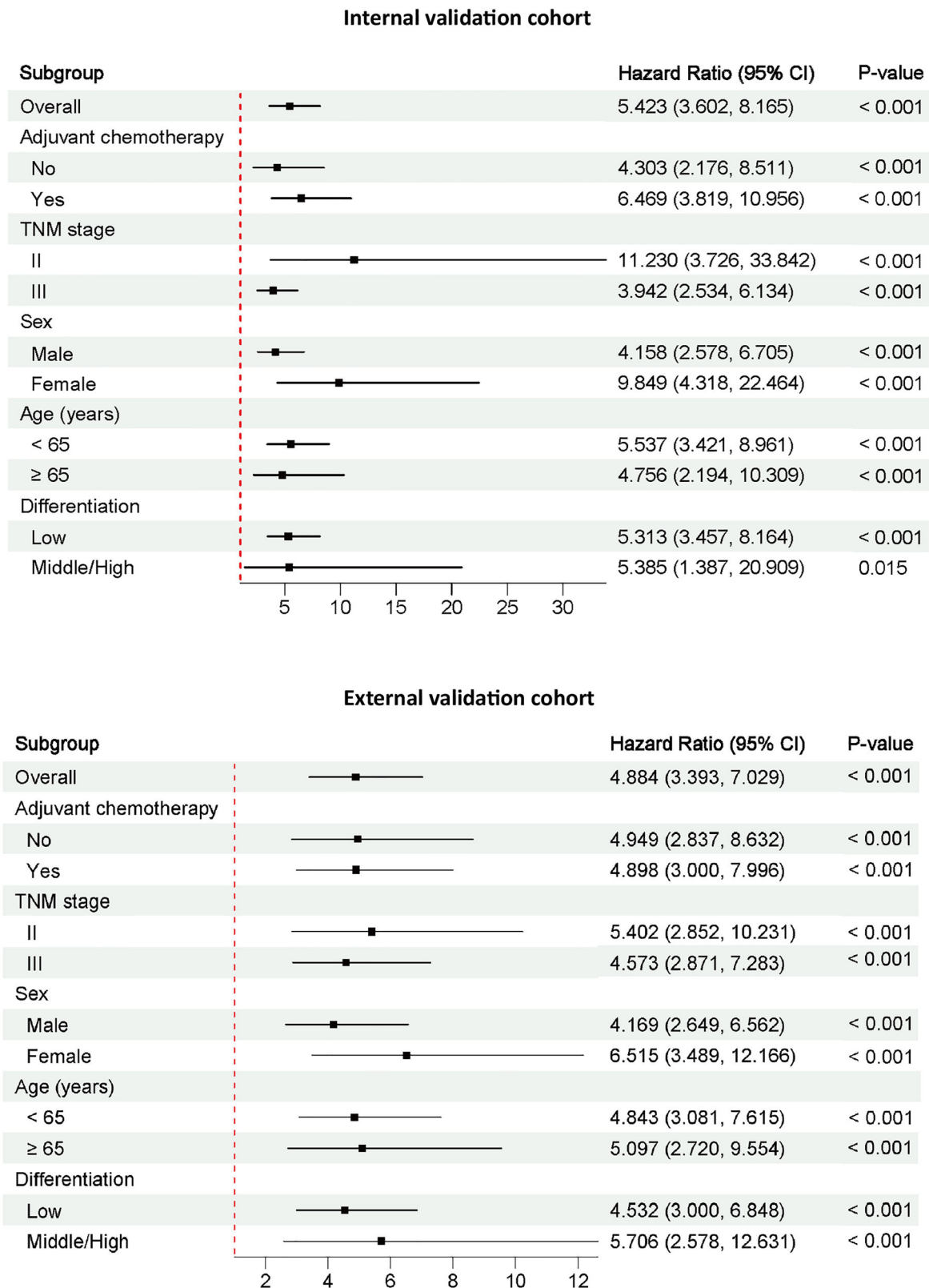


**Fig. 3 | Forest plots of subgroup analyses of the association between the SMAT-TC model score and RFS in the internal and external validation cohorts.** RFS recurrence-free survival.

1900 patients who underwent radical (R0) resection for GC between November 2010 and December 2020. The criteria for inclusion were as follows: (1) age  $\geq 18$  years, (2) histopathologically confirmed GC, (3) abdominal contrast-enhanced CT scan conducted within four weeks before surgery, and (4) complete clinical and follow-up data. Exclusion criteria

included: (1) preoperative anticancer treatment, (2) concurrent malignancies, (3) prior gastrectomy, (4) chronic hepatic/renal disease or wasting disorders, and (5) inadequate CT data (missing scans, poor image quality, incomplete L3 vertebral coverage, or non-visualized tumors). A total of 1553 patients were eligible and were randomly assigned to a training cohort (1242

**Table 4 | Univariate and multivariate Cox regression analyses of prognostic factors in patients with gastric cancer**

| Clinicopathological features | Internal validation cohort |                  |                             |                  | External validation cohort |                  |                             |                  |
|------------------------------|----------------------------|------------------|-----------------------------|------------------|----------------------------|------------------|-----------------------------|------------------|
|                              | RFS (Univariate analysis)  |                  | RFS (Multivariate analysis) |                  | RFS (Univariate analysis)  |                  | RFS (Multivariate analysis) |                  |
|                              | HR (95% CI)                | P-value          | HR (95% CI)                 | P-value          | HR (95% CI)                | P-value          | HR (95% CI)                 | P-value          |
| CEA level                    |                            |                  |                             |                  |                            |                  |                             |                  |
| Normal                       | Ref.                       |                  |                             |                  | Ref.                       |                  |                             |                  |
| Elevated                     | 1.3 (0.7, 2.2)             | 0.379            |                             |                  | 0.7 (0.4, 1.1)             | 0.116            |                             |                  |
| Unknown                      | 0.5 (0.2, 1.1)             | 0.088            |                             |                  | 0.0 (0.0, inf)             | 1.000            |                             |                  |
| T stage                      |                            |                  |                             |                  |                            |                  |                             |                  |
| T1                           | Ref.                       |                  |                             |                  | Ref.                       |                  |                             |                  |
| T2                           | 0.4 (0.1, 2.2)             | 0.317            |                             |                  | 0.6 (0.2, 1.5)             | 0.272            |                             |                  |
| T3                           | 1.4 (0.4, 4.9)             | 0.648            |                             |                  | 1.0 (0.5, 2.0)             | 0.970            |                             |                  |
| T4                           | 3.0 (1.0, 9.5)             | 0.060            |                             |                  | 0.7 (0.3, 1.5)             | 0.320            |                             |                  |
| N stage                      |                            |                  |                             |                  |                            |                  |                             |                  |
| N0                           | Ref.                       |                  | Ref.                        |                  | Ref.                       |                  | Ref.                        |                  |
| N1                           | 0.7 (0.3, 1.7)             | 0.476            | 0.8 (0.3, 1.8)              | 0.581            | 0.9 (0.5, 1.5)             | 0.659            | 0.8 (0.4, 1.3)              | 0.336            |
| N2                           | 2.0 (1.0, 4.1)             | <b>0.047</b>     | 1.0 (0.4, 2.4)              | 0.97             | 1.7 (1.1, 2.7)             | <b>0.026</b>     | 0.7 (0.4, 1.2)              | 0.196            |
| N3a                          | 3.5 (1.8, 6.7)             | <b>&lt;0.001</b> | 0.7 (0.2, 1.9)              | 0.455            | 1.2 (0.7, 1.9)             | 0.512            | 0.4 (0.2, 0.8)              | <b>0.013</b>     |
| N3b                          | 7.0 (3.4, 14.5)            | <b>&lt;0.001</b> | 0.6 (0.1, 2.2)              | 0.395            | 1.9 (1.1, 3.3)             | <b>0.019</b>     | 0.4 (0.2, 1.1)              | 0.066            |
| Tumor location               |                            |                  |                             |                  |                            |                  |                             |                  |
| Fundus                       | Ref.                       |                  | Ref.                        |                  | Ref.                       |                  | Ref.                        |                  |
| Body                         | 1.1 (0.6, 2.0)             | 0.807            | 0.9 (0.5, 1.8)              | 0.841            | 0.5 (0.3, 1.0)             | <b>0.034</b>     | 0.5 (0.3, 0.8)              | <b>0.008</b>     |
| Antrum                       | 0.6 (0.4, 1.2)             | 0.138            | 0.7 (0.4, 1.4)              | 0.339            | 0.5 (0.3, 0.8)             | <b>0.005</b>     | 0.4 (0.3, 0.7)              | <b>&lt;0.001</b> |
| ≥2 locations                 | 3.1 (1.1, 8.2)             | <b>0.026</b>     | 2.4 (0.8, 6.9)              | 0.100            | 0.7 (0.3, 1.4)             | 0.284            | 0.5 (0.3, 1.2)              | 0.121            |
| Lymph node ratio             | 17.5 (9.2, 33.2)           | <b>&lt;0.001</b> | 6.9 (1.5, 30.8)             | <b>0.012</b>     | 2.3 (1.4, 3.6)             | <b>0.001</b>     | 3.0 (1.2, 7.6)              | <b>0.017</b>     |
| Adjuvant chemotherapy        |                            |                  |                             |                  |                            |                  |                             |                  |
| No                           | Ref.                       |                  |                             |                  | Ref.                       |                  | Ref.                        |                  |
| Yes                          | 0.8 (0.5, 1.1)             | 0.202            |                             |                  | 1.7 (1.2, 2.3)             | <b>0.002</b>     | 1.6 (1.2, 2.3)              | <b>0.005</b>     |
| SMAT-TC score                | 595.9 (194.5, 1826.0)      | <b>&lt;0.001</b> | 255.2 (70.1, 929.1)         | <b>&lt;0.001</b> | 138.2 (51.2, 372.6)        | <b>&lt;0.001</b> | 125.7 (44.5, 355.2)         | <b>&lt;0.001</b> |

Ref. Reference, RFS Recurrence-free survival, HR Hazard ratio. The bolded P-values indicate that the results were statistically significant ( $P < 0.05$ ).

patients, 23821 tumor slices, and 1242 body composition slices at the L3 level and an internal validation cohort (311 patients, 5893 tumor slices, and 311 body composition slices at the L3 level) at an 4:1 ratio. Second, an external validation cohort from Center 2 ( $n = 326$ ) comprised patients who underwent R0 resection between April 2008 and January 2018, using the same inclusion and exclusion criteria. This cohort assessed model generalizability. Finally, 309 patients (7911 tumor slices and 309 body composition slices at the L3 level) were collected in the external validation cohort. The study protocol was approved by the local ethics committee and institutional review board of each hospital (First Affiliated Hospital of Jinan University: KY-2024-203; Hunan Cancer Hospital: KYJJ-2025-285), and complied with the Declaration of Helsinki. The ethics committee waived the requirement for informed consent due to the retrospective nature and the use of anonymized data for analysis.

All gastrectomies (i.e., distal, proximal, or total) were determined by primary lesion location and performed by specialized surgeons using either laparoscopic or open procedures. Lymph node dissection and gastric reconstruction were conducted in accordance with the contemporary Japanese GC Treatment guidelines. Combined multi-organ resection was performed in cases suspected of involvement of adjacent organs by locally advanced tumors. AC was administered to patients at high risk of recurrence within 4–6 weeks after surgery. The standard adjuvant regimen for stage II GC or patients ineligible for doublet chemotherapy consisted of S-1 monotherapy for one year. For stage III disease, a six-month course of fluorouracil-platinum-based combined therapy (e.g., CAPOX or SOX) was recommended. The median number of AC cycles was 3 (interquartile range, 0–6).

The clinicopathological characteristics were obtained from the medical records, including sex, age, CEA level, tumor differentiation grade, tumor location, TNM stage, positive LNR, and AC. Each record was collected by experienced clinicians and was audited by a specially trained medical team. Tumors were restaged according to the eighth edition of the American Joint Committee on Cancer staging system. The LNR was defined as the percentage of positive metastatic lymph nodes in the total number of lymph nodes dissected. The N-stage and LNR were important clinicopathological variables reflecting the lymph node burden. The clinical endpoint was RFS, which was defined as the time from surgery to tumor recurrence or the last follow-up.

### Image acquisition and segmentation

All patients underwent preoperative contrast-enhanced abdominal CT scans within four weeks before surgery. Detailed CT acquisition protocols and parameters are provided in Supplementary Table 1. Images were archived in the picture archiving and communication system. Venous phase images were analyzed due to their superior contrast between tumor tissue and adjacent normal gastric structures.

The original CT images were preprocessed as follows: First, serial transverse venous-phase slices containing tumor lesions were identified. Two board-certified abdominal radiologists (with 5 years of clinical experience) manually delineated tumor volumes using ITK-SNAP (<https://www.itksnap.org>) software. All bounding boxes at each slice were resized to  $128 \times 128 \times 48$  pixels. Second, the areas of SM and AT (including visceral adipose tissue, subcutaneous adipose tissue, and inter-muscular adipose

**Table 5 | Univariate and multivariate Cox regression analyses of prognostic factors in patients with gastric cancer**

| Clinicopathological features | Internal validation cohort |                  |                             |                  | External validation cohort |                  |                             |                  |
|------------------------------|----------------------------|------------------|-----------------------------|------------------|----------------------------|------------------|-----------------------------|------------------|
|                              | RFS (Univariate analysis)  |                  | RFS (Multivariate analysis) |                  | RFS (Univariate analysis)  |                  | RFS (Multivariate analysis) |                  |
|                              | HR (95% CI)                | P-value          | HR (95% CI)                 | P-value          | HR (95% CI)                | P-value          | HR (95% CI)                 | P-value          |
| CEA level                    |                            |                  |                             |                  |                            |                  |                             |                  |
| Normal                       | Ref.                       |                  |                             |                  | Ref.                       |                  |                             |                  |
| Elevated                     | 1.3 (0.7, 2.2)             | 0.379            |                             |                  | 0.7 (0.4, 1.1)             | 0.116            |                             |                  |
| Unknown                      | 0.5 (0.2, 1.1)             | 0.088            |                             |                  | 0.0 (0.0, inf)             | 1.000            |                             |                  |
| T stage                      |                            |                  |                             |                  |                            |                  |                             |                  |
| T1                           | Ref.                       |                  |                             |                  | Ref.                       |                  |                             |                  |
| T2                           | 0.4 (0.1, 2.2)             | 0.317            |                             |                  | 0.6 (0.2, 1.5)             | 0.272            |                             |                  |
| T3                           | 1.4 (0.4, 4.9)             | 0.648            |                             |                  | 1.0 (0.5, 2.0)             | 0.970            |                             |                  |
| T4                           | 3.0 (1.0, 9.5)             | 0.060            |                             |                  | 0.7 (0.3, 1.5)             | 0.320            |                             |                  |
| N stage                      |                            |                  |                             |                  |                            |                  |                             |                  |
| N0                           | Ref.                       |                  | Ref.                        |                  | Ref.                       |                  | Ref.                        |                  |
| N1                           | 0.7 (0.3, 1.7)             | 0.476            | 0.7 (0.3, 1.6)              | 0.358            | 0.9 (0.5, 1.5)             | 0.659            | 0.8 (0.5, 1.5)              | 0.527            |
| N2                           | 2.0 (1.0, 4.1)             | <b>0.047</b>     | 0.8 (0.3, 1.8)              | 0.562            | 1.7 (1.1, 2.7)             | <b>0.026</b>     | 0.9 (0.5, 1.7)              | 0.853            |
| N3a                          | 3.5 (1.8, 6.7)             | <b>&lt;0.001</b> | 0.7 (0.3, 1.9)              | 0.484            | 1.2 (0.7, 1.9)             | 0.512            | 0.7 (0.3, 1.4)              | 0.268            |
| N3b                          | 7.0 (3.4, 14.5)            | <b>&lt;0.001</b> | 0.7 (0.2, 2.5)              | 0.546            | 1.9 (1.1, 3.3)             | <b>0.019</b>     | 0.4 (0.2, 1.1)              | 0.074            |
| Tumor location               |                            |                  |                             |                  |                            |                  |                             |                  |
| Fundus                       | Ref.                       |                  | Ref.                        |                  | Ref.                       |                  | Ref.                        |                  |
| Body                         | 1.1 (0.6, 2.0)             | 0.807            | 0.7 (0.4, 1.4)              | 0.330            | 0.5 (0.3, 1.0)             | <b>0.034</b>     | 0.7 (0.4, 1.3)              | 0.274            |
| Antrum                       | 0.6 (0.4, 1.2)             | 0.138            | 0.6 (0.3, 1.0)              | 0.061            | 0.5 (0.3, 0.8)             | <b>0.005</b>     | 0.6 (0.4, 0.9)              | <b>0.019</b>     |
| ≥2 locations                 | 3.1 (1.1, 8.2)             | <b>0.026</b>     | 0.3 (0.1, 0.9)              | <b>0.030</b>     | 0.7 (0.3, 1.4)             | 0.284            | 0.7 (0.3, 1.5)              | 0.334            |
| Lymph node ratio             | 17.4 (9.2, 33.2)           | <b>&lt;0.001</b> | 10.5 (2.6, 42.0)            | <b>0.001</b>     | 2.3 (1.4, 3.6)             | <b>0.001</b>     | 4.0 (1.7, 9.3)              | <b>0.002</b>     |
| Adjuvant chemotherapy        |                            |                  |                             |                  |                            |                  |                             |                  |
| No                           | Ref.                       |                  |                             |                  | Ref.                       |                  | Ref.                        |                  |
| Yes                          | 0.8 (0.5, 1.1)             | 0.202            |                             |                  | 1.7 (1.2, 2.3)             | <b>0.002</b>     | 1.3 (0.9, 1.9)              | 0.101            |
| Tumor score                  | 271.1 (88.9, 826.3)        | <b>&lt;0.001</b> | 21.0 (6.0, 73.1)            | <b>&lt;0.001</b> | 347.0 (97.1, 1240.4)       | <b>&lt;0.001</b> | 40.7 (11.5, 144.6)          | <b>&lt;0.001</b> |
| SM score                     | 328.7 (100.3, 1077.7)      | <b>&lt;0.001</b> | 21.0 (5.3, 83.5)            | <b>&lt;0.001</b> | 76.4 (29.1, 200.7)         | <b>&lt;0.001</b> | 12.7 (4.4, 36.3)            | <b>&lt;0.001</b> |
| AT score                     | 173.7 (60.9, 494.9)        | <b>&lt;0.001</b> | 25.0 (7.8, 80.2)            | <b>&lt;0.001</b> | 71.9 (27.7, 187.0)         | <b>&lt;0.001</b> | 17.1 (6.0, 48.9)            | <b>&lt;0.001</b> |

Ref. Reference, RFS Recurrence-free survival, HR Hazard ratio. The bolded P-values indicate that the results were statistically significant ( $P < 0.05$ ).

tissue) at the L3 level were semi-automatically segmented using Darwin's research platform (<https://arxiv.org/abs/2009.00908>). Tissue differentiation among SM, AT, and other tissues was based on Hounsfield units (HU). A threshold range of −29 to 150 HU was used to define SM, and that of −190 to −30 HU was used to define AT. Finally, all segmentations were verified and corrected by a senior abdominal radiologist with 10 years of experience.

Intra-reader reproducibility for manual tumor segmentation was quantitatively evaluated using the DSC, HD95, and ASD on a randomly selected 10% sample from the training cohort. Specifically, the DSC measured volumetric overlap between repeated segmentations; the HD95 quantified the maximum boundary discrepancy at the 95th percentile to mitigate outlier effects; and the ASD provided the mean surface-to-surface difference between contours.

### DL procedures and the optimal model selection

DL models require large-scale datasets, as limited data may cause overfitting or training failure. To mitigate this limitation, we applied data augmentation techniques including mirroring, rotation, resizing, cropping, translation, and Gaussian noise addition. Each technique introduced stochastic transformations through technique-specific parameters (e.g., rotation angles or noise levels).

Our DL framework adopted a multi-stream Transformer architecture specifically designed to integrate complementary information from body composition and tumor characteristics. The model architecture adheres to established Vision Transformer principles, with specific adaptations for medical imaging and multimodal data fusion, informed by successful applications in computational pathology and radiology. As shown in Fig. 5, the model utilized a series of Transformer layers to effectively capture both local and global contextual information, thereby enhancing feature representation robustness. Each Transformer layer consisted of a multi-head self-attention (MSA) module followed by a multi-layer perceptron (MLP) block.

Self-attention (SA) was a variant of the attention mechanism designed to establish internal correlations between each pixel in the abstract feature map, allowing the model to focus on the global spatial relationships within the feature map. The process was as follows:

When the input feature map obtained matrices Q, K, and V through the mapping matrix, the similarity between matrix Q and matrix K was calculated. To preserve spatial information, learnable absolute positional encodings were added to the patch embeddings before computing Q, K, and V. Then, after splicing with matrix K, the weight of each pixel was mapped using soft-max. Finally, the weight obtained and the mapping matrix V were

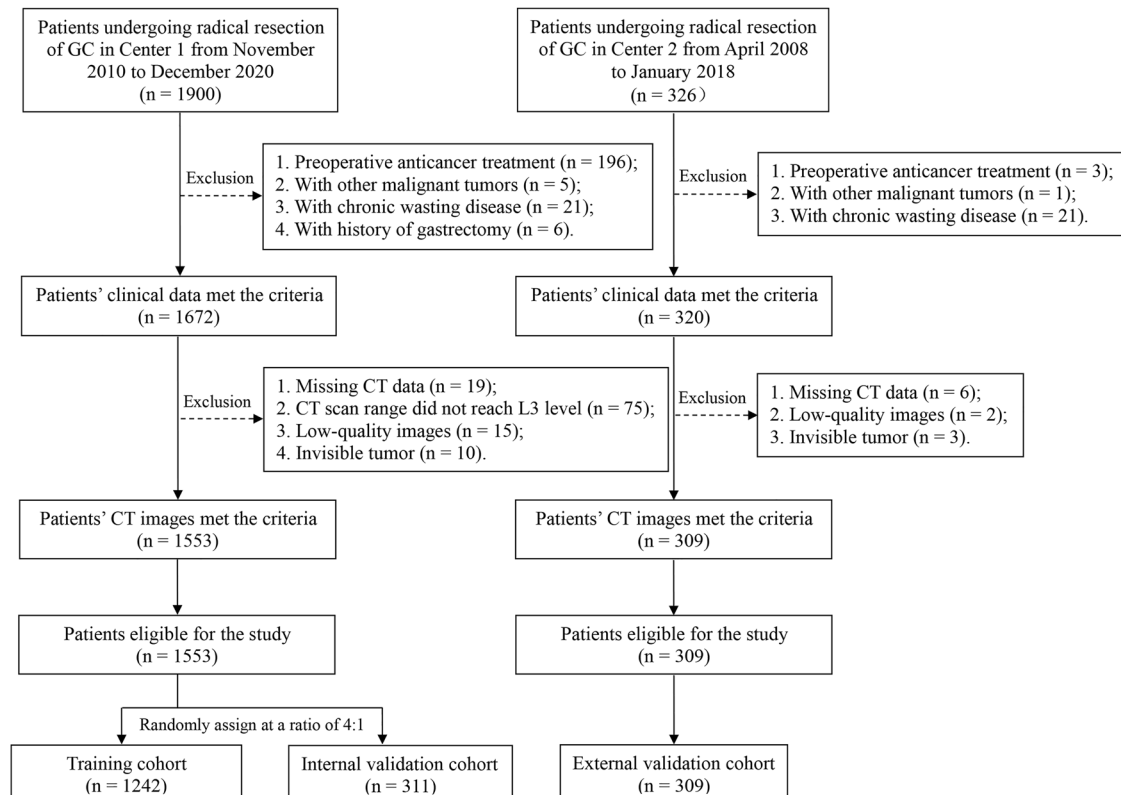


**Fig. 4 | Patient flowchart.** GC gastric cancer, L3 the third lumbar vertebra.

weighted as the output. Attention was calculated as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V = AV$$

The MSA module was formed by the parallel connection of six SA modules to produce multiple output maps. The final output was typically a concatenation and projection of all outputs of SA modules, which could be given by:

$$Z_i = \text{SA}(X \times W_i^q, X \times W_i^k X \times W_i^v),$$

$$\text{MSA}(Q, K, V) = \text{Concat}[Z_1, \dots, Z_h] \times W^o$$

Each MSA module was followed by an MLP block. MLP was a crucial component as it injected inductive bias into the Transformer, which was lacking in the MSA. This was because MLP was local and translation-equivariant, while SA computation was a global operation. The MLP was comprised of two feed-forward networks with a Gaussian Error Linear Unit activation in between:

$$\text{MLP}(x) = \varnothing(xW_1 + b_1)W_2 + b_2$$

Where  $x$  denotes the input, and  $W$  and  $b$  denote, respectively, the weight matrix and the bias of the corresponding linear layer. The dimensions of the weight matrices,  $W_1$  and  $W_2$ , were typically set as  $D \times 4D$  and  $4D \times D$ . Since the input was a matrix of flattened and tokenized patches, applying  $W$  to  $x$  was analogous to applying a convolutional layer with a kernel size of  $1 \times 1$ . Consequently, the MLPs in the Transformer were highly localized and equivariant to translation.

We developed seven models for predicting RFS probabilities: Clinical, Tumor, SM, AT, TC, SM-TC, and SMAT-TC models. The Clinical model

included significant clinical risk factors identified through multivariable Cox regression. The SM, AT, and Tumor models were developed using two-dimensional (2D) SM images at the L3 level, 2D AT images at the L3 level, and three-dimensional (3D) tumor images, respectively. The TC model integrated tumor imaging signatures with significant clinical risk factors. The SM-TC model combined tumor and SM imaging signatures with significant clinical risk factors. The SMAT-TC model further incorporated all imaging signatures (i.e., tumor, SM, and AT) with significant clinical risk factors. Details of the imaging-based prediction models' construction were as follows:

- (1) SM model: The SM model processed 2D venous-phase CT images at the L3 level with corresponding SM masks. The input images were first segmented into non-overlapping patches, which were then linearly projected into 512-dimensional embedding vectors. Learnable positional encodings were incorporated to preserve spatial relationships within the image. The embedded patches passed through a series of four Transformer layers, each comprising an MSA mechanism followed by an MLP block with Gaussian Error Linear Unit activation. Layer normalization was applied before each sub-component, and residual connections ensured stable gradient flow during training.
- (2) AT model: The AT model maintained an identical architectural configuration to the SM model, processing L3-level CT images with AT masks while employing separate, non-shared parameters to capture modality-specific feature representations. This parallel processing strategy allowed each stream to develop specialized feature detectors tailored to its respective tissue type.
- (3) Tumor model: For volumetric tumor analysis, the Tumor model processed consecutive CT slices encompassing the entire tumor volume. The architecture adapted the standard Transformer framework to handle 3D data through volumetric patch extraction and 3D positional encoding. While maintaining the same fundamental layer configuration as the 2D models, the tumor stream incorporated 3D-aware operations that captured spatial relationships across multiple slices, enabling comprehensive tumor characterization.

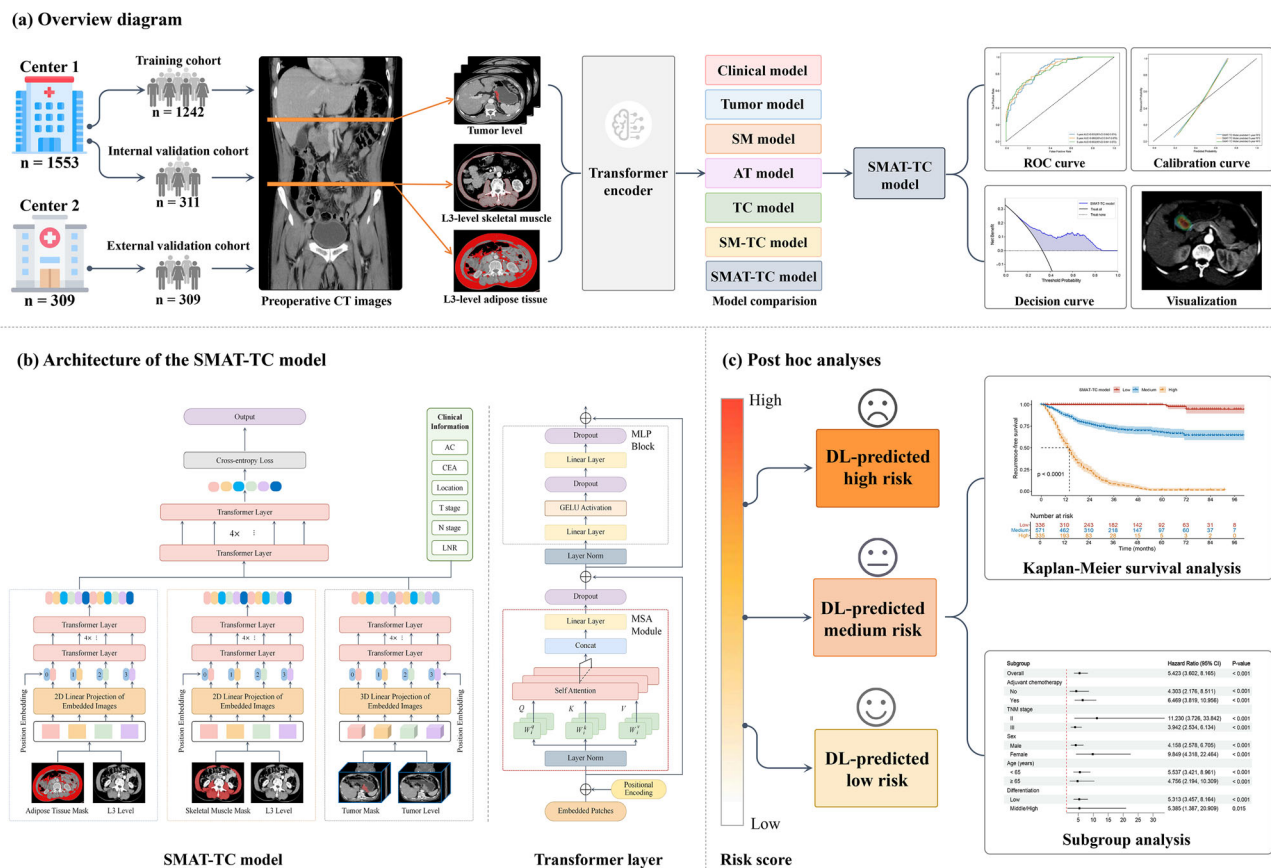


**Fig. 5 | Overview of the workflow of this study.** AC adjuvant chemotherapy, AT adipose tissue, CEA carcinoembryonic antigen, DL deep learning, LNR lymph node ratio, L3 the third lumbar vertebra, MLP multi-layer perceptron, MSA multi-head

self-attention, ROC receiver operating characteristic, SM skeletal muscle, SM-TC SM-tumor-clinical, SMAT-TC SM-AT-tumor-clinical, TC tumor-clinical, 2D two-dimensional, 3D three-dimensional.

- TC model: Building upon the Tumor framework, the TC model integrated tumor features with significant clinical risk factors, including CEA level, tumor location, T stage, N stage, LNR, and AC. The clinical variables were embedded via a linear projection layer. The resulting feature representations from various modalities were concatenated and processed via an additional series of four Transformer layers that learned cross-modal interactions and captured complementary prognostic information.
- SM-TC model: The SM-TC model extended the analytical scope beyond the tumor by integrating SM, alongside tumor imaging and clinical factors. The model used separate Transformer encoders to process SM and tumor imaging data, thereby retaining their modality-specific characteristics. Similar to the TC model, the features from three streams—SM, tumor, and embedded clinical variables—were concatenated and fused through a subsequent series of four Transformer layers.
- SMAT-TC model: Building on the SM-TC framework, the SMAT-TC model introduces AT as a third imaging modality. The expanded architecture processed SM, AT, and tumor regions through dedicated Transformer encoders that operated in parallel, preserving the unique characteristics of each data source. The model then integrated all imaging-derived features with clinical risk factors through a unified Transformer-based fusion module that implemented cross-modal attention mechanisms, enabling the network to learn synergistic relationships between different data types and ultimately enhancing predictive performance for recurrence risk assessment.

During training, model predictions were evaluated using a cross-entropy loss function to measure the discrepancy between outputs and ground-truth labels. Model parameters were iteratively updated via back-propagation using an AdamW optimizer to progressively minimize the loss function. Through this optimization process, the model's predictions gradually converged toward the ground-truth labels, yielding optimized model parameters. The final model output provided quantitative recurrence risk scores for each GC patient, enabling stratification of patients based on their recurrence probability.

The performance of various models was compared using the C-index. The incremental value of adding body composition was quantified using the NRI and IDI. The model achieving the highest C-index was selected for subsequent analyses.

### Predictive performance evaluation of the optimal model

The predictive ability of the optimal model was evaluated using time-dependent ROC analysis, calibration curve, and DCA. Patients were stratified into high-, medium-, or low-risk groups using the IQR of predictive scores from the training cohort. Specifically, patients with risk scores above the third quartile (Q3) were classified as high-risk, those with scores between the first quartile (Q1) and Q3 as medium-risk, and those with scores below Q1 as low-risk. Survival outcomes among the three risk groups were compared using Kaplan-Meier curves with the Log-rank test. Post hoc analyses examined the relationships between risk groups across subgroups stratified by sex, age, TNM stage, differentiation grade, and AC.

## Statistical analysis

Statistical analyses were conducted using Python (version 3.8), SPSS (version 26.0), and MedCalc (version 20.21). Continuous variables were assessed for normality using the Shapiro–Wilk test. Variables following a normal distribution were presented with mean  $\pm$  standard deviation (SD), while non-normally distributed variables were presented with median (IQR). Continuous variables were compared using the one-way analysis of variance or the Kruskal–Wallis test as appropriate, and categorical variables were compared using the Chi-square test. Calibration curves, along with the Hosmer–Lemeshow test, were used to compare the predicted survival probabilities with the actual probabilities. DCA was performed to evaluate the net benefit associated with the optimal model at various threshold probabilities. The association between the model-derived risk score and GC prognosis was assessed using univariate and multivariable Cox regression analyses. The variables that were significant in the univariate Cox regression analysis ( $P < 0.05$ ) were included in the multivariable Cox regression analysis. A two-tailed  $P$ -value  $< 0.05$  was considered statistically significant.

## Data availability

The imaging data used in this study are protected and not publicly available due to personal information protection, patient privacy regulations, and institutional data-sharing policies. However, the data can be made available to qualified researchers for academic purposes upon reasonable request. Interested parties may contact the corresponding author, who will review each request for scientific merit and compliance with data governance requirements. The source codes used in this study are available at <https://github.com/MedAI-Lab/SMAT-TC>, which can only be used for non-commercial research purposes.

Received: 15 August 2025; Accepted: 18 November 2025;

Published online: 03 December 2025

## References

- Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **74**, 229–263 (2024).
- Muro, K. et al. Pan-Asian adapted ESMO Clinical Practice Guidelines for the management of patients with metastatic gastric cancer: a JSMO-ESMO initiative endorsed by CSCO, KSMO, MOS, SSO and TOS. *Ann. Oncol.* **30**, 19–33 (2019).
- Pentheroudakis, G. & Committee, E. G. Recent eUpdates to the ESMO Clinical Practice Guidelines on hepatocellular carcinoma, cancer of the pancreas, soft tissue and visceral sarcomas, cancer of the prostate and gastric cancer. *Ann. Oncol.* **30**, 1395–1397 (2019).
- Amin, M. B. 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.* **67**, 93–99 (2017).
- Bang, Y. J. et al. Adjuvant capecitabine and oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): a phase 3 open-label, randomised controlled trial. *Lancet* **379**, 315–321 (2012).
- Dong, Q. T. et al. Influence of body composition, muscle strength, and physical performance on the postoperative complications and survival after radical gastrectomy for gastric cancer: A comprehensive analysis from a large-scale prospective study. *Clin. Nutr.* **40**, 3360–3369 (2021).
- Arends, J. et al. ESPEN guidelines on nutrition in cancer patients. *Clin. Nutr.* **36**, 11–48 (2017).
- Lim, H. S., Lee, B., Cho, I. & Cho, G. S. Nutritional and Clinical Factors Affecting Weight and Fat-Free Mass Loss after Gastrectomy in Patients with Gastric Cancer. *Nutrients* **12**, <https://doi.org/10.3390/nu12071905> (2020).
- White, J. V. et al. Consensus statement of the Academy of Nutrition and Dietetics/American Society for Parenteral and Enteral Nutrition: characteristics recommended for the identification and documentation of adult malnutrition (undernutrition). *J. Acad. Nutr. Diet.* **112**, 730–738 (2012).
- Magudia, K. et al. Population-Scale CT-based Body Composition Analysis of a Large Outpatient Population Using Deep Learning to Derive Age-, Sex-, and Race-specific Reference Curves. *Radiology* **298**, 319–329 (2021).
- Li, C. et al. Deep learning-based AI model for signet-ring cell carcinoma diagnosis and chemotherapy response prediction in gastric cancer. *Med. Phys.* **49**, 1535–1546 (2022).
- Jin, C. et al. Deep learning analysis of the primary tumour and the prediction of lymph node metastases in gastric cancer. *Br. J. Surg.* **108**, 542–549 (2021).
- Martin, L. et al. Cancer Cachexia in the Age of Obesity: Skeletal Muscle Depletion Is a Powerful Prognostic Factor, Independent of Body Mass Index. *J. Clin. Oncol.* **31**, 1539–1547 (2013).
- Prado, C. M. et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol.* **9**, 629–635 (2008).
- Miao, S. et al. Deep learning radiomics under multimodality explore association between muscle/fat and metastasis and survival in breast cancer patients. *Brief Bioinform.* **23**, <https://doi.org/10.1093/bib/bbac432> (2022).
- Jiang, Y. et al. Noninvasive Prediction of Occult Peritoneal Metastasis in Gastric Cancer Using Deep Learning. *JAMA Netw. Open* **4**, e2032269 (2021).
- Zhang, J. et al. Deep learning predicts resistance to neoadjuvant chemotherapy for locally advanced gastric cancer: a multicenter study. *Gastric Cancer* **25**, 1050–1059 (2022).
- Vaswani, A. et al. Attention is all you need. In *31st Conference on Neural Information Processing Systems (NIPS 2017)*.
- Wang, R. S. et al. Medical image segmentation using deep learning: A survey. *IET Image Process.* **16**, 1243–1267 (2022).
- Liu, Z. et al. A ConvNet for the 2020s. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 11966–11976 (IEEE, 2022).
- Chen, J. Y. et al. TransMorph: Transformer for unsupervised medical image registration. *Med. Image Anal.* **82**, <https://doi.org/10.1016/j.media.2022.102615> (2022).
- Hatamizadeh, A. et al. UNETR: Transformers for 3D Medical Image Segmentation. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 1748–1758 (IEEE, 2022).
- Zhang, Z. C., Yu, L. Q., Liang, X. K., Zhao, W. & Xing, L. TransCT: Dual-Path Transformer for Low Dose Computed Tomography. *Lect. Notes Comput. Sc.* **12906**, 55–64 (2021).
- Mourtzakis, M. et al. A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl. Physiol. Nutr. Metab.* **33**, 997–1006 (2008).
- Davis, M. P. & Panikkar, R. Sarcopenia associated with chemotherapy and targeted agents for cancer therapy. *Ann. Palliat. Med.* **8**, 86–101 (2019).
- Huang, D. D. et al. Sarcopenia predicts 1-year mortality in elderly patients undergoing curative gastrectomy for gastric cancer: a prospective study. *J. Cancer Res. Clin. Oncol.* **142**, 2347–2356 (2016).
- Kawamura, T. et al. Long-Term Outcomes of Gastric Cancer Patients with Preoperative Sarcopenia. *Ann. Surg. Oncol.* **25**, 1625–1632 (2018).
- Kudou, K. et al. Prognostic Significance of Sarcopenia in Patients with Esophagogastric Junction Cancer or Upper Gastric Cancer. *Ann. Surg. Oncol.* **24**, 1804–1810 (2017).
- Lu, J. et al. A Novel Preoperative Skeletal Muscle Measure as a Predictor of Postoperative Complications, Long-Term Survival and Tumor Recurrence for Patients with Gastric Cancer After Radical Gastrectomy. *Ann. Surg. Oncol.* **25**, 439–448 (2018).

30. Zhuang, C. L. et al. Sarcopenia is an Independent Predictor of Severe Postoperative Complications and Long-Term Survival After Radical Gastrectomy for Gastric Cancer Analysis from a Large-Scale Cohort. *Medicine* **95**, <https://doi.org/10.1097/MD.0000000000003164> (2016).
31. Lan, Q. et al. Radiomics in Addition to Computed Tomography-Based Body Composition Nomogram May Improve the Prediction of Postoperative Complications in Gastric Cancer Patients. *Ann. Nutr. Metab.* **78**, 316–327 (2022).
32. Anconina, R. et al. Combined 18 F-FDG PET/CT Radiomics and Sarcopenia Score in Predicting Relapse-Free Survival and Overall Survival in Patients With Esophagogastric Cancer. *Clin. Nucl. Med.* **47**, 684–691 (2022).
33. Chen, Q. et al. Transformer-based skeletal muscle deep-learning model for survival prediction in gastric cancer patients after curative resection. *Gastric Cancer* **28**, 684–695 (2025).
34. Li, J. et al. Transforming medical imaging with Transformers? A comparative review of key properties, current progresses, and future perspectives. *Med. Image Anal.* **85**, 102762 (2023).

## Acknowledgements

We acknowledge financial support from the National Key Research and Development Program of China (2023YFF1204600); National Natural Science Foundation of China (82402262, 82227802); Fundamental Research Funds for the Central Universities (21624304); Science and Technology Projects in Guangzhou (2025A04J3480, 2025A04J7006); Outstanding Young Talents of Guangdong Special Support Program (Health Commission of Guangdong Province) (0720240213); Clinical Frontier Technology Program of the First Affiliated Hospital of Jinan University (No. JNU1AF-CFTP-2022-a01201); Science and Technology Projects in Guangzhou (202201020022, 2023A03J1036, 2023A03J1038).

## Author contributions

Conceptualization, Methodology, Writing-Original Draft: Q.Y.C., H.X., Y.Y.L., L.J., and L.Z. Investigation: B.L. and X.W.W. Software, Resources: J.J.Y., Z.J., and H.S. Analysis and interpretation of data: W.L.H. and J.S. Funding acquisition: Q.Y.C., S.X.Z., and B.Z. Writing-Review & Editing, Supervision: S.X.Z. and B.Z.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41746-025-02183-z>.

**Correspondence** and requests for materials should be addressed to Shuixing Zhang or Bin Zhang.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

© The Author(s) 2025