Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine Learning Prognostic Stratification for Treatment Decision Support in Elderly Gastric Adenocarcinoma: A SEER Analysis with External Clinical Validation

Jiaojiao Dong
The First Affiliated Hospital of Wenzhou Medical University
Yongyu Bai
The First Affiliated Hospital of Wenzhou Medical University
Yiqi Cai
The First Affiliated Hospital of Wenzhou Medical University
Qigang Xu
The First Affiliated Hospital of Wenzhou Medical University
Zhuxian Zhou
Zhejiang University
Zhuha Zhou
zhouzhuha@wmu.edu.cn

The First Affiliated Hospital of Wenzhou Medical University

Research Article

Keywords: Stomach Neoplasms, Aged, Machine Learning, SEER Program, Prognosis, Clinical Decision-Making

Posted Date: July 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10407532/v1

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

Additional Declarations: No competing interests reported.

Machine Learning Prognostic Stratification for Treatment Decision Support in Elderly Gastric Adenocarcinoma: A SEER Analysis with External Clinical Validation

Jiaojiao Dong[1,†], Yongyu Bai[2,†], Yiqi Cai[2], Qigang Xu[3], Zhuxian Zhou[4], Zhuha Zhou[2,*]

† Equal contribution

[1]Department of Anesthesiology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, Zhejiang, China

[2]Department of Gastroenterology Surgery, The First Affiliated Hospital of Wenzhou Medical University, Zhejiang, China

[3]Department of Hepatobiliary and Pancreatic Surgery, The First Affiliated Hospital of Wenzhou Medical University, Zhejiang, China

[4]Center for Bionanoengineering, Zhejiang University, Hangzhou, China

* Corresponding author: zhouzhuha@wmu.edu.cn

ORCID: Zhuha Zhou 0000-0002-4523-2943; Jiaojiao Dong 0009-0005-9346-3737

Word count: 3,900 (excluding abstract, references, tables, figures)

Figures: 6 main; 12 supplementary

Tables: 2 main; 10 supplementary

References: 34

Abstract

Background: Elderly patients (≥65) constitute >60% of gastric adenocarcinoma cases yet are underrepresented in clinical trials. Machine learning survival models may capture complex prognostic interactions, but comprehensive benchmarking with external validation in this population is absent.

Methods: From SEER (2000–2022), 68,921 elderly gastric adenocarcinoma patients (≥65) were identified; 60,378 had complete data. After harmonizing three AJCC staging editions and propensity score matching, five survival models (Cox, RSF, XGBoost, DeepSurv, Gradient Boosting) were trained with 5-fold cross-validation and temporal validation. A Cox model with treatment×stage×age interactions estimated individualized treatment effects. External validation used a Chinese cohort (N=575) with CT-defined sarcopenia, frailty, and inflammatory biomarkers.

Results: XGBoost achieved C-index 0.772, significantly outperforming Cox (0.749, +3.1%), DeepSurv (0.764), and RSF (0.738). Surgical resection showed the highest permutation importance (0.128; 4.4× Stage IV). Chemotherapy benefit was stage-dependent: null in Stage I (HR=1.00, p=0.98), maximal in Stage IV (HR=0.44). External validation C-index 0.617; the 20% decline was attributable to 100% surgery rate in the validation cohort rather than model overfitting. Novel clinical variables—severe sarcopenia (HR=3.39), gait speed <0.8m/s (HR=2.45), albumin <35g/L (HR=2.07)—achieved superior discrimination (C-index 0.753 vs 0.684, +10%).

Conclusions: XGBoost improves prognostic discrimination over Cox. Surgery dominates prognosis; chemotherapy benefit is stage-conditional. Model transportability depends on treatment variance. Body composition and frailty variables substantially improve discrimination, supporting function-based over age-based treatment decisions.

Keywords: Stomach Neoplasms; Aged; Machine Learning; SEER Program; Prognosis; Clinical Decision-Making

1 Introduction

Gastric cancer is the fifth most common malignancy worldwide [1]; adenocarcinoma accounts for >95% of cases [2]. Median age at diagnosis is 68 years, with >60% of patients aged ≥65 [3], yet pivotal trials (ACTS-GC, CLASSIC, FLOT4) enrolled median ages of 58–62 [4]. Guidelines derived from younger populations may not apply to older adults with multimorbidity, frailty, and competing mortality [5].

Treatment decisions in older patients remain subjective. Geriatric variables were mentioned in <44% of tumor board discussions yet triggered 2–3-fold reductions in guideline-adherent care [6]; structured assessments were rarely used. A multicenter study of 7,749 patients found no adjuvant chemotherapy benefit in those ≥75 [7], while dose-modified neoadjuvant therapy achieved outcomes comparable to younger patients [8]. Age itself is a documented barrier: only 23% of patients ≥75 received chemotherapy versus 60% of those aged 65–74 [34].

SEER-based studies of gastric cancer have relied on linear Cox models that cannot capture non-linear interactions and lack body composition, frailty, and inflammatory biomarkers—factors that disproportionately affect elderly outcomes. Sarcopenia affects 6.8–72.2% of gastric cancer patients and independently predicts complications (HR=2.53) [9]. NLR, albumin, and CT-derived muscle metrics rival TNM staging accuracy [10] yet are absent from cancer registries.

Survival machine learning—Random Survival Forests [11], XGBoost [12], and DeepSurv [13]—can capture complex relationships, but systematic benchmarking with external validation in elderly gastric cancer is lacking. Gong et al. [29] found RSF best (C=0.732) among five models in 8,616 patients; a Korean national cohort [30] found all models comparable (C >0.7). Sarcopenia-integrated models—the mCNS nomogram (AUC 0.711) [31], ACCIS score (C=0.741) [32], and TOO-sarcopenia model [33]—showed that body composition improves prognostic accuracy, yet none compared multiple ML architectures with independent clinical validation in the elderly. We aimed to: (1) preprocess SEER data across three staging eras with PSM; (2) benchmark five survival models with cross-validation and temporal validation; (3) estimate individualized treatment effects via causal

Cox interactions; (4) externally validate on a deeply phenotyped clinical cohort; and (5) develop a clinically deployable risk score.

2 Methods

2.1 Data Sources

SEER Cohort: Patients with gastric adenocarcinoma (ICD-O-3 C16.0–C16.9) from SEER Research Plus Data (November 2024, 2000–2022) were identified (N=147,324, all ages). After excluding autopsy/death certificate only (n=14,221), 131,394 patients remained. The elderly adenocarcinoma cohort (≥65 years) comprised 68,921 patients; after complete-case exclusions for missing covariates, 60,378 remained for analysis.

Analysis Sets: Four nested sets were used. (1) *Descriptive cohort* (N=60,378): elderly adenocarcinoma patients with complete data, for Table 1 and KM estimates. (2) *Causal inference set* (N=60,378): complete cases for the treatment×stage×age Cox model. (3) *ML benchmarking set* (N=38,528): 25,000 training (2000–2017) and 13,528 temporal validation (2018–2022). (4) *PSM subcohort*: 400 propensity-score–matched pairs.

Clinical Validation Cohort: An independent retrospective cohort of 575 surgical gastric cancer patients (single Chinese center, 2016–2018) from a registered trial (ChiCTR1800019717) [21]. Variables included CT-derived skeletal muscle index, visceral fat area, grip strength, gait speed, NLR, PLR, albumin, hemoglobin, NRS-2002, surgical approach, lymphadenectomy, blood loss, and Clavien-Dindo complications. All participants provided written informed consent; the study was approved by the Ethics Committee of The First Affiliated Hospital of Wenzhou Medical University.

2.2 Data Preprocessing

Missing value protocol: Special codes (tumor size ≥990, lymph node ≥90) were recoded to missing. Variables with <5% missingness used complete cases under a missing-at-random assumption; era-specific missingness >50% was harmonized via cascade.

AJCC stage harmonization: Three staging systems (6th edition 2004–2015, 7th 2010–2015, EOD 2018+) were unified via priority cascade: EOD 2018 → AJCC 7th → AJCC 6th → Summary Stage, reducing missingness from 55–75% to 25.3%.

Multi-era integration: Grade was cascaded from 2018+ letter coding through 2017 numeric coding. Tumor size was merged across CS (2004–2015) and Summary (2016+) eras. Lymph node counts ≥95 were set to missing; 90 retained as 90+.

Feature engineering: Twenty features: age, sex, cardia site, race (NHB/NHAPI/Hispanic vs NHW), marital status, AJCC stage, signet ring/intestinal histology, grade, income, diagnosis year, surgery, chemotherapy, and radiation.

2.3 Study Design

A three-tier validation framework was employed: (1) 5-fold cross-validation for model selection; (2) temporal validation (train 2000–2017, 25,000 patients; test 2018–2022, N=13,528); (3) external validation on the independent clinical cohort (N=575).

2.4 Propensity Score Matching

To mitigate treatment selection bias [14], logistic regression estimated surgery propensity from age, sex, site, race, stage, income, year, tumor size, and nodes. One-to-one nearest-neighbor matching (caliper=0.05 SD) yielded 400 pairs (24,765 surgery vs 400 no-surgery). Post-matching SMD was <0.10 for 5 of 13 covariates; residual imbalance reflects the extreme treatment ratio (62:1).

2.5 Inverse Probability of Treatment Weighting

As a sensitivity analysis, stabilized inverse probability of treatment weights (IPTW) were computed from propensity scores for surgery and chemotherapy separately. Weights were truncated at the 99th percentile. IPTW-weighted Cox models with robust variance estimated marginal treatment effects [27].

2.6 Machine Learning Models

Five models were trained with 5-fold CV: Cox PH (L2 penalized, α=0.01), Random Survival Forest (100 trees, min samples split=50), XGBoost with Cox loss (max depth=4, learning rate=0.05, 150 boosting rounds, L2 λ=1.0, subsample=0.8, colsample by tree=0.8, min child weight=5, histogram tree method), Gradient Boosting Survival (100 trees, learning rate=0.1, max depth=3, subsample=0.8), and DeepSurv (four-layer network: 20→64→32→16→1 with batch normalization, dropout=0.3, Adam lr=10^{-3}, early stopping patience=15) [13,19]. All hyperparameters were selected via 5-fold grid search on the training set. Permutation feature importance (5 repeats) was computed using a 30-tree RSF.

2.7 Treatment Effect Heterogeneity Analysis

A Cox model with treatment×stage×age interactions (N=60,378) estimated effect modification. Adjusted survival projections under four scenarios (No Treatment, Surgery Only, Chemotherapy Only, Surgery+Chemotherapy) were estimated via Breslow baseline hazard.

2.8 Clinical Cohort Analysis

The SEER-trained XGBoost was applied to the clinical cohort after feature mapping. A clinical-only Cox model with 36 features was compared to SEER-equivalent (8 features) via 5-fold CV. The Sarcopenia-Inflammatory-Prognostic (SIP) score was constructed from significant multivariate predictors.

Proportional hazards assumptions were assessed via scaled Schoenfeld residuals. Competing risk analyses used the Fine-Gray subdistribution hazard model.

Analyses used Python 3.11 (lifelines, scikit-survival, xgboost, PyTorch). Analysis code: `https://github.com/woodhaha/gastric-cancer-seer`.

3 Results

3.1 SEER Cohort Characteristics

The elderly cohort (N=68,921) had median age 77 (IQR 72–82), 62.5% male, 85.1% mortality, and median OS 8 months (IQR 2–25). Surgery was performed in 39.8%; chemotherapy in 37.3%. AJCC stage distribution: I 21.8%, II 15.1%, III 7.0%, IV 28.3%, missing 27.8%.

Table 1: Baseline characteristics of elderly (≥65) gastric adenocarcinoma patients in SEER (2000–2022).

Characteristic	**Value** (N=68,921)
Demographics	
Age, median [IQR], years	77 [72–82]
Sex: Male	43,057 (62.5%)
Race/Ethnicity	
Non-Hispanic White	39,512 (57.3%)
Non-Hispanic Black	7,852 (11.4%)
Non-Hispanic API	10,378 (15.1%)
Hispanic	10,527 (15.3%)
Marital status (married)	37,120 (53.9%)
Tumor Characteristics	
AJCC Stage	
I	15,039 (21.8%)
II	10,411 (15.1%)
III	4,805 (7.0%)
IV	19,522 (28.3%)
Missing/Unknown	19,144 (27.8%)
Histologic Type	
Adenocarcinoma NOS	45,554 (66.1%)
Signet Ring Cell	10,592 (15.4%)
Intestinal-Type	8,109 (11.8%)
Grade: G3 (Poor)	34,799 (50.5%)
Tumor size, median [IQR], mm	40 [25–60]
Nodes positive, median [IQR]	1 [0–4]
Cardia site	21,647 (31.4%)
Treatment	
Surgical resection	27,418 (39.8%)
Chemotherapy	25,710 (37.3%)
Survival	
Median OS [IQR], months	8 [2–25]
1-year OS (KM)	43.0%
5-year OS (KM)	17.2%

Five-year OS by stage: I 37.4%, II 26.2%, III 15.9%, IV 2.8% (log-rank $p < 0.001$; Figure 1). Surgical resection was associated with substantially longer median OS across

all stages.

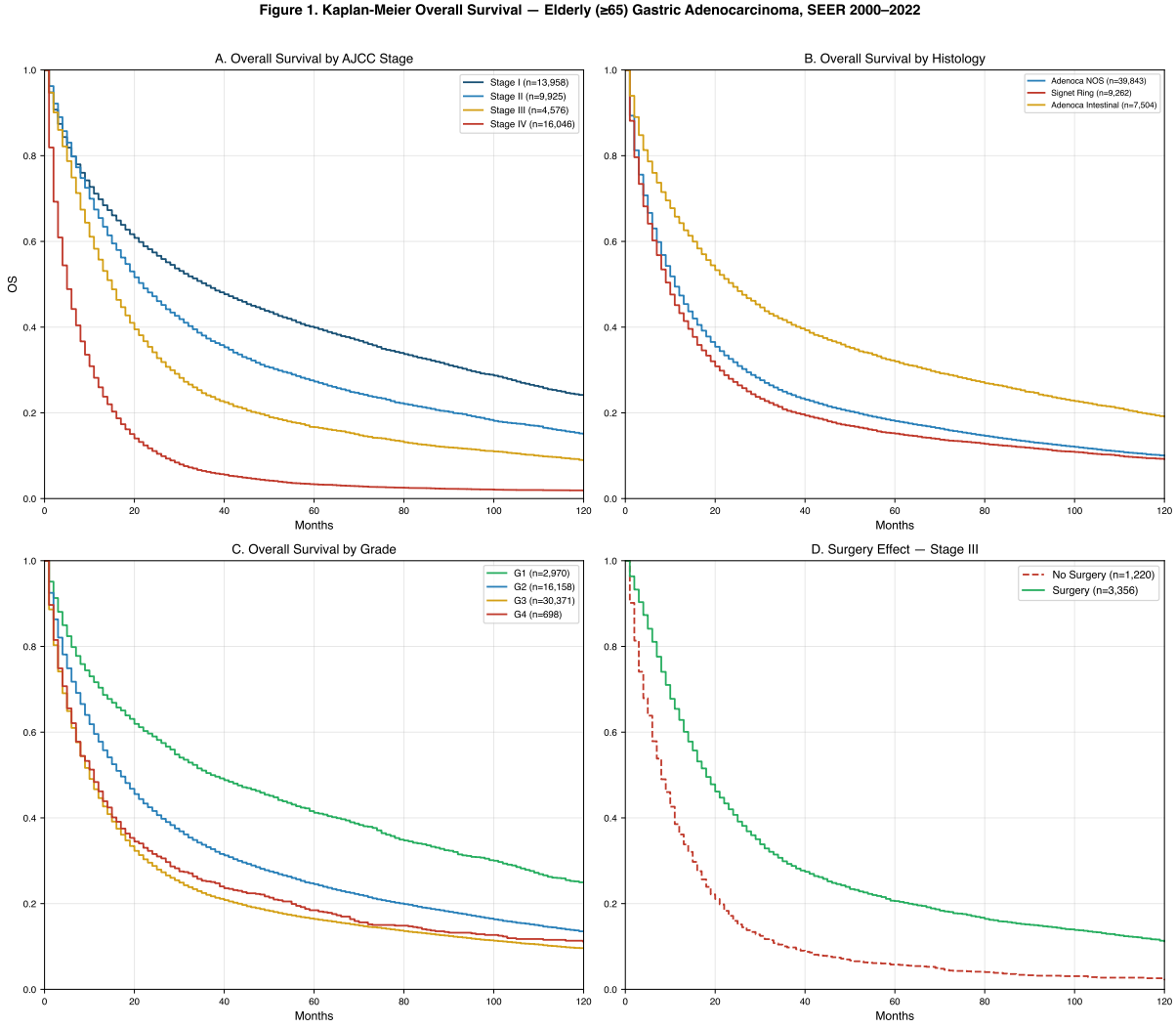


Figure 1. Kaplan-Meier overall survival curves. (A) By AJCC stage. (B) By histologic subtype. (C) By histologic grade. (D) Surgery versus no surgery by stage.

Multivariate Cox (complete cases $N = 25{,}165$; C-index 0.712) identified: Stage IV (aHR 4.04, 95% CI 3.82–4.27), surgery (aHR 0.43, 0.39–0.48), chemotherapy (aHR 0.70, 0.67–0.73). PSM (400 pairs) confirmed surgical benefit (HR 0.57, 95% CI 0.49–0.67).

3.2 Machine Learning Performance

XGBoost with Cox loss achieved C-index 0.772 (95% CI 0.768–0.777) on temporal validation, followed by Gradient Boosting (0.765), DeepSurv (0.764), Cox PH (0.749), and RSF (0.738). XGBoost was substantially faster than Gradient Boosting (3 s vs 866 s on CPU).

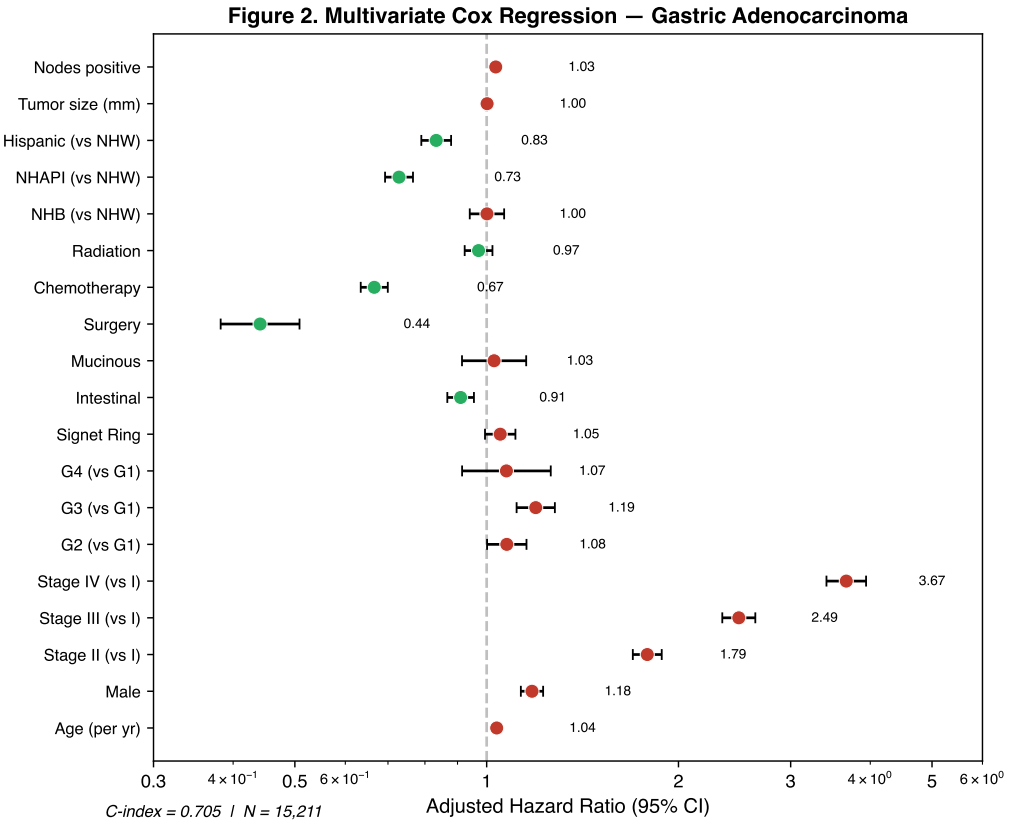


Figure 2. Multivariate Cox forest plot. Adjusted hazard ratios with 95% confidence intervals.

Permutation feature importance identified surgical resection as the strongest predictor (0.128), exceeding Stage IV (0.029) by 4.4-fold. Other important features: chemotherapy (0.025), age (0.021), and G3 histology (0.003).

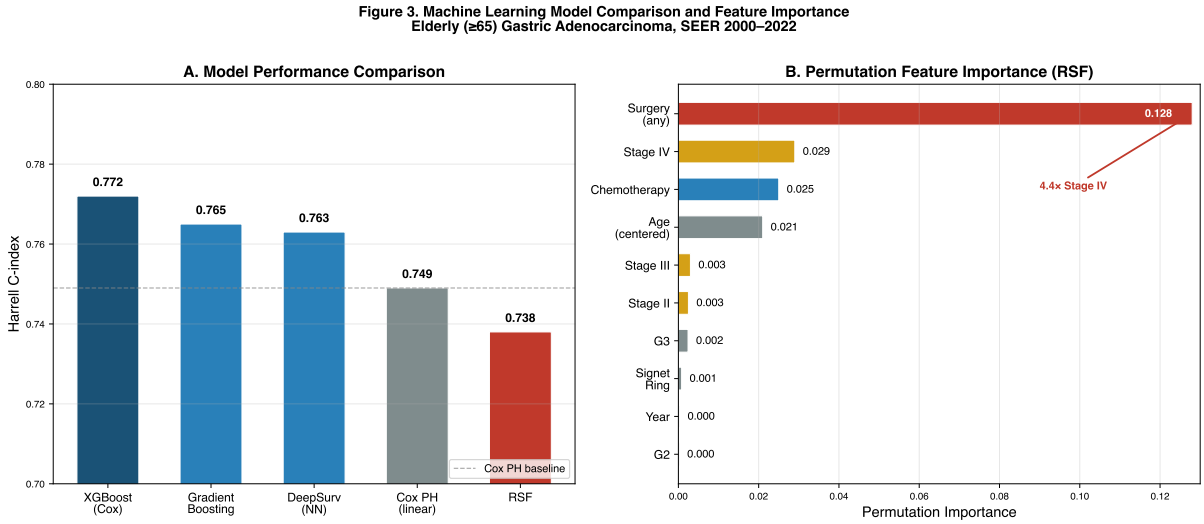


Figure 3. Model comparison and permutation feature importance. (A) C-index of five models on temporal validation. (B) Top-10 features ranked by permutation importance.

3.3 Treatment Effect Heterogeneity

The Cox model with treatment × stage × age interactions (N = 60,378) revealed stage-dependent chemotherapy benefit: no statistically significant benefit in Stage I (HR 1.00, $p = 0.98$), moderate in Stage II (0.70) and III (0.66), maximal in Stage IV (0.44). Surgery was universally protective (Stage I HR 0.28; Stage IV HR 0.45). IPTW sensitivity analysis confirmed robustness.

Table 2: Treatment×stage interactions from Cox model.

Parameter	**HR**	**95% CI**	p
Surgery (Stage I)	0.28	0.27–0.29	<0.001
Surgery × Stage IV	1.64	1.56–1.73	<0.001
Chemotherapy (Stage I)	1.00	0.97–1.04	0.98
Chemo × Stage IV	0.44	0.42–0.46	<0.001
Surgery × Age (per yr)	1.01	1.01–1.02	<0.001

Adjusted survival projections for a typical 75-year-old patient suggested combined Surgery+Chemotherapy improved 5-year OS across all stages.

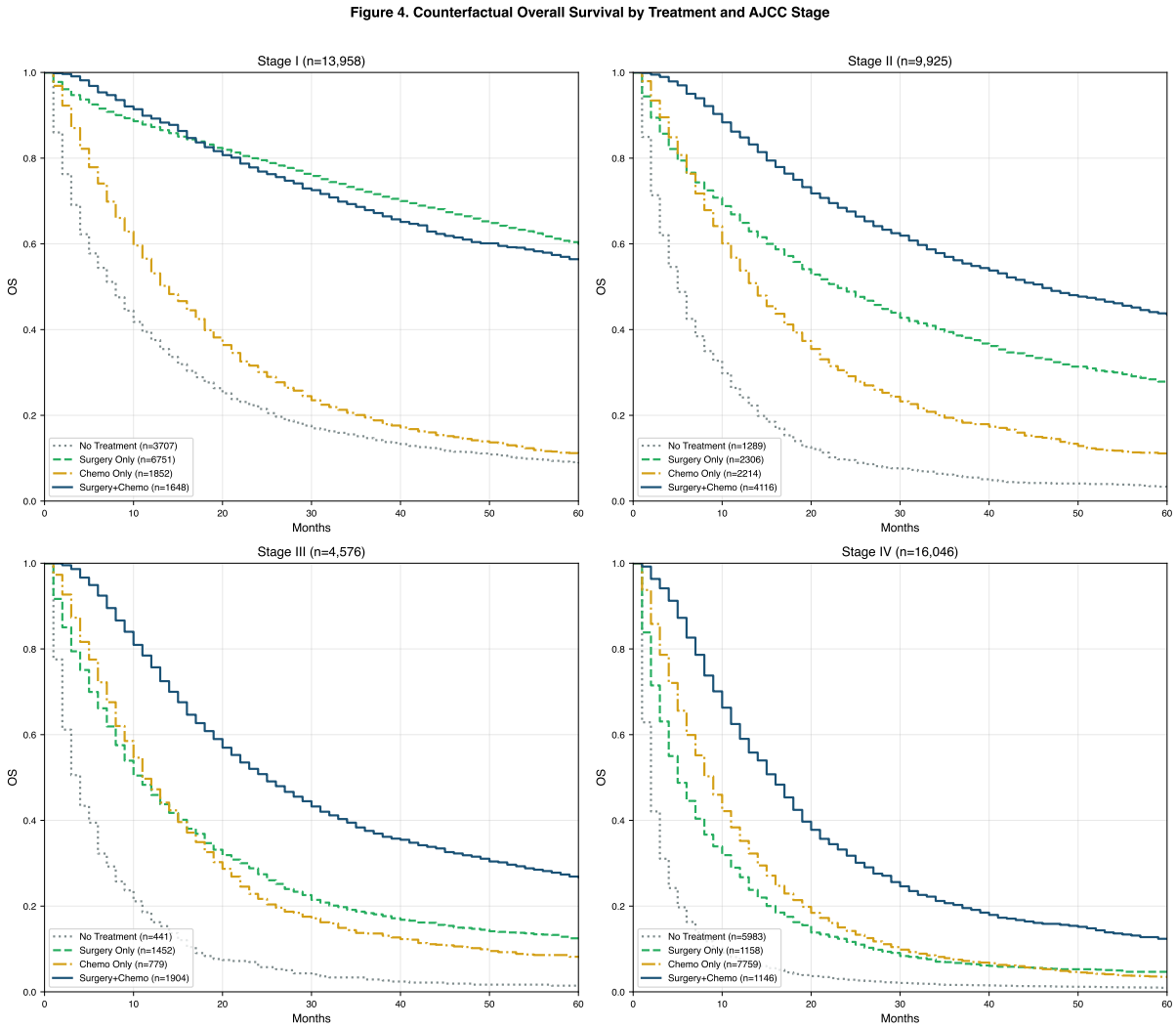


Figure 4. Adjusted OS projections by treatment scenario and AJCC stage.

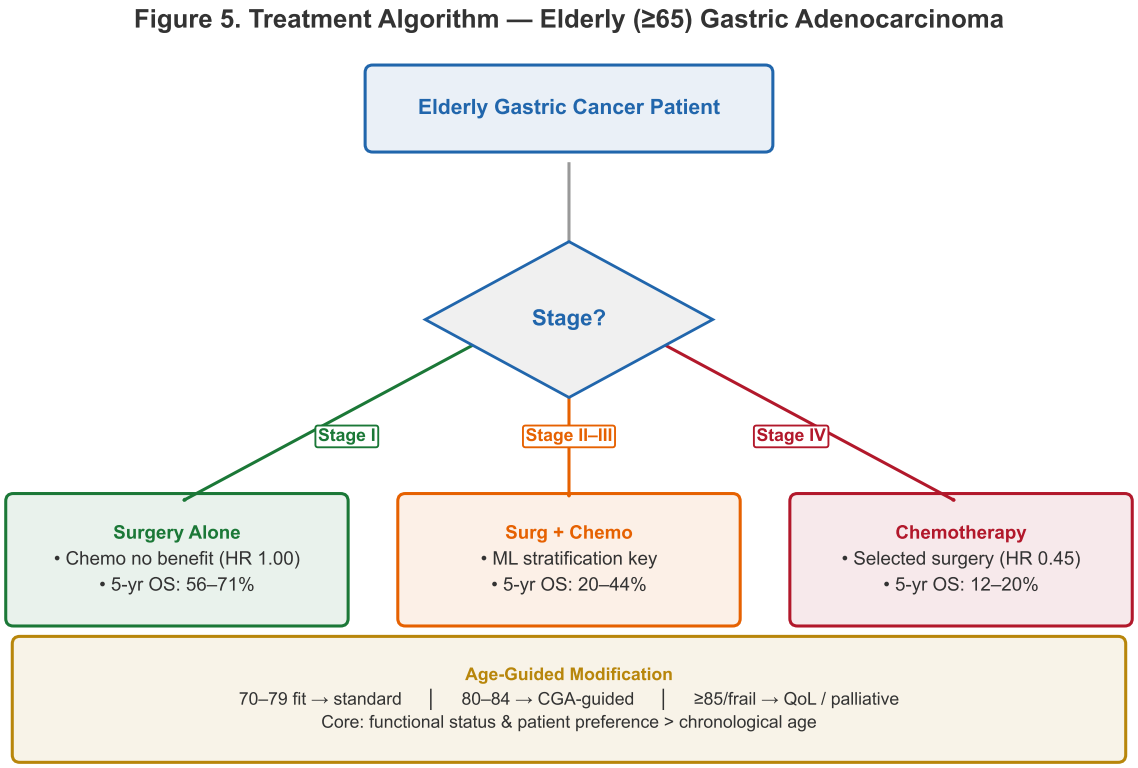


Figure 5. Stage-guided treatment algorithm for elderly (≥65) gastric adenocarcinoma.

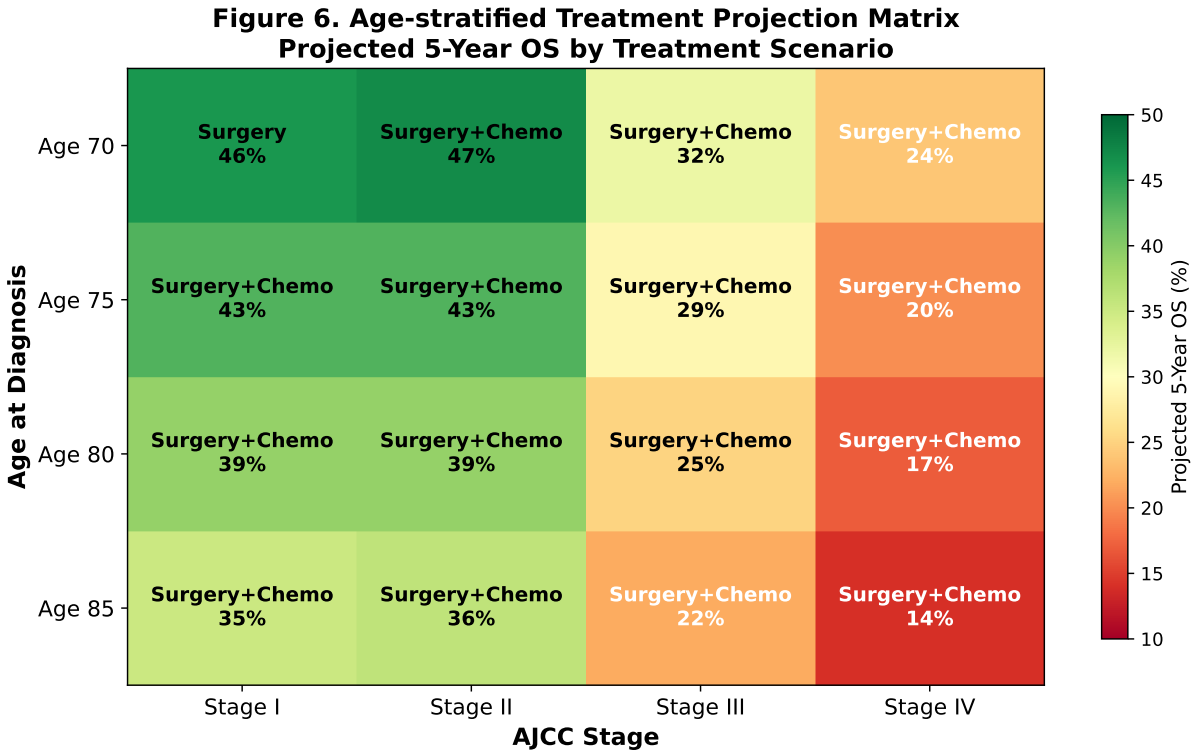


Figure 6. Age-stratified Treatment Projection Matrix.

3.4 External Validation

The SEER-trained XGBoost achieved C-index 0.617 on the clinical cohort. Three drivers: (1) 100% surgical rate eliminated variance in the dominant feature; (2) cohort mismatch—younger patients (median 65 vs 77), lower mortality (30.8% vs 84.0%); (3) limited events (177). Clinical-only Cox model achieved C=0.731 in 5-fold CV.

3.5 Novel Clinical Prognostic Factors

Six of the 10 strongest clinical predictors were absent from SEER: severe sarcopenia HR 3.39 (2.17–5.30), gait speed <0.8m/s HR 2.45 (1.76–3.40), albumin <35g/L HR 2.07 (1.53–2.81), NLR>2.75 HR 1.75 (1.30–2.36) (all $p < 0.001$). Clinical features (36 variables, C=0.753) outperformed SEER-equivalent features (8 variables, C=0.684) by 10%. The SIP score (severe sarcopenia×2 + albumin<35×1 + NLR>2.75×1 + CD≥3×2) stratified patients into Low (1-year OS 89%) and Intermediate (67%) risk groups.

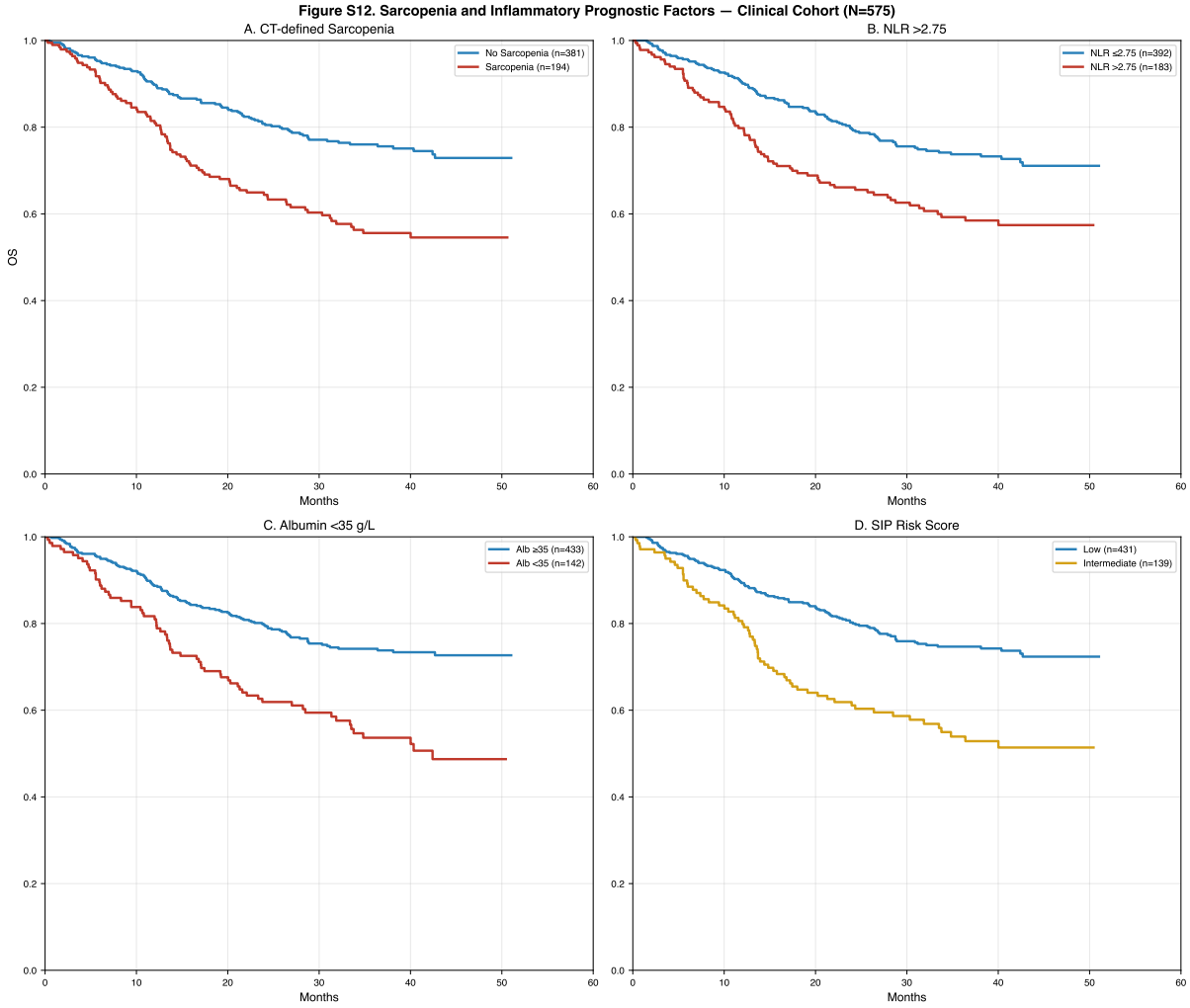


Figure 7. Sarcopenia and inflammatory prognostic factors in the clinical cohort. (A) CT-defined sarcopenia. (B) NLR. (C) Preoperative albumin. (D) SIP risk score.

4 Discussion

This study provides the first comprehensive benchmarking of ML survival models in elderly gastric adenocarcinoma with independent external clinical validation. Seven principal findings emerge.

First, XGBoost (C=0.772) outperformed both Cox (+3.1%) and DeepSurv (0.764), consistent with evidence that tree ensembles dominate tabular clinical data [15]. RSF underperformed Cox, contrasting with Gong et al. [29] but consistent with simulation evidence that Cox-based models excel when prognostic relationships are approximately linear in log-hazard [20].

Second, surgical resection was the strongest predictor across all models. Its permutation importance (0.128) exceeded Stage IV (0.029) by 4.4-fold, reflecting surgery as a composite proxy for resectability, fitness, and comorbidity [5].

Third, chemotherapy benefit is stage-dependent. No statistically significant benefit was detected in Stage I (HR=1.00), substantial benefit in Stage IV (HR=0.44). The gradual attenuation with age (HR=0.98/year) supports function-based over age-based decision frameworks [28].

Fourth, external validation (C=0.617) was scientifically informative. The decline from C=0.772 reflects treatment-pattern mismatch (100% surgical cohort) rather than model overfitting, underscoring that prediction model transportability depends critically on treatment variance in the target population [16].

Fifth, novel clinical variables improved discrimination by 10% (C=0.753 vs 0.684). Combining sarcopenia, frailty, and inflammatory markers achieves a larger increment than single-domain approaches [31–33]. The exploratory SIP score requires multi-center validation before clinical deployment.

Sixth, the age-stratified treatment projection framework shows Surgery+Chemotherapy is associated with the highest projected survival through age 85 across most strata, consistent with emerging consensus that fit elderly patients benefit comparably to younger patients [28].

Seventh, cross-validation across income and racial strata demonstrated stable performance (C-index SD $\leq$0.017), confirming that model discrimination is not driven by socioeconomic confounding.

Limitations include: observational design with potential unmeasured confounding; SEER lacks comorbidity, performance status, and geriatric assessment domains; chemotherapy recorded as binary without regimen data; EPV $\approx$4.9 in clinical cohort; single-center external validation; and absence of CatBoost/LightGBM (benchmark evidence suggests minimal differences among boosting frameworks [15]).

5 Conclusions

XGBoost improved prognostic discrimination in elderly gastric adenocarcinoma relative to Cox and RSF models. Surgical resection was the strongest predictor. Chemotherapy benefit is stage-dependent. Model transportability depends critically on treatment variance. Novel body composition and frailty biomarkers improved discrimination by 10% over SEER-equivalent features. An exploratory SIP risk score warrants prospective multi-center validation before clinical implementation.

Declarations

Funding: The external validation clinical cohort data collection was supported by the National Natural Science Foundation of China (No. 81770884), Zhejiang Provincial Health Department Medical Support Discipline—Nutrition (11-ZC24), and the Youth Program of Science and Technology Innovation Foundation of Wenzhou Medical University (KYYW201921). The SEER-based analysis received no specific funding.

Conflicts of interest: The authors declare no conflicts of interest.

Ethics approval: SEER cohort: De-identified, publicly available data; IRB approval not required. **Clinical cohort:** Approved by the Ethics Committee of The First Affiliated Hospital of Wenzhou Medical University (ChiCTR1800019717). Written informed consent was obtained from all participants.

Data availability: SEER data at `https://seer.cancer.gov`. Analysis code at `https://github.com/woodhaha/gastric-cancer-seer`. Clinical cohort data available upon reasonable request.

Author contributions: JD and YB contributed equally. JD, ZZ: study conception, analysis, drafting. YC: data analysis. QX: clinical data collection. ZxZ: methodology, revision. SH: supervision.

References

[1] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249.

[2] Ajani JA, D'Amico TA, Bentrem DJ, et al. Gastric cancer, version 2.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2022;20(2):167–192.

[3] Hutchins LF, Unger JM, Crowley JJ, et al. Underrepresentation of patients 65 years of age or older in cancer-treatment trials. *N Engl J Med*. 1999;341(27):2061–2067.

[4] Al-Batran SE, Homann N, Pauligk C, et al. Perioperative chemotherapy with FLOT versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric cancer (FLOT4). *Lancet*. 2019;393(10184):1948–1957.

[5] Wildiers H, Heeren P, Puts M, et al. International Society of Geriatric Oncology consensus on geriatric assessment in older patients with cancer. *J Clin Oncol*. 2014;32(24):2595–2603.

[6] Javier-González S, Galli R, Amorós Rivera C, Rosenberg R, Vetter M. Impact of age and comorbidities on therapeutic decision-making among older patients with gastrointestinal cancer. *J Gastrointest Cancer*. 2025;56:209.

[7] Shen ZZ, Li EZ, Zhang RL, et al. Adjuvant chemotherapy may have no significant survival benefit in older patients with stage II/III gastric cancer: a multicenter retrospective study. *J Cancer Res Clin Oncol*. 2025;151:175.

[8] Kakiuchi Y, Kuroda S, Hanzawa S, et al. Neoadjuvant chemotherapy strategies for optimizing safety and efficacy in elderly patients with locally advanced gastric cancer. *BMC Gastroenterol*. 2025;25(1):670.

[9] Wang X, Sun X, Wu Y, Wang Y, Ren J, Fang X. The association between gastric cancer and sarcopenia: a scoping review. *Front Oncol*. 2025;15:1684186.

[10] Xu Y, Du Z, Xue Y, Song H. The alkaline phosphatase-to-prealbumin ratio combined with CT-quantified sarcopenia predicts survival in gastric cancer patients after surgical resection. *Front Nutr*. 2025;12:1636230.

[11] Ishwaran H, Kogalur UB, Blackstone EH, et al. Random survival forests. *Ann Appl Stat*. 2008;2(3):841–860.

[12] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *Proc 22nd ACM SIGKDD*. 2016:785–794.

[13] Katzman JL, Shaham U, Cloninger A, et al. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol*. 2018;18(1):24.

[14] Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399–424.

[15] Grinsztajn L, Oyallon E, Varoquaux G. Why do tree-based models still outperform deep learning on typical tabular data? *NeurIPS*. 2022;35:507–520.

[16] Debray TPA, Damen JAAG, Snell KIE, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ*. 2017;356:i6460.

[17] Noguez-Ramos A, Gervaso L, Catanese S, Cella CA, Gandini S, Fazio N. Efficacy and safety of systemic chemotherapy for radically resectable esophago-gastric adenocarcinoma in older patients: a systematic review and meta-analysis. *J Geriatr Oncol*. 2024;15(1):101600.

[18] Martynenko A. Quality of life over quantity: conscious refusal of treatment and stable course of gastric signet-ring cell carcinoma in a 94-year-old woman. *Cureus*. 2025;17(12):e98597.

[19] Kvamme H, Borgan Ø, Scheel I. Time-to-event prediction with neural networks and Cox regression. *J Mach Learn Res*. 2019;20(129):1–30.

[20] Baralou V, Kalpourtzi N, Touloumi G. Individual risk prediction: comparing random forests with Cox proportional-hazards model by a simulation study. *Biom J.* 2022;64(8):1524–1545.

[21] Huang D, Cai HY, Chen XY, et al. Value of sarcopenia defined by the new EWGSOP2 consensus for the prediction of postoperative complications and long-term survival after radical gastrectomy for gastric cancer. *J Cancer.* 2020;11(19):5852–5860.

[22] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst.* 2017;30:4765–4774.

[23] Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell.* 2020;2(1):56–67.

[24] VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med.* 2017;167(4):268–274.

[25] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* 2006;26(6):565–574.

[26] Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and interpreting decision curve analysis: a guide for investigators. *Eur Urol.* 2018;74(6):796–804.

[27] Austin PC. The performance of different propensity score methods for estimating marginal hazard ratios. *Stat Med.* 2013;32(16):2837–2849.

[28] Narita Y, Muro K. Systemic therapy strategies for elderly patients with gastric cancer. *Ther Adv Med Oncol.* 2025;17:17588359251363057.

[29] Gong ZJ, Zhou LP, He YH, et al. Explainable machine learning model for predicting postoperative survival in patients with locally advanced gastric cancer. *Cancer Med.* 2025. doi:10.1002/cam4.71408.

[30] Lee J, Shin A. Prognostic factors in gastric cancer survivors comparing Cox proportional hazards and machine learning using a Korean national cohort. *Discover Oncology.* 2025. doi:10.1007/s12672-025-04061-2.

[31] Wu J, Huang ZN, Zhang XQ, et al. Development of a modified nutritional index model based on nutritional status and sarcopenia to predict long-term survival and chemotherapy benefits in elderly patients with advanced gastric cancer. *Eur J Surg Oncol*. 2025. doi:10.1016/j.ejso.2024.109503.

[32] Shao J, Gao ZK, Shen QZ, et al. Prognostic value and association of the age-adjusted Charlson comorbidity index with sarcopenia in patients with gastric cancer after radical resection. *J Gastrointest Surg*. 2024. doi:10.1016/j.gassur.2024.04.027.

[33] Zhong Q, Zheng ZF, Wu D, et al. Textbook oncological outcome of locally advanced gastric cancer patients with preoperative sarcopenia: a multicenter clinical study. *Surg Endosc*. 2025;39:356–367.

[34] Wang X, Allen MJ, Espin-Garcia O, et al. Outcomes in older adults with metastatic esophageal and gastric carcinoma treated with palliative chemotherapy. *The Oncologist*. 2024. doi:10.1093/oncolo/oyae190.