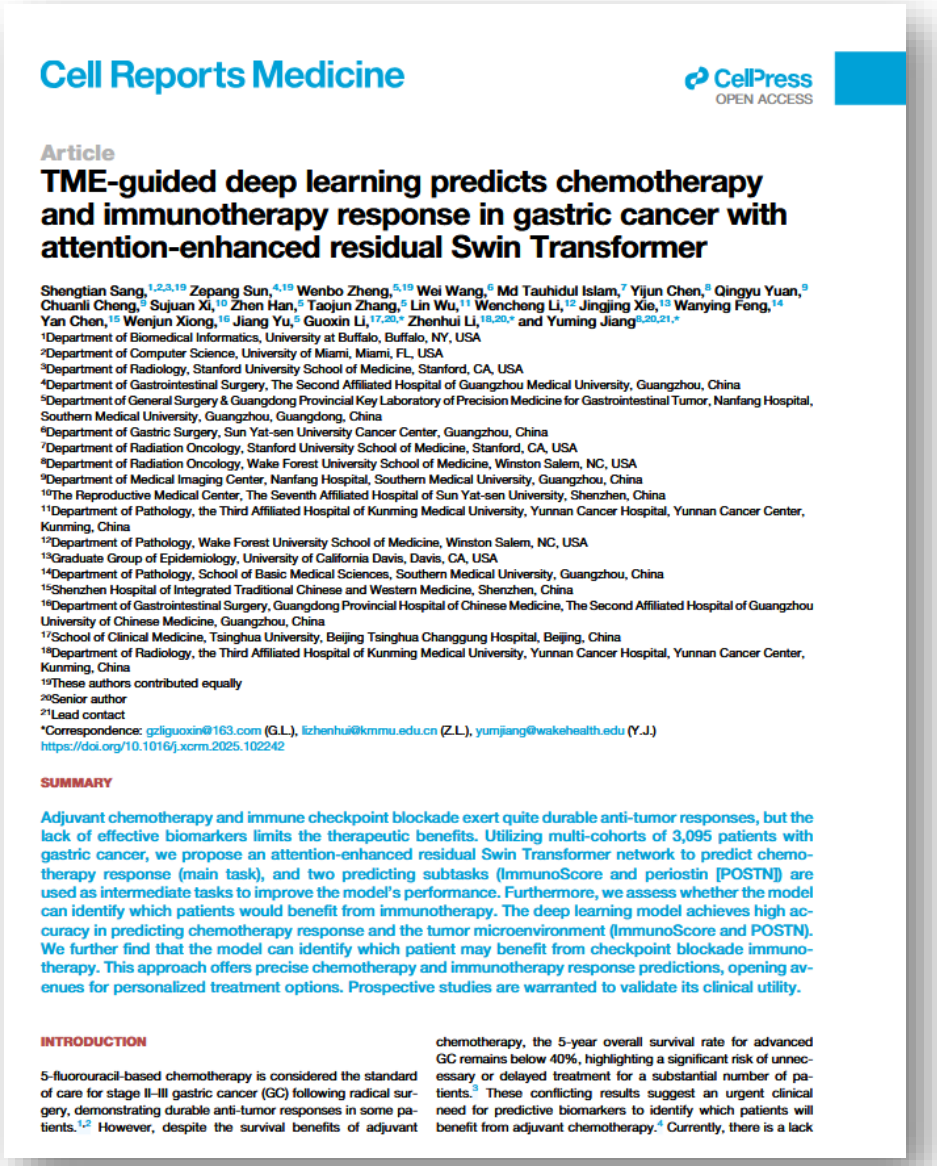
IPLab 智能感知与信息处理实验室

论文分享会

余海军　2025 年 10 月 14日

#多任务学习

#swintransformer



Cell Reports Medicine

CellPress
OPEN ACCESS

Article

TME-guided deep learning predicts chemotherapy and immunotherapy response in gastric cancer with attention-enhanced residual Swin Transformer

Shengtian Sang,[1,2,3,19] Zepang Sun,[4,19] Wenbo Zheng,[5,19] Wei Wang,[6] Md Tauhidul Islam,[7] Yijun Chen,[8] Qingyu Yuan,[9] Chuanli Cheng,[5] Sujuan Xi,[10] Zhen Han,[5] Taojun Zhang,[5] Lin Wu,[11] Wencheng Li,[12] Jingjing Xie,[13] Wanying Feng,[14] Yan Chen,[15] Wenjun Xiong,[16] Jiang Yu,[5] Guoxin Li,[17,20,*] Zhenhui Li,[18,20,*] and Yuming Jiang[8,20,21,*]

[1]Department of Biomedical Informatics, University at Buffalo, Buffalo, NY, USA
[2]Department of Computer Science, University of Miami, Miami, FL, USA
[3]Department of Radiology, Stanford University School of Medicine, Stanford, CA, USA
[4]Department of Gastrointestinal Surgery, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China
[5]Department of General Surgery & Guangdong Provincial Key Laboratory of Precision Medicine for Gastrointestinal Tumor, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, China
[6]Department of Gastric Surgery, Sun Yat-sen University Cancer Center, Guangzhou, China
[7]Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA, USA
[8]Department of Radiation Oncology, Wake Forest University School of Medicine, Winston Salem, NC, USA
[9]Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, Guangzhou, China
[10]The Reproductive Medical Center, The Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen, China
[11]Department of Pathology, the Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, China
[12]Department of Pathology, Wake Forest University School of Medicine, Winston Salem, NC, USA
[13]Graduate Group of Epidemiology, University of California Davis, Davis, CA, USA
[14]Department of Pathology, School of Basic Medical Sciences, Southern Medical University, Guangzhou, China
[15]Shenzhen Hospital of Integrated Traditional Chinese and Western Medicine, Shenzhen, China
[16]Department of Gastrointestinal Surgery, Guangdong Provincial Hospital of Chinese Medicine, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China
[17]School of Clinical Medicine, Tsinghua University, Beijing Tsinghua Changgung Hospital, Beijing, China
[18]Department of Radiology, the Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, China
[19]These authors contributed equally
[20]Senior author
[21]Lead contact
*Correspondence: gzliguoxin@163.com (G.L.), lizhenhui@kmmu.edu.cn (Z.L.), yumjiang@wakehealth.edu (Y.J.)
https://doi.org/10.1016/j.xcrm.2025.102242

SUMMARY

Adjuvant chemotherapy and immune checkpoint blockade exert quite durable anti-tumor responses, but the lack of effective biomarkers limits the therapeutic benefits. Utilizing multi-cohorts of 3,095 patients with gastric cancer, we propose an attention-enhanced residual Swin Transformer network to predict chemotherapy response (main task), and two predicting subtasks (ImmunoScore and periostin [POSTN]) are used as intermediate tasks to improve the model's performance. Furthermore, we assess whether the model can identify which patients would benefit from immunotherapy. The deep learning model achieves high accuracy in predicting chemotherapy response and the tumor microenvironment (ImmunoScore and POSTN). We further find that the model can identify which patient may benefit from checkpoint blockade immunotherapy. This approach offers precise chemotherapy and immunotherapy response predictions, opening avenues for personalized treatment options. Prospective studies are warranted to validate its clinical utility.

INTRODUCTION

5-fluorouracil-based chemotherapy is considered the standard of care for stage II–III gastric cancer (GC) following radical surgery, demonstrating durable anti-tumor responses in some patients.[1,2] However, despite the survival benefits of adjuvant chemotherapy, the 5-year overall survival rate for advanced GC remains below 40%, highlighting a significant risk of unnecessary or delayed treatment for a substantial number of patients.[3] These conflicting results suggest an urgent clinical need for predictive biomarkers to identify which patients will benefit from adjuvant chemotherapy.[4] Currently, there is a lack

论文基本信息

- 论文题目：
  - **TME-guided deep learning predicts chemotherapy and immunotherapy response in gastric cancer with attention-enhanced residual Swin Transformer**
  - TME引导的深度学习利用注意力增强残差swin-transformer预测胃癌的化疗和免疫疗法反应
  - TME:肿瘤微环境（**Tumor Microenvironment**）

- 作者信息：美国纽约布法罗大学等
- 会议分区：2025 Cell Reports Medicine

Summary

Adjuvant chemotherapy and immune checkpoint blockade exert quite durable anti-tumor responses, but the lack of effective biomarkers limits the therapeutic benefits. Utilizing multi-cohorts of 3,095 patients with gastric cancer, we propose an attention-enhanced residual Swin Transformer network to predict chemotherapy response (main task), and two predicting subtasks (ImmunoScore and periostin [POSTN]) are used as intermediate tasks to improve the model's performance. Furthermore, we assess whether the model can identify which patients would benefit from immunotherapy. The deep learning model achieves high accuracy in predicting chemotherapy response and the tumor microenvironment (ImmunoScore and POSTN). We further find that the model can identify which patient may benefit from checkpoint blockade immunotherapy. This approach offers precise chemotherapy and immunotherapy response predictions, opening avenues for personalized treatment options. Prospective studies are warranted to validate its clinical utility.

任务背景：辅助化疗和免疫检查点抑制剂已被证明对胃癌有持久的抗肿瘤作用。但目前缺乏有效的生物标志物，（无法准确预测哪些患者会真正从这些治疗中获益）。这限制了个体化治疗的发展。

实验效果：该深度学习模型在预测化疗反应和肿瘤微环境（免疫评分和POSTN）方面表现出较高的准确性。**我们进一步发现，该模型能够识别出哪些患者可能从检查点阻断免疫治疗中受益**。

本文方法与任务：我们提出了一种注意力增强的残差Swin Transformer网络，用于**预测化疗反应（主要任务）**，并且使用两个预测子任务（免疫评分和骨膜素[POSTN]）作为中间任务，以提高模型的性能。（利用3,095名胃癌患者的多组队列数据）

这一方法提供了精准的化疗和免疫治疗反应预测，为个性化治疗方案开辟了新的途径。未来的前瞻性研究有望验证其临床应用价值。

强调临床意义：此外，我们还评估了该模型是否能识别哪些患者会从免疫治疗中受益。

引言读解

TME-guided deep learning predicts chemotherapy and immunotherapy response in gastric cancer with attention-enhanced residual Swin Transformer

第一段：任务急迫性1：缺乏有效生物标注物预测化疗疗效

5-fluorouracil-based chemotherapy is considered the standard of care for stage II–III gastric cancer (GC) following radical surgery, demonstrating durable anti-tumor responses in some patients.[1,2] However, despite the survival benefits of adjuvant chemotherapy, the 5-year overall survival rate for advanced GC remains below 40%, highlighting a significant risk of unnecessary or delayed treatment for a substantial number of patients.[3] These conflicting results suggest an urgent clinical need for predictive biomarkers to identify which patients will benefit from adjuvant chemotherapy.[4] Currently, there is a lack of effective biomarkers for predicting the benefit of 5-fluorouracil-based chemotherapy, making the identification of predictive biomarkers to personalize treatment for patients with GC crucial and long overdue.

化疗是胃癌标准治疗手段：以5-氟尿嘧啶为基础的化疗被认为是根治性手术后II－III期胃癌（GC）的标准治疗，能够在一些患者中产生持久的抗肿瘤反应。

化疗的治疗矛盾：然而，尽管辅助化疗具有生存获益，晚期胃癌的5年总生存率仍低于40%，这突显出**相当一部分患者可能面临不必要或延误治疗的显著风险**。

引出任务重要性：这些矛盾的结果表明，迫切需要临床预测生物标志物，以识别哪些患者能够从辅助化疗中获益。

挖坑：目前，缺乏有效的生物标志物来预测5-氟尿嘧啶基础化疗的疗效，因此，为胃癌患者个性化治疗寻找预测性生物标志物显得尤为关键且迫切。

引言读解

TME-guided deep learning predicts chemotherapy and immunotherapy response in gastric cancer with attention-enhanced residual Swin Transformer

第二段：任务急迫性2：缺乏有效生物标注物预测免疫治疗疗效

Recently, the emergence of immunotherapies, including immune checkpoint inhibitors (ICIs), cancer vaccines, oncolytic viruses, and cell therapies, has revolutionized cancer treatment.[5–7] Among these, ICIs have become a first-line treatment option for advanced GC.[5,8,9] Preliminary positive outcomes have also been observed with perioperative immunotherapy.[10–13] However, the low response rate of ICIs and the difficulty in identifying benefit groups remain major clinical challenges.[14] Although biomarkers such as programmed cell death ligand 1 (PD-L1), Epstein-Barr virus, microsatellite instability, and tumor mutation burden have been approved to guide ICI therapy, none are fully predictive of immunotherapy responses.[14–17] Therefore, accurately identifying patients who will benefit from immunotherapy to maximize therapeutic outcomes is a critical issue that needs to be addressed.

免疫治疗：近年来，免疫治疗的出现，包括免疫检查点抑制剂（ICIs）、癌症疫苗、溶瘤病毒和细胞治疗，彻底改变了癌症治疗的格局。

免疫检查点抑制治疗：在这些治疗方法中，免疫检查点抑制剂已成为晚期胃癌的首选治疗方案。初步的积极结果也表明围手术期免疫治疗具有一定的疗效。

引出任务重要性：然而，免疫检查点抑制剂的低反应率以及难以识别获益患者群体仍然是主要的临床挑战。尽管像程序性死亡配体1（PD-L1）、埃布斯坦-巴尔病毒、微卫星不稳定性和肿瘤突变负荷等生物标志物已被批准用于指导免疫检查点抑制剂治疗，但没有任何一种能够完全预测免疫治疗的反应。

挖坑：因此，准确识别能够从免疫治疗中获益的患者，以最大化治疗效果，是一个亟待解决的关键问题。

引言读解

TME-guided deep learning predicts chemotherapy and immunotherapy response in gastric cancer with attention-enhanced residual Swin Transformer

第三段：有效评估TME对免疫治疗/化疗的价值

The tumor microenvironment (TME) consists of innate immune cells, adaptive immune cells, cytokines, and extracellular matrix components.[18–20] These elements form a complex regulatory network that plays a crucial role in tumor progression and treatment response.[19–21] Accurate evaluation of the TME can improve the assessment of immunotherapy effectiveness in GC.[22] Additionally, the tumor immune status of patients is closely correlated with chemotherapy response, highlighting the value of incorporating TME evaluation into chemotherapy response assessment.[23] Our previous studies have shown that ImmunoScore and periostin (POSTN) expression are significantly associated with chemotherapy benefit.[21,24] However, the assessment of the TME relying on histopathological staining has limitations, such as invasiveness, sample heterogeneity, high cost, time consuming nature, and technical complexity.[25–27] Therefore, developing a non-invasive and cost-effective model to assess chemotherapy response while integrating TME status to enhance model performance is imperative.

TME介绍：**肿瘤微环境**由固有免疫细胞、适应性免疫细胞、细胞因子和细胞外基质成分组成。这些元素形成了一个复杂的调控网络，**在肿瘤进展和治疗反应中发挥着至关重要的作用**。

TME在免疫治疗反应/化疗反应的价值：准确评估肿瘤微环境可以提高胃癌免疫治疗效果的评估。此外，患者的肿瘤免疫状态与化疗反应密切相关，这凸显了将肿瘤微环境评估纳入化疗反应评估中的价值。

引出子任务：我们的前期研究表明，免疫评分（ImmunoScore）和骨膜素（POSTN）表达与化疗获益显著相关。

挖坑（评估TME难点）：然而，依赖组织病理学染色评估肿瘤微环境存在一定局限性，如侵入性强、样本异质性大、成本高、耗时且技术复杂。因此，开发一种非侵入性且具有成本效益的模型来评估化疗反应，并整合肿瘤微环境状态以提升模型性能，

引言读解

TME-guided deep learning predicts chemotherapy and immunotherapy response in gastric cancer with attention-enhanced residual Swin Transformer

第四段：多任务深度学习是一种非侵入且高效的方法

挖坑（评估TME难点）：然而，依赖组织病理学染色评估肿瘤微环境存在一定局限性，如侵入性强、样本异质性大、成本高、耗时且技术复杂。因此，开发一种非侵入性且具有成本效益的模型来评估化疗反应，并整合肿瘤微环境状态以提升模型性能，

Radiological imaging is a non-invasive tool routinely used for diagnosis, staging, and treatment evaluation in patients with cancer, including those with GC.[28,29] Recently, deep learning has emerged as a transformative methodology for automatically learning representative features from annotated tumor images for disease evaluation. Traditionally, deep learning models have been designed for single-task purposes, such as predicting specific clinical outcomes (e.g., lymph node metastasis).[30,31] In contrast, multitask deep learning enables the simultaneous analysis of different tasks within a single model.[32] By sharing feature representations and interactions among related tasks, multitask learning is more data efficient and has been shown to reduce overfitting and improve model generalization across various applications, including computer vision, disease diagnosis, and drug discovery.[33–35]

放射影像深度学习是非侵入预测工具：放射影像是癌症患者（包括胃癌患者）常规用于诊断、分期和治疗评估的非侵入性工具。近年来，深度学习作为一种变革性的方法，已经能够自动从标注的肿瘤图像中学习代表性特征，以进行疾病评估。

单任务模型：传统上，深度学习模型通常是为单一任务设计的，例如预测特定的临床结果（如淋巴结转移）。

多任务模型：相比之下，多任务深度学习使得在一个模型中同时分析不同的任务成为可能。通过共享特征表示和相关任务之间的交互，多任务学习具有更高的数据利用效率，并且已被证明能够减少过拟合并提高模型在各种应用中的泛化能力，包括计算机视觉、疾病诊断和药物发现。

引言读解

TME-guided deep learning predicts chemotherapy and immunotherapy response in gastric cancer with attention-enhanced residual Swin Transformer

第五段：介绍本文方法与优势

Here, we developed a multitask deep learning model called attention-enhanced residual Swin Transformer (AER-SwinT) for the simultaneous prediction of chemotherapy response and TME characteristics using preoperative computed tomography (CT) images. Compared to existing convolutional neural networks (CNNs) or transformer-based methods, our approach captures global feature relationships while leveraging multiple granularities for data analysis. AER-SwinT employs the Swin Transformer to extract hierarchical features from CT images, enabling analysis at various levels of granularity, from fine to coarse.[36–41] Additionally, channel average-based attention maps and residual connections enhance the AER-SwinT model's ability to dynamically refine its focus on spatial regions, improving feature extraction across multiple stages.[42] Furthermore, this multitask learning approach enables the model to capture more comprehensive features from the training data. By using chemotherapy response as the primary task and ImmunoScore and POSTN expression as intermediate tasks, the model leverages these clinically relevant immune biomarkers to improve predictive accuracy over single-task methods. Given the established clinical importance of these immune biomarkers, we also evaluated the model's ability to predict immunotherapy outcomes.

总结本文方法：在这项研究中，我们开发了一种多任务深度学习模型，称为注意力增强的残差Swin Transformer（AER-SwinT），用于通过术前计算机断层扫描（CT）图像同时预测化疗反应和肿瘤微环境（TME）特征。

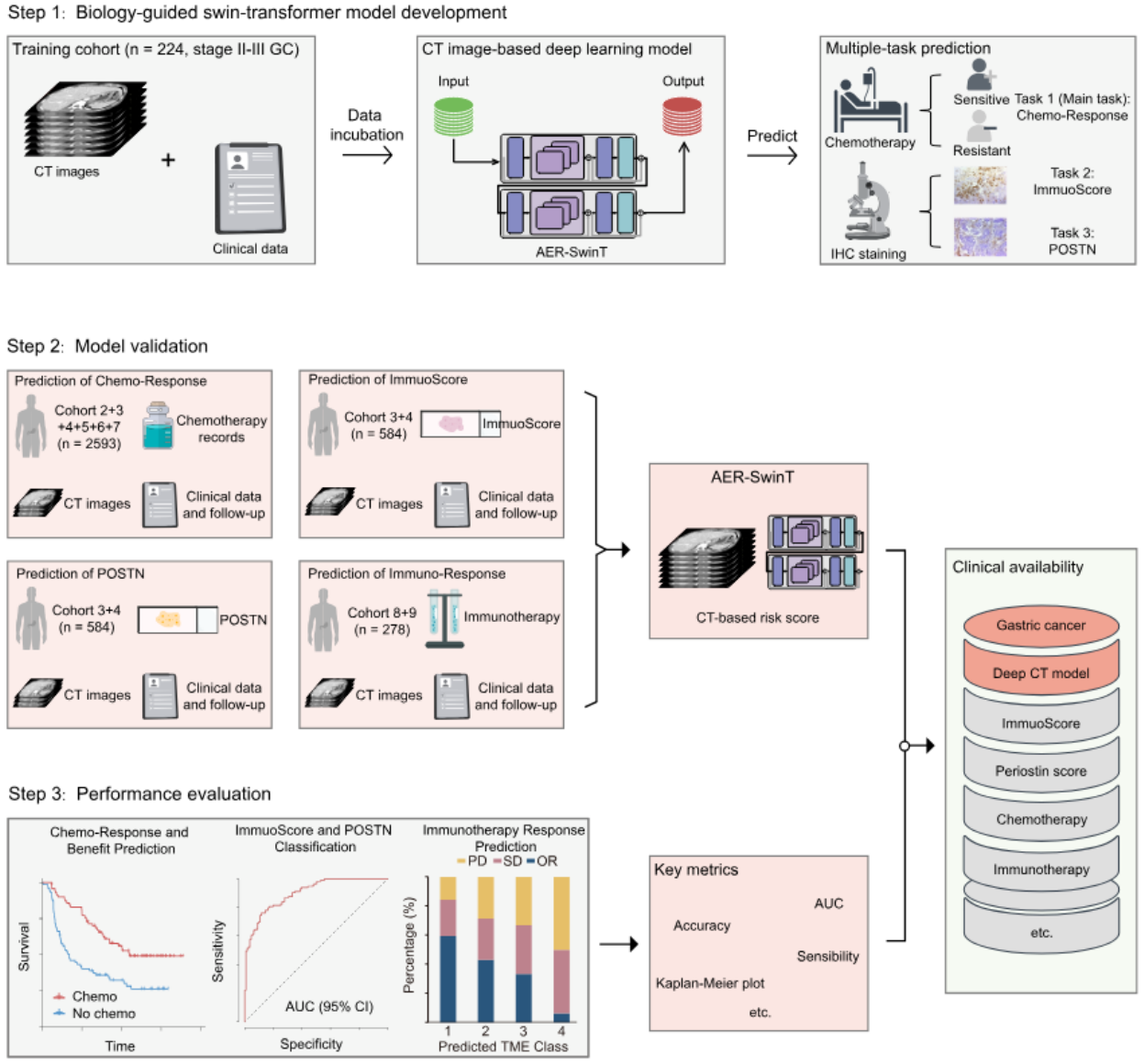
方法优势1（多层次粒度分析）：与现有的卷积神经网络（CNN）或基于变换器的方法相比，我们的方法能够捕捉全局特征关系，同时利用多层次粒度进行数据分析。AER-SwinT使用SwinTransformer从CT图像中提取层次化特征，能够在从精细到粗略的不同粒度级别进行分析。

注意力增强+残差动机：此外，基于通道平均的注意力图和残差连接增强了AER-SwinT模型动态调整焦点的能力，从而改善了多阶段特征提取。

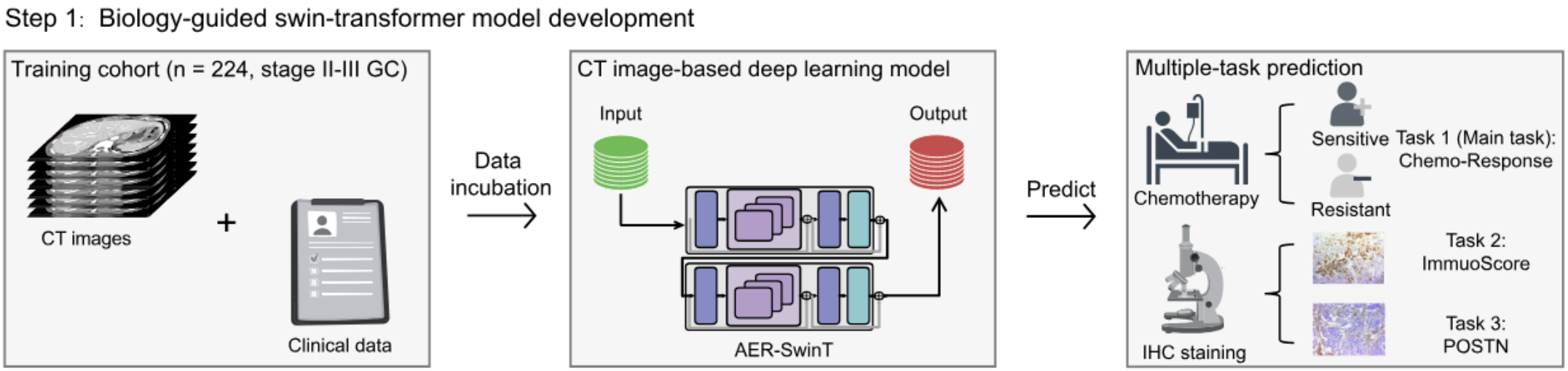
方法优势2（多任务学习）：更重要的是，这种多任务学习方法使模型能够从训练数据中捕捉更全面的特征。通过将化疗反应作为主要任务，免疫评分（ImmunoScore）和骨膜素（POSTN）表达作为中间任务，该模型利用这些具有临床相关性的免疫生物标志物，提升了预测准确性，相比单任务方法更为精准。

鉴于这些免疫生物标志物在临床上的重要性，我们还评估了该模型预测免疫治疗效果的能力。

后续-方法学-研究流程


Step 1: Biology-guided swin-transformer model development
Training cohort (n = 224, stage II-III GC)
CT images
Clinical data
Data incubation
CT image-based deep learning model
Input
Output
AER-SwinT
Predict
Multiple-task prediction
Chemotherapy
Sensitive
Resistant
Task 1 (Main task): Chemo-Response
Task 2: ImmuoScore
Task 3: POSTN
IHC staining
Step 2: Model validation
Prediction of Chemo-Response
Cohort 2+3 +4+5+6+7 (n = 2593)
Chemotherapy records
CT images
Clinical data and follow-up
Prediction of ImmuoScore
Cohort 3+4 (n = 584)
ImmuoScore
CT images
Clinical data and follow-up
Prediction of POSTN
Cohort 3+4 (n = 584)
POSTN
CT images
Clinical data and follow-up
Prediction of Immuno-Response
Cohort 8+9 (n = 278)
Immunotherapy
CT images
Clinical data and follow-up
AER-SwinT
CT-based risk score
Clinical availability
Gastric cancer
Deep CT model
ImmuoScore
Periostin score
Chemotherapy
Immunotherapy
etc.
Step 3: Performance evaluation
Chemo-Response and Benefit Prediction
Survival
Chemo
No chemo
Time
ImmuoScore and POSTN Classification
Sensitivity
AUC (95% CI)
Specificity
Immunotherapy Response Prediction
PD SD OR
Percentage (%)
1 2 3 4
Predicted TME Class
Key metrics
AUC
Accuracy
Sensibility
Kaplan-Meier plot
etc.

后续-方法学-研究流程



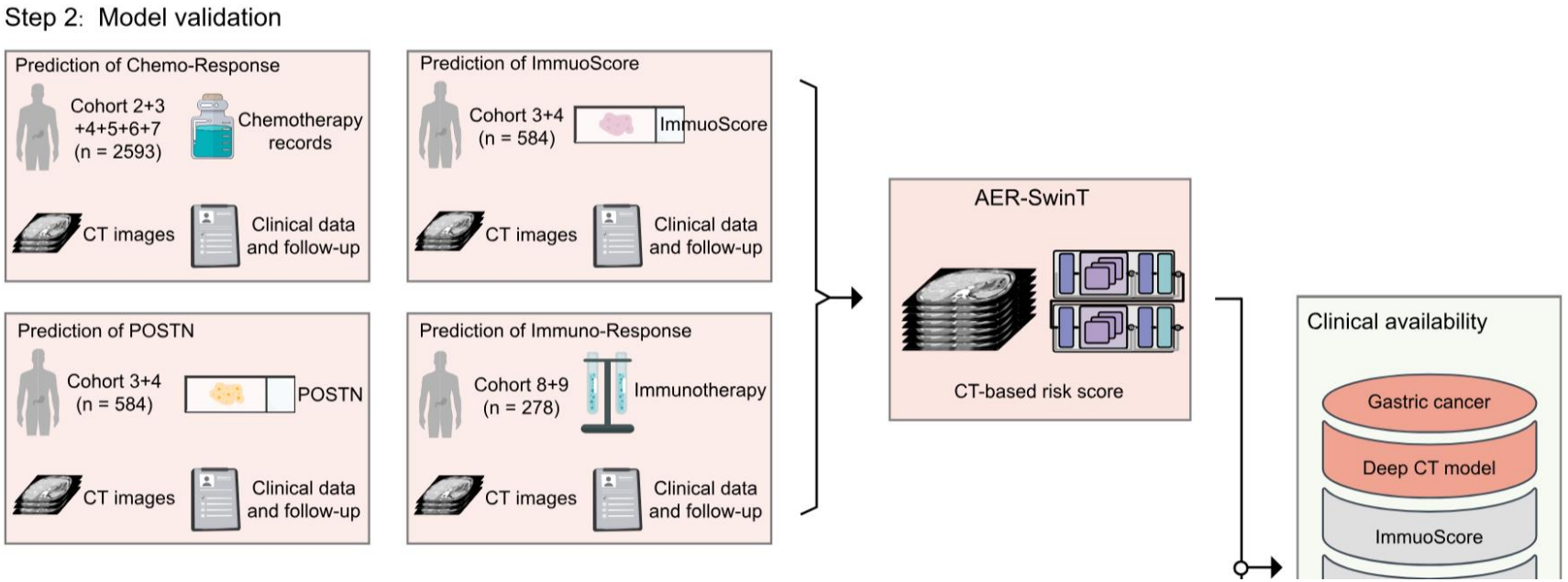
步骤1：生物学引导的Swin-Transformer模型开发

•训练队列（n=224，II-III期胃癌）：使用CT影像数据和临床数据来训练AER-SwinT模型。

•数据处理：CT影像和临床数据被送入深度学习模型（AER-SwinT）进行训练和预测。

•多任务预测：主任务：预测化疗反应（敏感/耐药性），进行免疫组化染色（子任务）：预测免疫评分（ImmunoScore）和骨髓素表达。

后续-方法学-研究流程



步骤2：模型验证

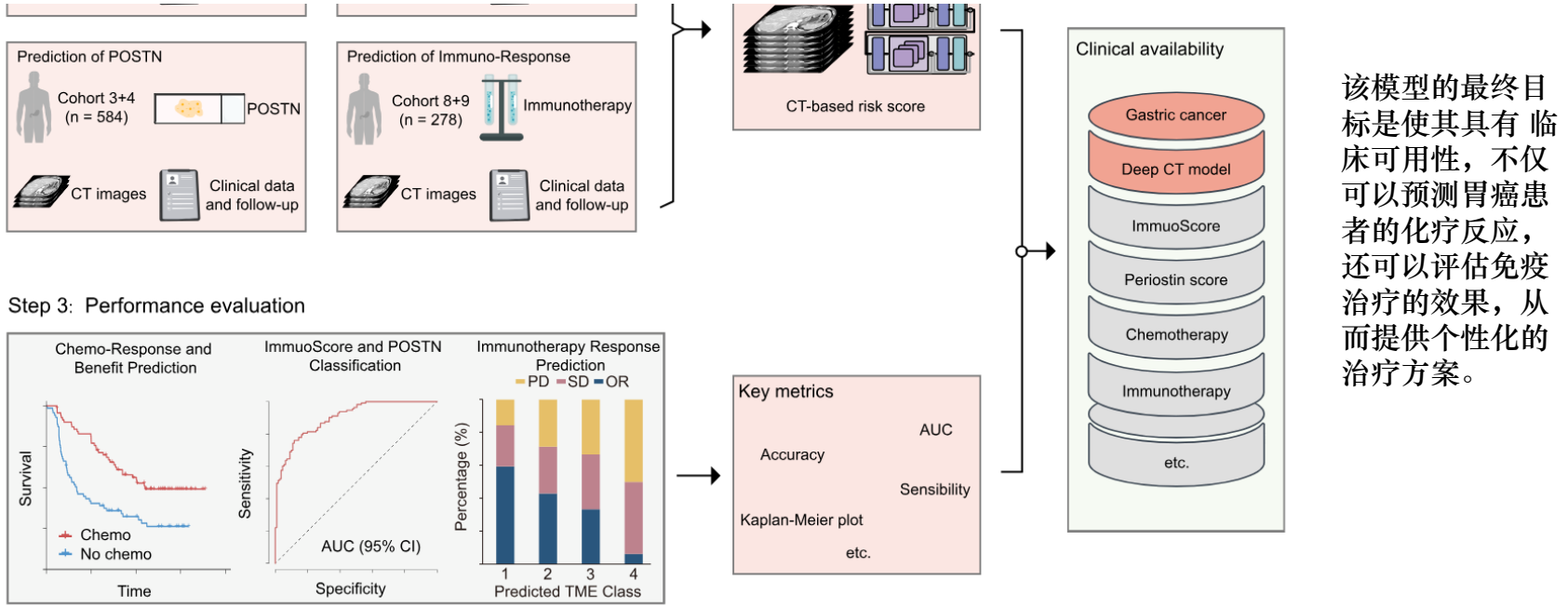
化疗反应预测（主要任务）：通过使用化疗记录、CT图像、临床数据和随访数据进行预测（队列2+3+4+5+6+7，共2,593名患者）。

免疫评分预测（ImmunoScore）：通过CT影像和临床数据来预测免疫评分（队列3+4，共584名患者）。

POSTN预测：预测骨膜素（POSTN）的表达（队列3+4，共584名患者）。

免疫治疗反应预测（免疫反应）：评估哪些患者可能从免疫治疗中受益（队列8+9，共276名患者）。

后续-方法学-研究流程



该模型的最终目标是使其具有 临床可用性，不仅可以预测胃癌患者的化疗反应，还可以评估免疫治疗的效果，从而提供个性化的治疗方案。

步骤3：性能评估

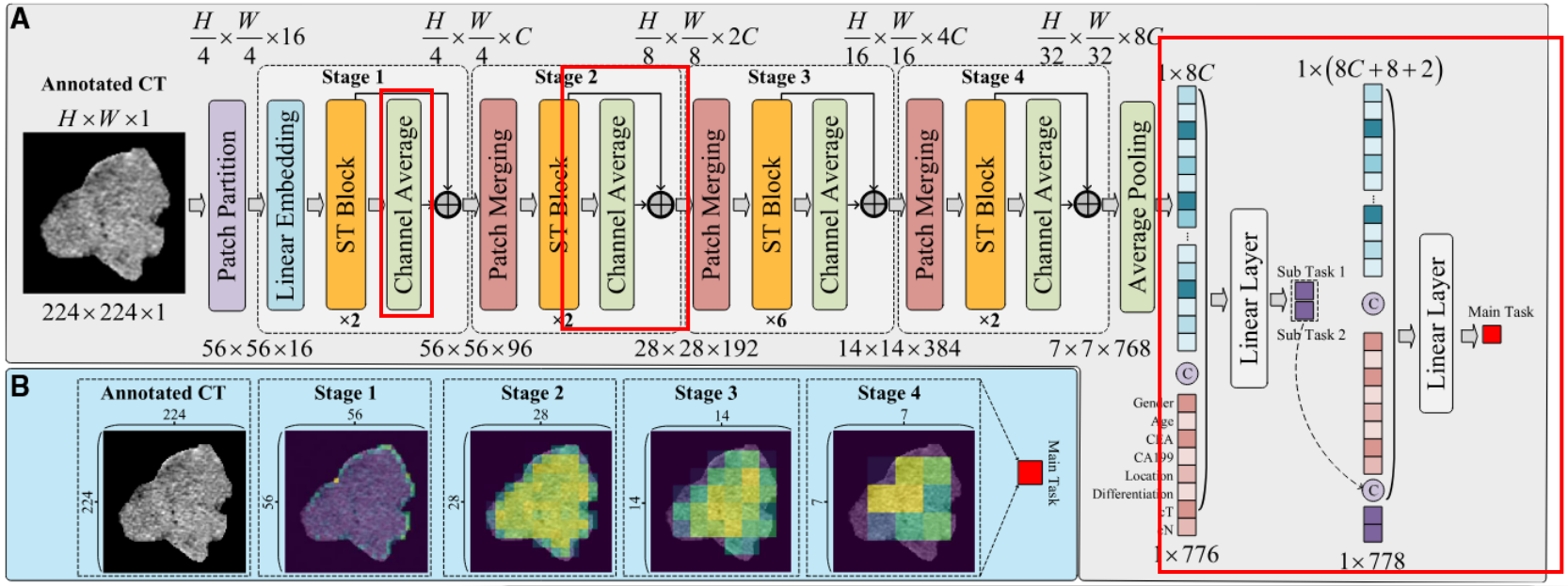
化疗反应与获益预测：通过Kaplan-Meier生存曲线，评估化疗反应的影响（生存期）。

免疫评分和POSTN分类：评估免疫评分和POSTN预测的效果，包括AUC（曲线下面积）和其他指标。

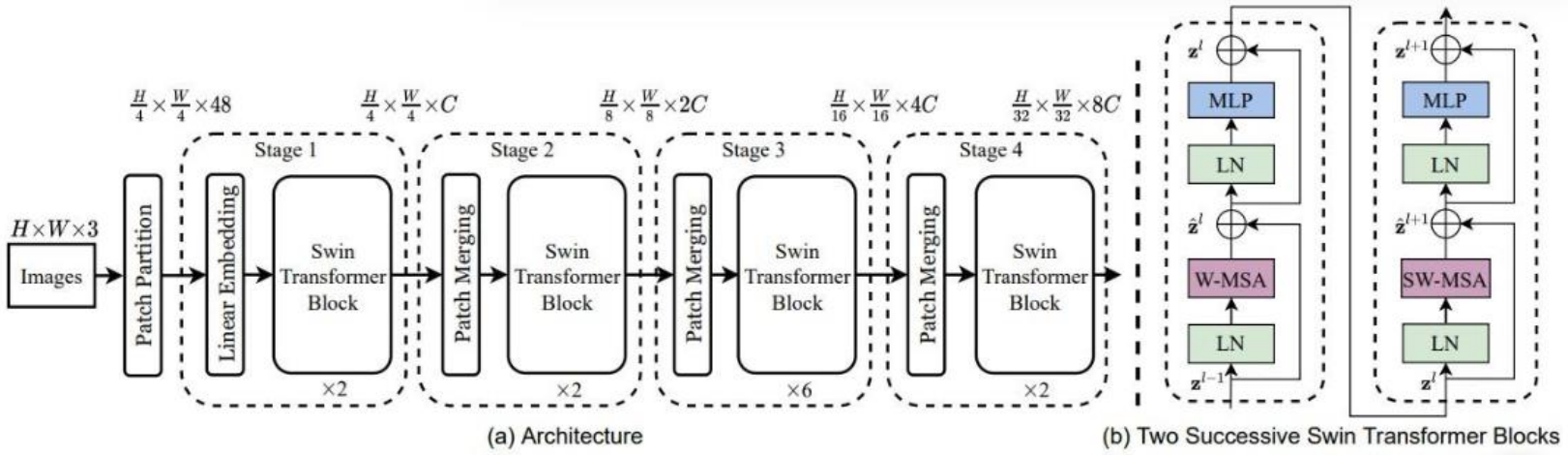
免疫治疗反应预测：通过预测免疫治疗反应（例如，PD、SD或OR）。

后续-方法学

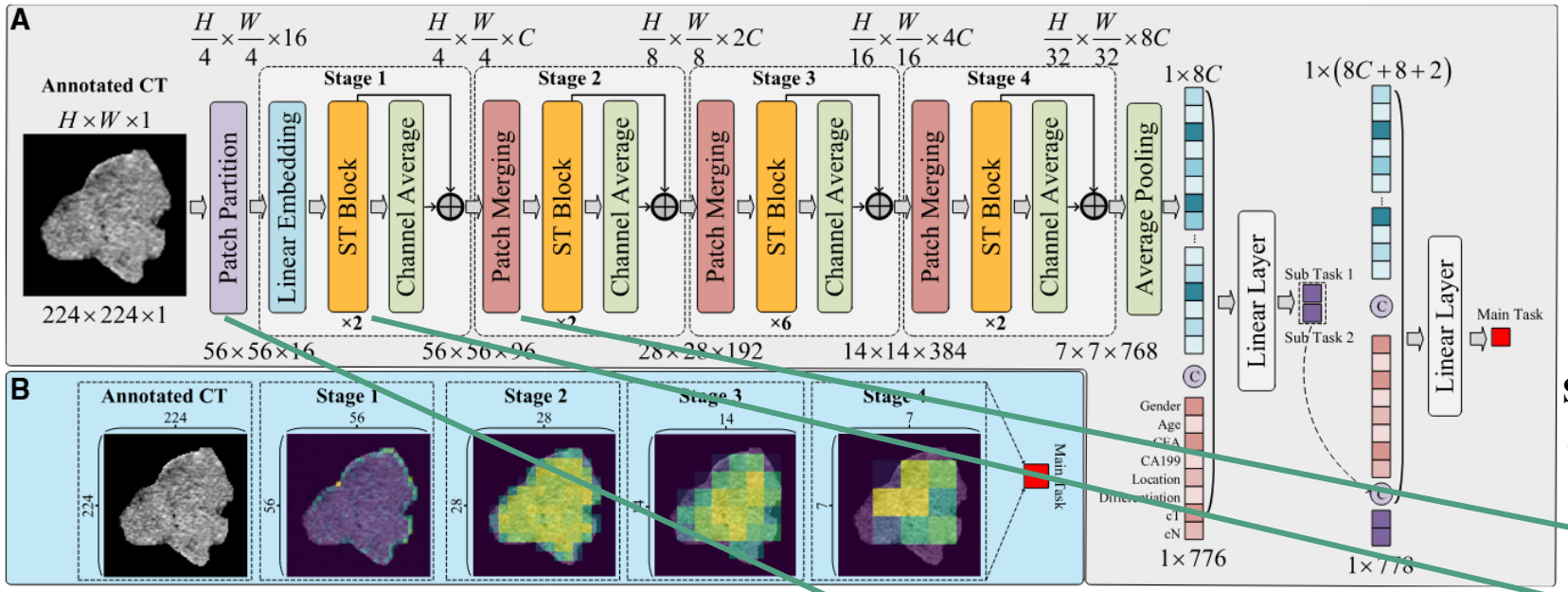
AER-SwinT



Swin Transformer



后续-方法学

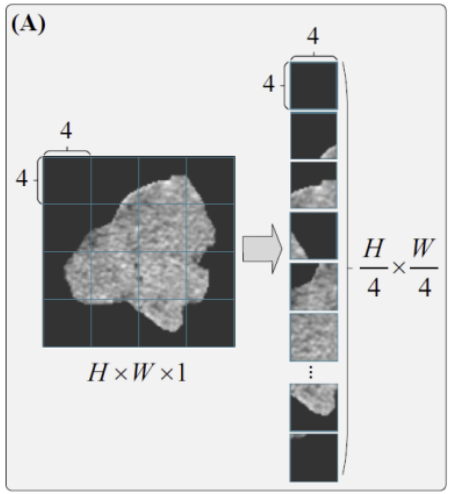


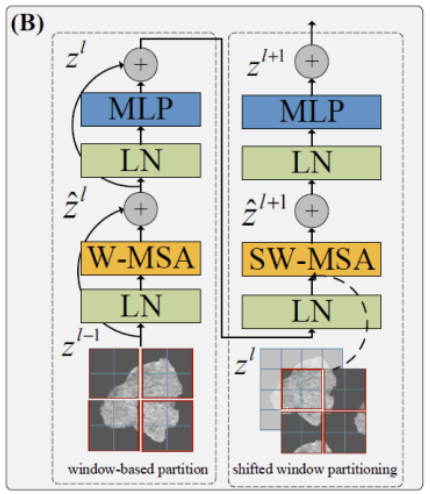
ST block：包含窗口注意力机制+滑窗后注意力机制

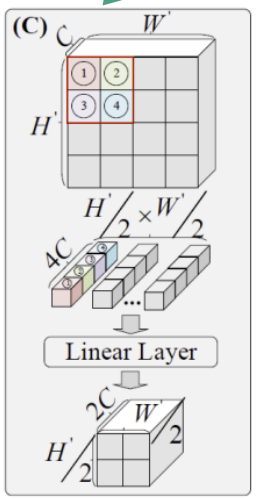
降采样，缩小分辨率，以形成层次

这里和SW-T无差别：

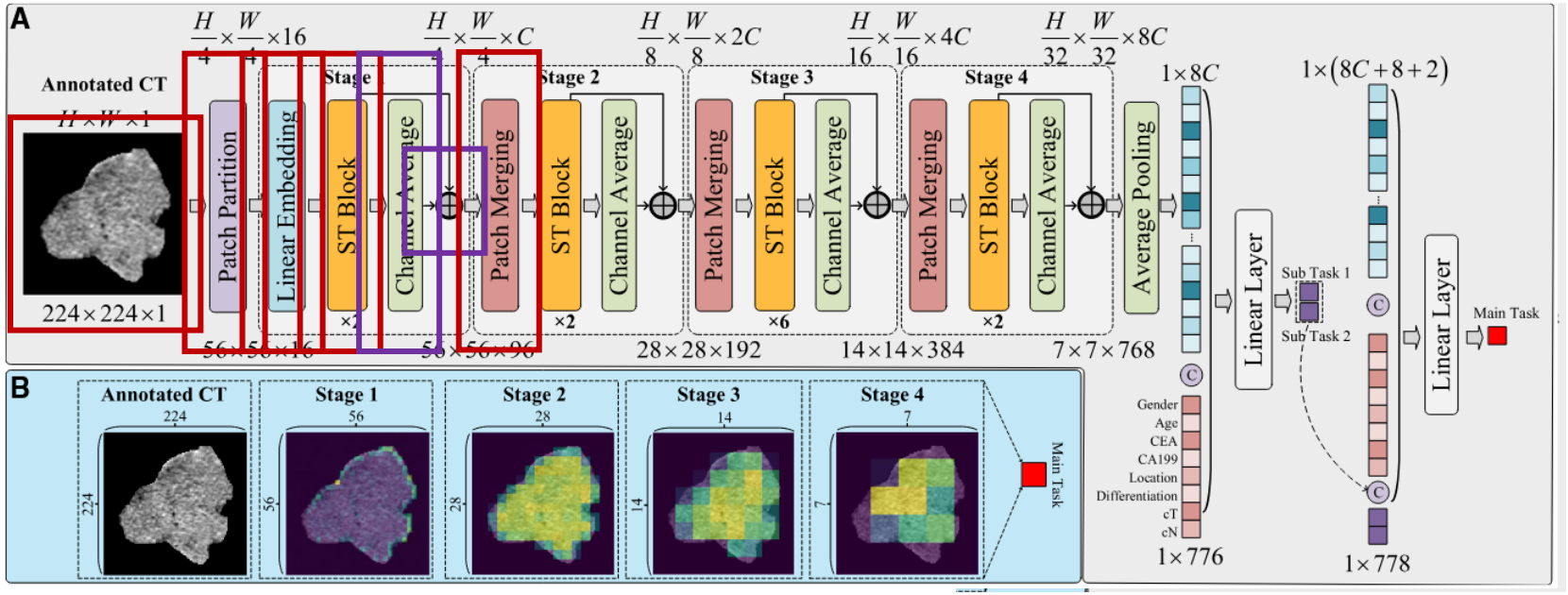
分成多个窗口

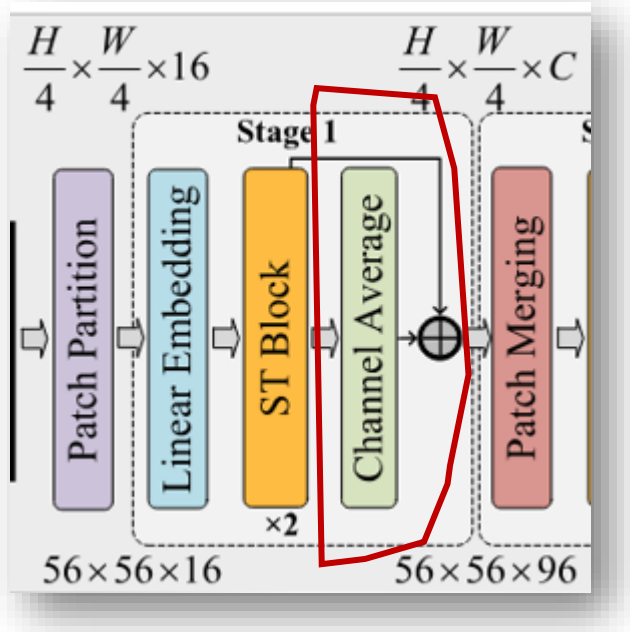






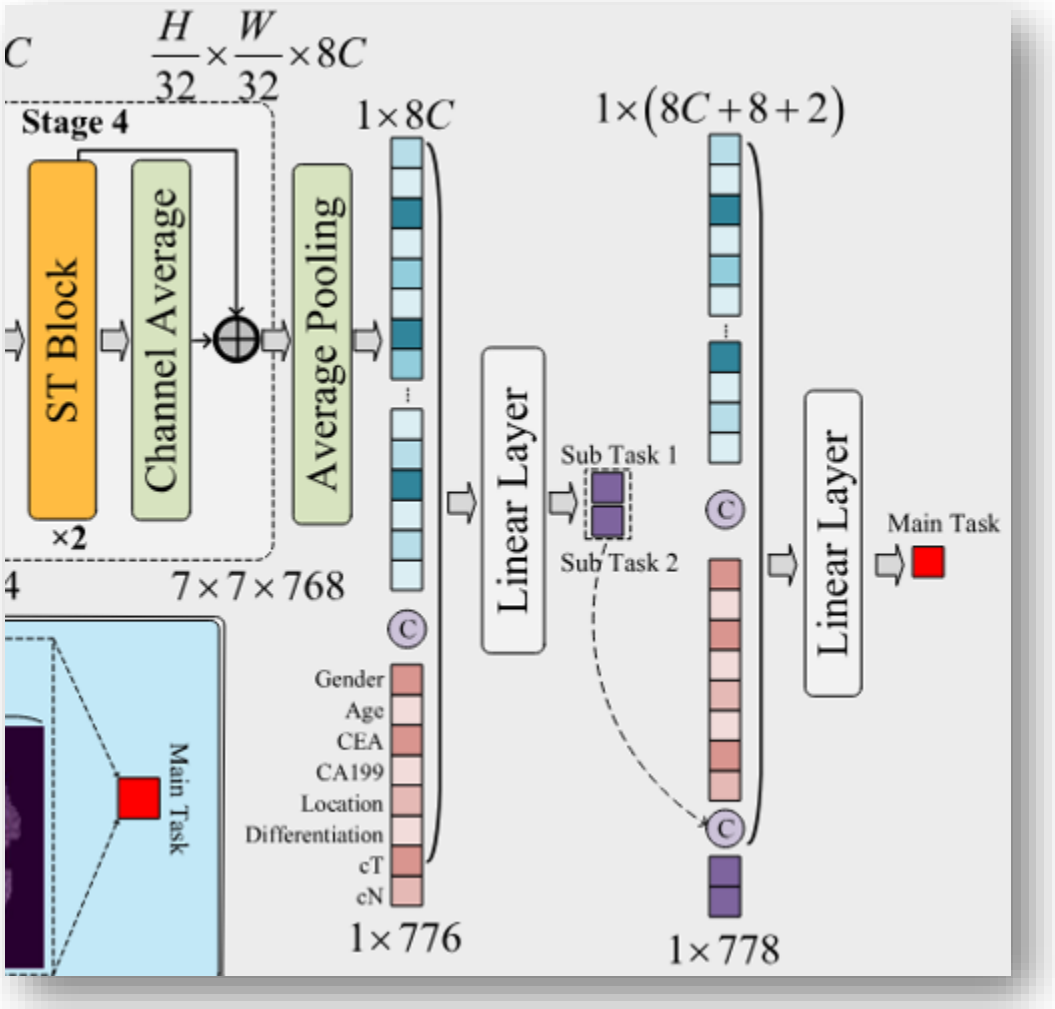
后续-方法学





- 图像原始尺寸是224*224*1。
- 按照4*4为一个窗，patch partition将原始图像分为了（224/4 * 224/4）个窗，即形状为56*56*16。
- 接着Linear embedding 进行16→C的嵌入，形状变为56*56*C。
- ST Block分别对每个窗口进行内部注意力计算，接着滑窗再进行一次内部注意力计算，最后得到了新的特征图 56*56*C。
- （new）Channel average 对每个通道进行平均，获得每一个窗的注意力得分，即获得注意力图 56*56*1。
- （new）注意力图 通过逐个点乘 和 残差连接提取的ST-W特征 结合 获得 stage1 的输出。56*56*C。
- 以上被包装为 基于通道平均的注意力增强。
- 进入stage2 patch Merging 会将特征图尺寸长款各缩小1/2，通道数翻倍，降低了分辨率，即28*28*2C。

后续-方法学



- 四个阶段最终特征图尺寸：7*7*8C 即 7*7*768。
- 使用平均池化转化成一个向量 1*768。
- 此外本文还使用了8个临床特征，拼接后即特征维度是 768+8=776。

- 用776维向量预测2个子任务获得 2维向量。
- 拼接获得 776+2=778维向量 预测主任务。

$$L_{main} = -\frac{1}{N}\sum_{i=1}^{N} y_{main} \cdot log(\hat{y}_{main}) + (1 - y_{main}) \cdot log(1 - \hat{y}_{main})$$

$$L_{sub1} = -\frac{1}{N}\sum_{i=1}^{N} y_{sub1} \cdot log(\hat{y}_{sub1}) + (1 - y_{sub1}) \cdot log(1 - \hat{y}_{sub1})$$

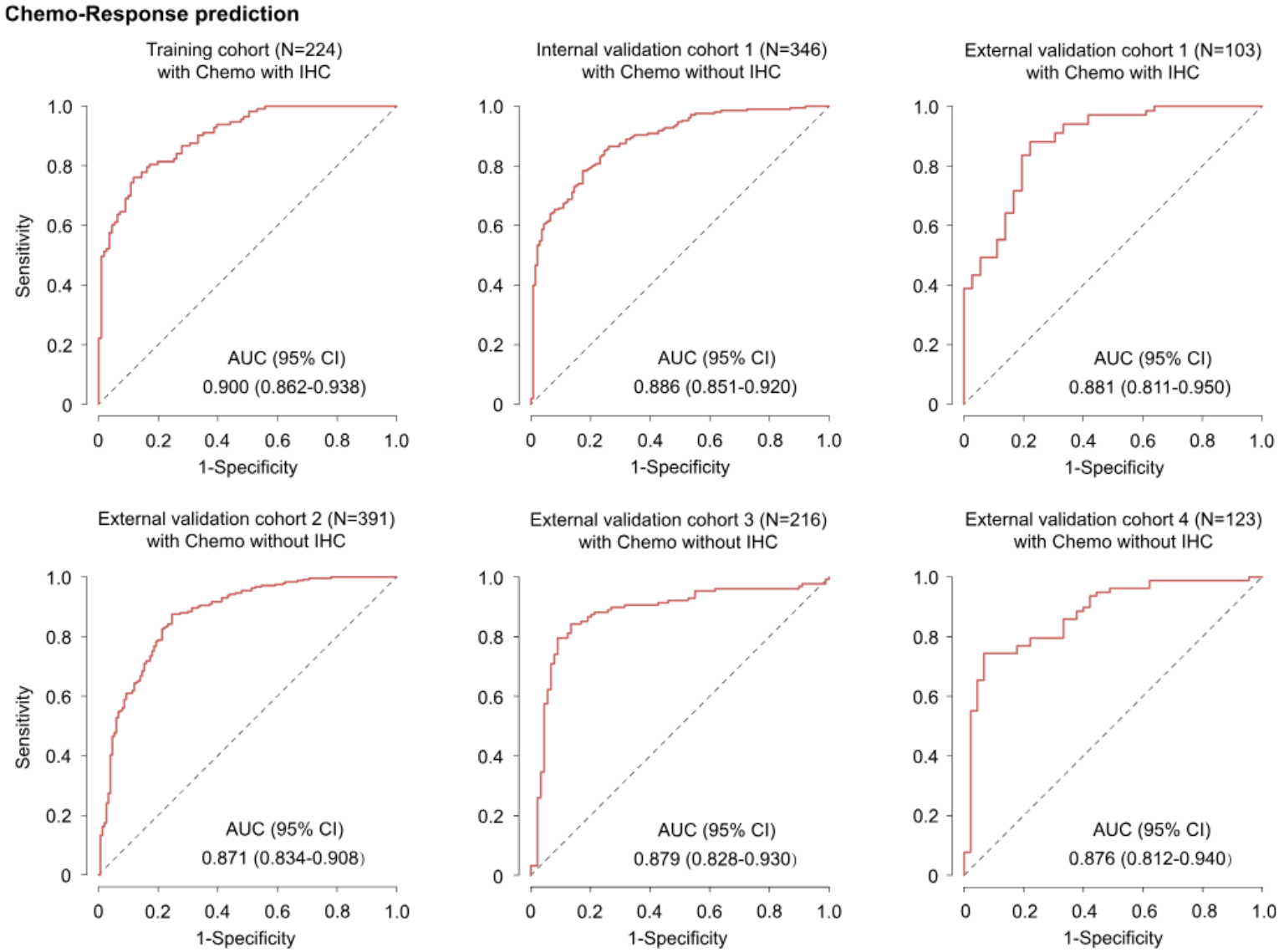
$$L_{sub2} = -\frac{1}{N}\sum_{i=1}^{N} y_{sub2} \cdot log(\hat{y}_{sub2}) + (1 - y_{sub2}) \cdot log(1 - \hat{y}_{sub2})$$

$$L = L_{main} + \frac{1}{2}(L_{sub1} + L_{sub2})$$

实验结果-数据

- **总体情况**:
- **男性患者**：1,926例（68.4%）
- **中位年龄**：57.0岁（四分位数范围：49.0－65.0岁）
- **疾病分期**:
- **II期或III期疾病**：2,579例（91.6%）
- **治疗情况**:
- **接受氟尿嘧啶类辅助化疗的患者**：1,403例（54.4%）
    - 这部分患者用于化疗反应预测任务（主要任务）
- **具有免疫组化（IHC）切片的患者**：808例（28.7%）
    - 这部分患者用于TME评估预测任务（子任务）
- **免疫治疗队列**:
- **男性患者**：161例（57.9%）
- **中位年龄**：57.0岁（四分位数范围：48.0－65.0岁）
- **队列特点**:
- 不同队列之间的人口学、临床病理和治疗特征大致相似且均衡。

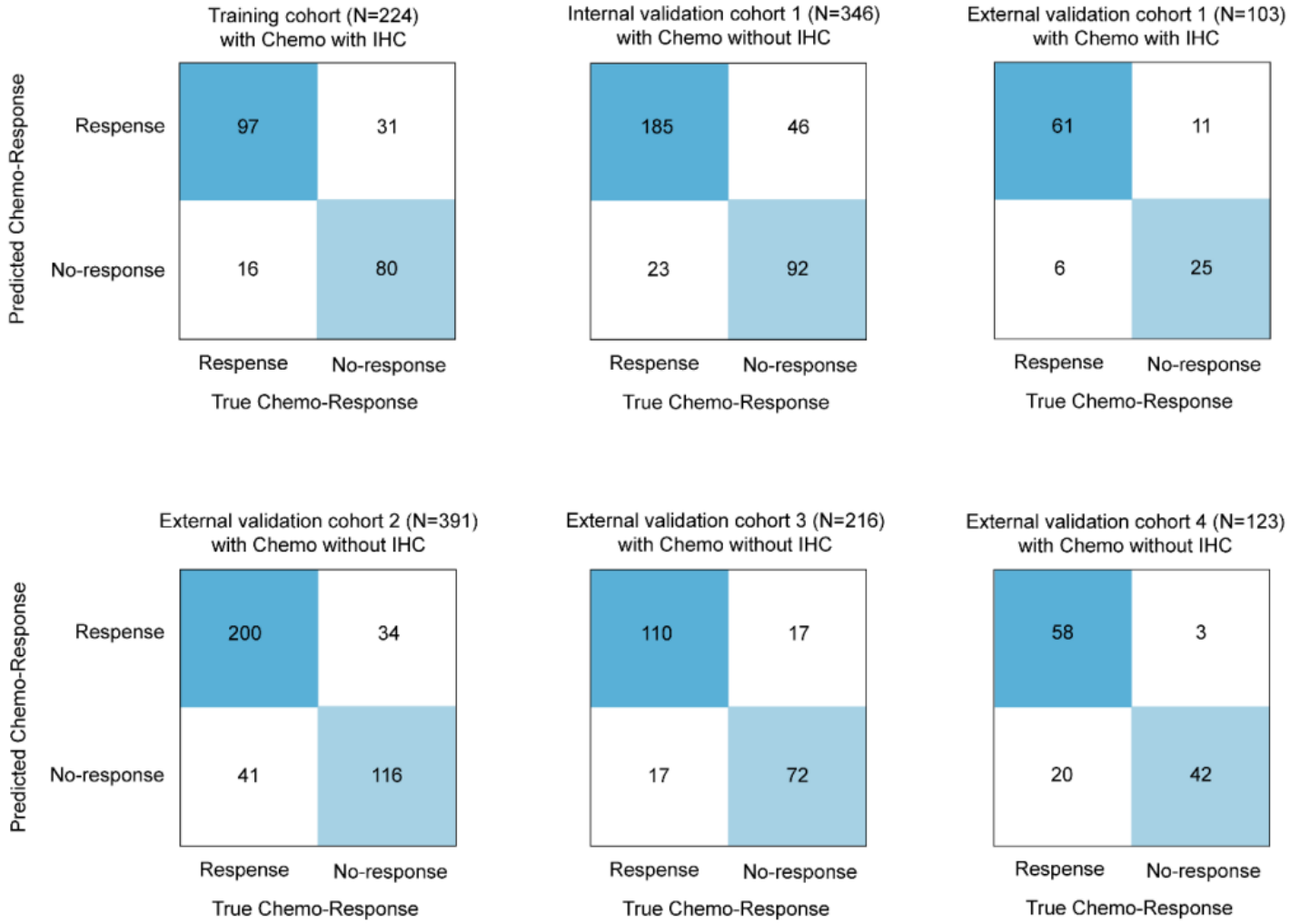
实验结果-化疗反应预测模型的准确性



- 在训练组和内部验证组中的曲线下面积（AUC）分别为0.900（95%置信区间：0.862–0.938）和0.886（95%：0.851–0.920）。
- 在外部验证队列1中也观察到了类似的结果，AUC为0.881（95% CI：0.811–0.950），外部验证队列2的AUC为0.871（95% CI：0.834–0.908）
- 外部队列3-4也显示出相对较高的准确性，AUC分别为0.879（95% CI：0.828–0.930）和0.876（95% CI：0.812–0.940）

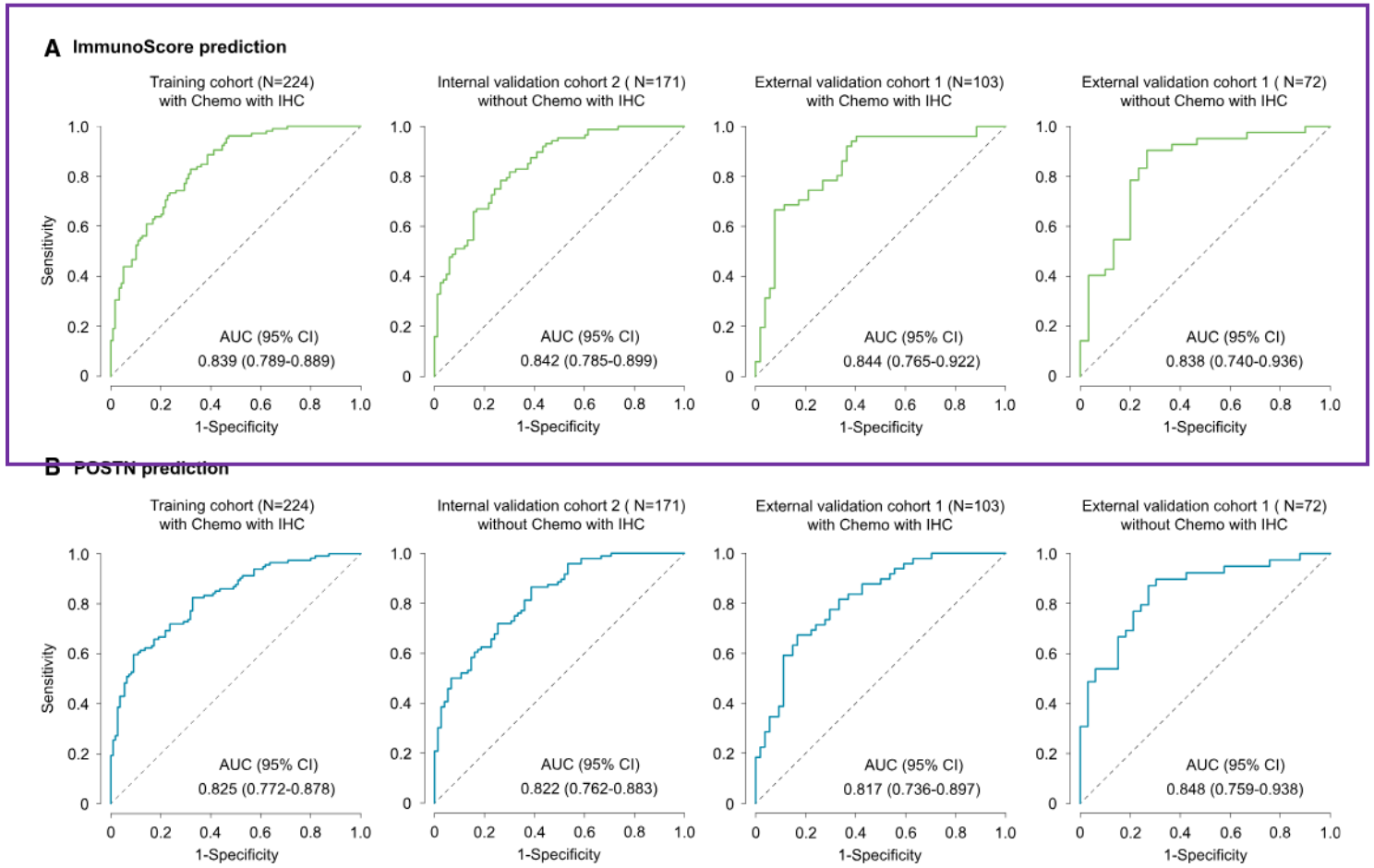
实验结果-化疗反应预测模型的准确性

附件 Chemo-Response prediction



混淆矩阵显示，模型预测与通过无病生存期（DFS）<或≥2年分类定义的实际化疗反应一致。

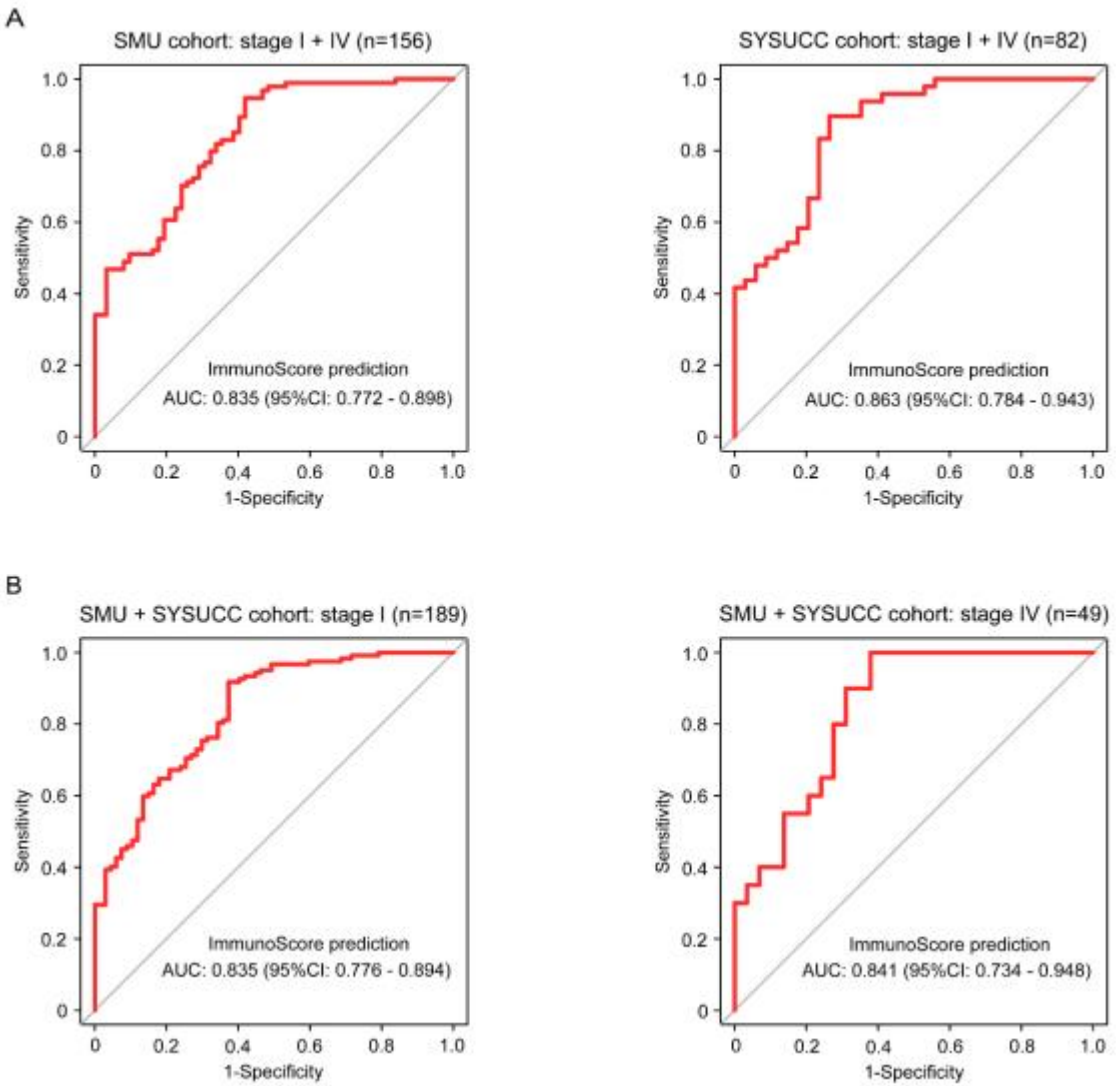
实验结果-免疫评分（*ImmunoScore*）的准确性



所提议的模型在预测免疫评分方面表现出一致的高准确性，在II期和III期患者中的曲线下面积（AUC）分别为0.839（95%置信区间[CI]：0.789－0.889）和0.842（95% CI：0.785－0.899）。在SYSUCC队列中也观察到了类似的结果，AUC分别为0.844（95% CI：0.765－0.922）和0.838（95% CI：0.740－0.936）。

实验结果-免疫评分（*ImmunoScore*）的准确性

附件

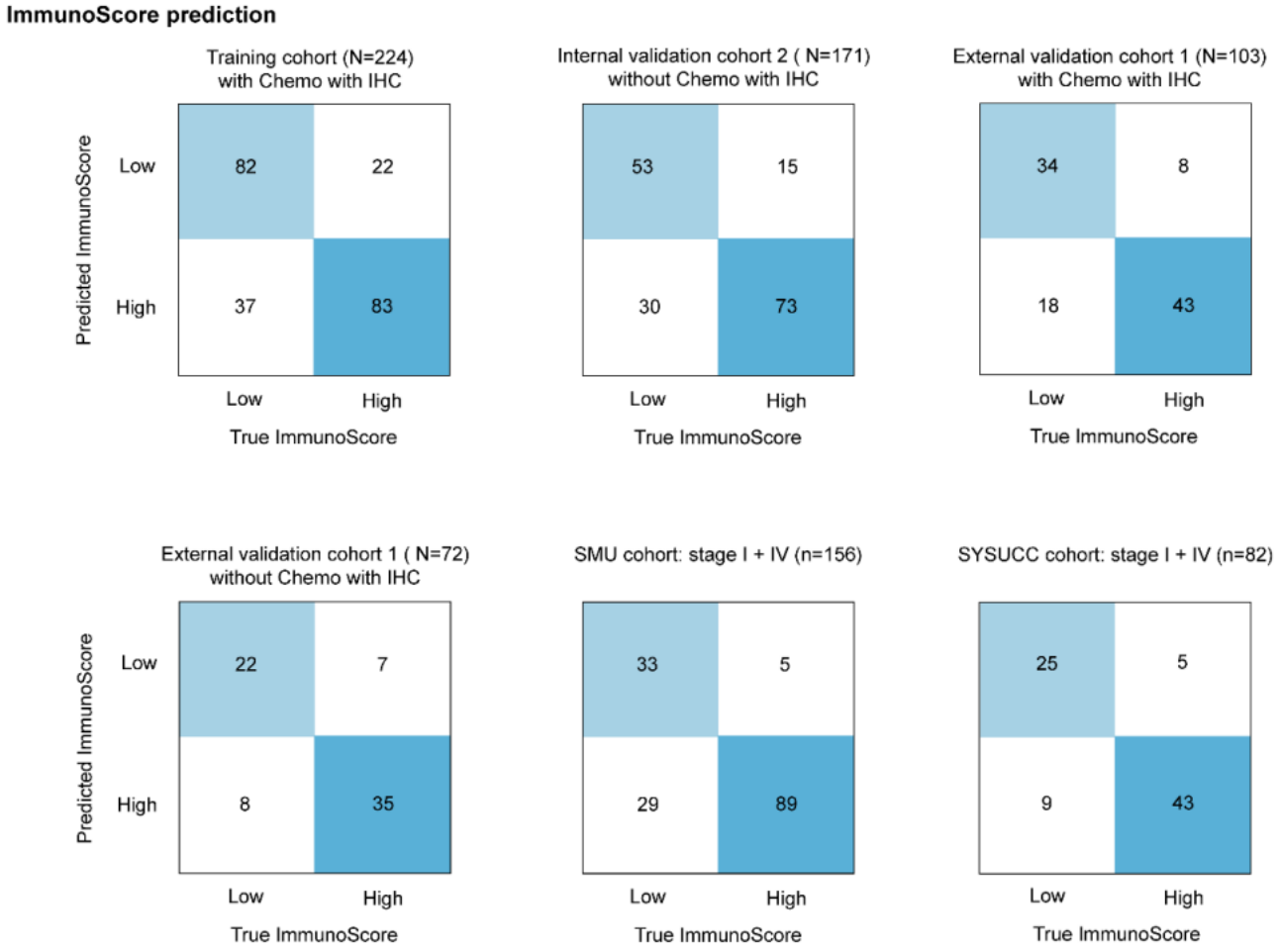


我们还探讨了该模型在I期和IV期患者中的潜力。I期和IV期患者的AUC分别为0.835（95% CI：0.776－0.894）和0.841（95% CI：0.734－0.948）

深度学习免疫评分预测模型与临床病理变量之间的关联见表S6和S7。灵敏度为79%－95%。阳性预测值和阴性预测值为71%－87%（见表S8）。

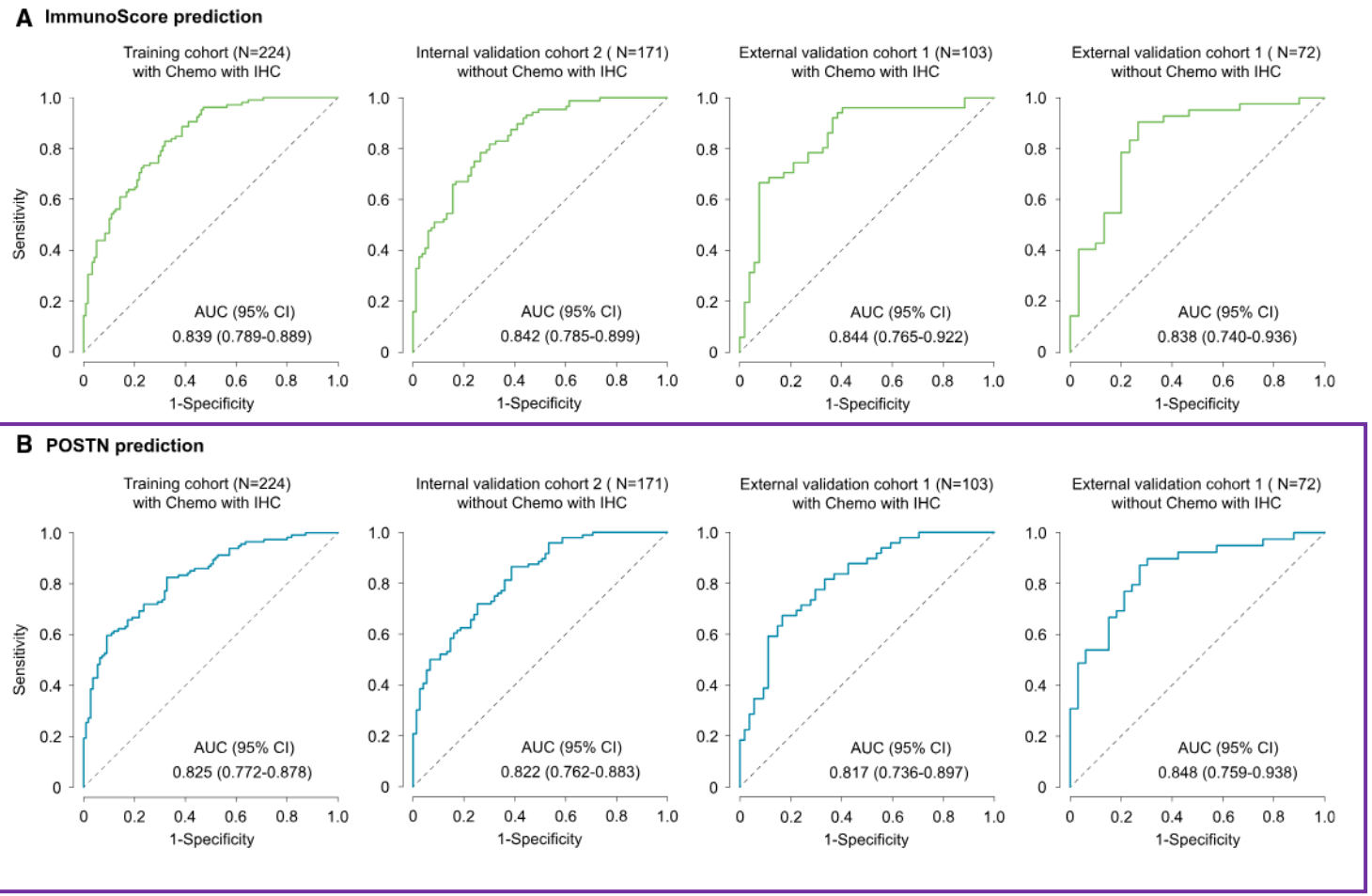
实验结果-免疫评分（*ImmunoScore*）的准确性

附件



混淆矩阵显示，模型预测与实际免疫评分类别一致

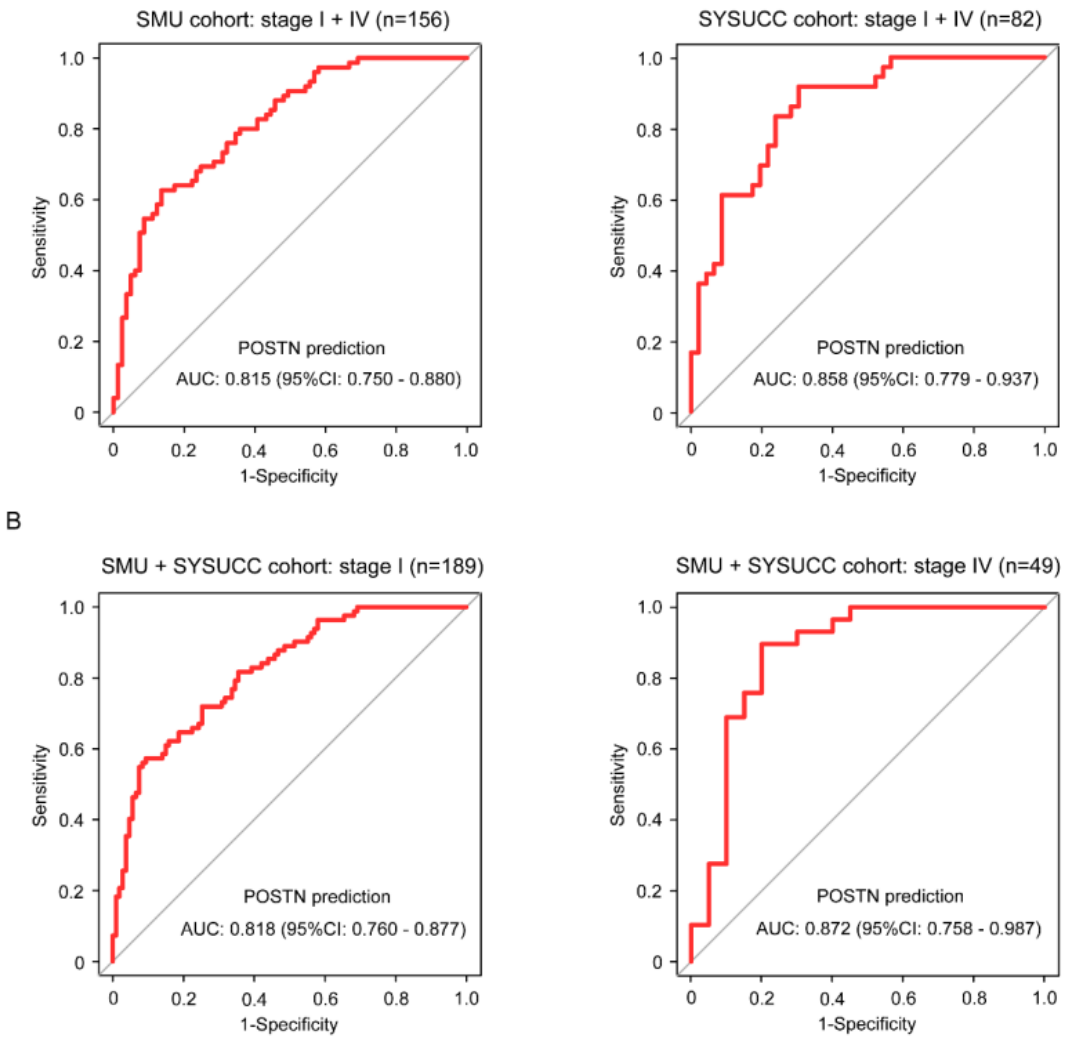
实验结果-*POSTN*预测的准确性



在预测POSTN时，所提出的模型展示了持续较高的准确性，在II期和III期患者中的曲线下面积（AUC）分别为0.825（95%CI：0.772－0.878）和0.822（95% CI：0.762－0.883）。在SYSUCC队列中也观察到了类似的结果，AUC分别为0.817（95% CI：0.736－0.897）和0.848（95% CI：0.759－0.938）。

实验结果- *POSTN*预测的准确性

附件

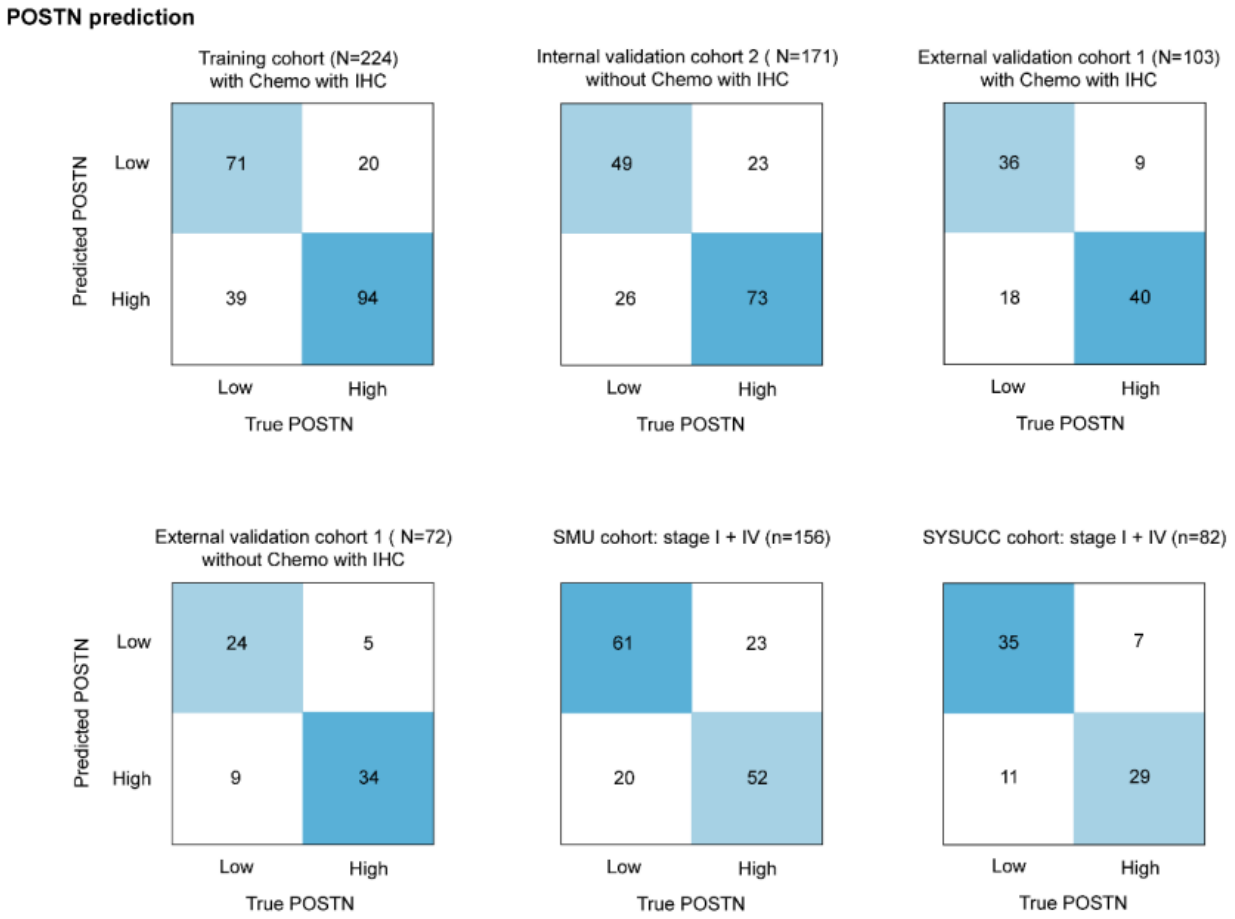


在I期和IV期患者中的AUC分别为0.818（95% CI：0.760－0.877）和0.872（95% CI：0.758－0.987）

深度学习POSTN预测模型与临床病理变量之间的关联见表S9和S10。灵敏度和特异性分别为65%－75%，阳性预测值或阴性预测值为68%－83%（见表S11）

实验结果- *POSTN*预测的准确性

附件



如混淆矩阵所示，POSTN类别预测的整体准确率为71%－81%

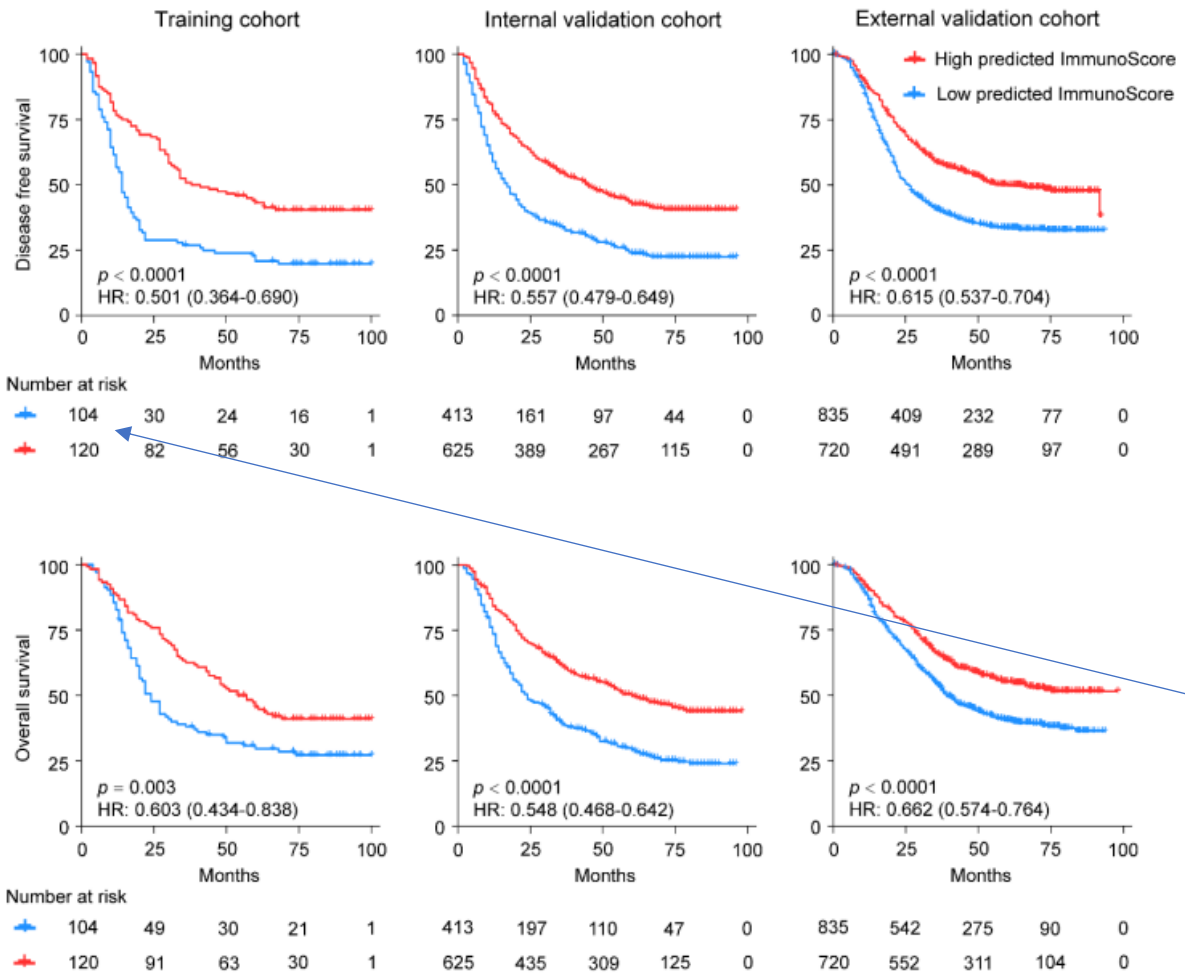
实验结果-TME预测模型预后价值

免疫评分的预后价值

Kaplan-Meier生存曲线

作者继续探究了TME对预后是否有价值（俩个子任务ImmunoScore得分/POSTN表达 和预后的关系）

附件



ImmunoScore的TME预测模型（风险比[HR]：0.501–0.662，p < 0.005）在各个队列中均与生存结局显著相关。

预测为高ImmunoScore组的5年无病生存期（DFS）和总生存期（OS）分别为40.1%和45.7%，显著高于低ImmunoScore组的23.2%和28.4%（p < 0.001）。

1. 风险比（HR）

风险比（**Hazard Ratio，HR**）是用来衡量**不同组之间**在某一特定时间点发生事件（如死亡、复发等）的**相对风险**的指标。

HR离1越远越显著，P越小越有统计意义

如何理解HR：

- **HR = 1**：表示两组之间的风险是一样的，即没有差异。
- **HR > 1**：表示**危险（事件发生的概率）**较高的一组，比参考组风险大。例如：HR = 2 意味着某组的风险是参考组的2倍。
- **HR < 1**：表示**危险较低的一组**，比参考组风险小。例如：HR = 0.5 意味着某组的风险是参考组的一半。

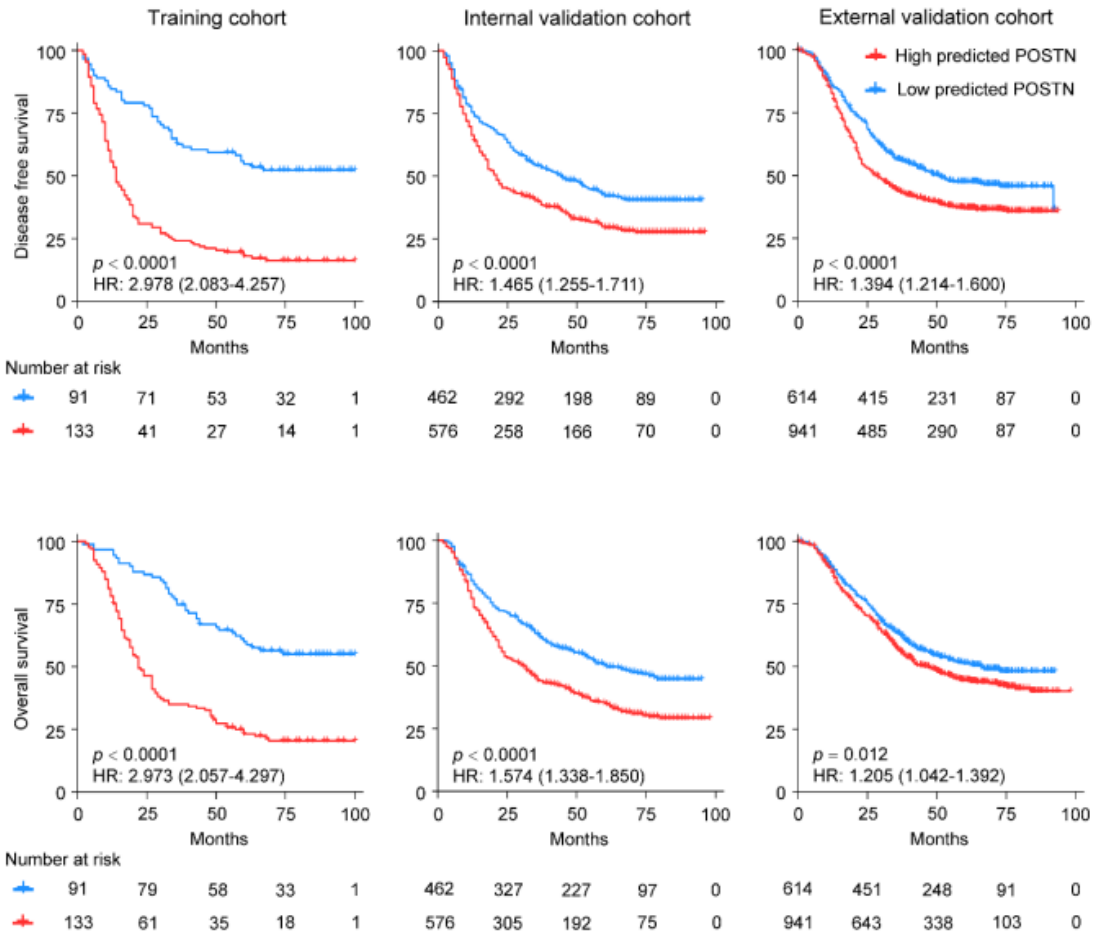
实验结果-TME预测模型预后价值

POSTN表达的预后价值

Kaplan-Meier生存曲线

作者继续探究了TME对预后是否有价值（俩个子任务ImmunoScore得分/POSTN表达 和预后的关系）

附件



POSTN的TME预测模型（HR：1.205－2.978，p < 0.05）在各个队列中均与生存结局显著相关。

在低POSTN组中，5年DFS和OS分别为39.8%和45.1%，也显著优于高POSTN组的26.4%和31.8%（p < 0.001）。

实验结果-TME预测模型预后价值

POSTN表达的预后价值

作者继续探究了TME对预后是否有价值（俩个子任务ImmunoScore得分/POSTN表达 和预后的关系）

附件

	A	B	C	D	E	F
4	**Training cohort**					
5	**Predicted TME biomarkers**					
6	Predicted ImmunoScore (high vs. low)	0.567 (0.406-0.791)	0.001		0.705 (0.503-0.989)	0.043
7	Predicted POSTN (high vs. low)	2.539 (1.758-3.666)	<0.0001		2.642 (1.812-3.853)	<0.0001
8	Gender (male vs. female)	1.514 (1.044-2.196)	0.029		1.393 (0.950-2.042)	0.09
9	Tumor size (>4 cm vs. ≤4 cm)	1.412 (1.009-1.974)	0.044		1.304 (0.923-1.844)	0.133
10	TNM stage (III vs. II)	2.554 (1.702-3.832)	<0.0001		2.956 (1.897-4.605)	<0.0001
11	**Internal validation cohort 1**					
12	**Predicted TME biomarkers**					
13	Predicted ImmunoScore (high vs. low)	0.729 (0.615-0.866)	<0.001		0.746 (0.625-0.891)	0.001
14	Predicted POSTN (high vs. low)	1.128 (0.952-1.337)	0.165		1.235 (1.034-1.474)	0.02
15	Age (years) (≥60 vs. <60)	1.009 (0.925-1.101)	0.842		1.037 (0.946-1.135)	0.439
16	Tumor size (>4 cm vs. ≤4 cm)	0.939 (0.801-1.102)	0.441		0.924 (0.782-1.091)	0.35
17	Differentiation	1.114 (0.999-1.243)	0.053		1.191 (1.060-1.338)	0.003
18	CEA (elevated vs. normal)	1.549 (1.242-1.932)	<0.001		1.556 (1.236-1.959)	<0.001
19	CA19-9 (elevated vs. normal)	1.347 (1.088-1.666)	0.006		1.274 (1.021-1.589)	0.032
20	TNM stage (IV vs. III vs. II vs. I)	1.841 (1.618-2.095)	<0.0001		1.780 (1.561-2.030)	<0.0001
21	Chemotherapy (yes vs. no)	0.732 (0.624-0.860)	<0.001		0.693 (0.586-0.820)	<0.001
22	**External validation cohort 1**					
23	**Predicted TME biomarkers**					
24	Predicted ImmunoScore (high vs. low)	0.709 (0.614-0.820)	<0.0001		0.732 (0.627-0.854)	<0.0001
25	Predicted POSTN (high vs. low)	1.188 (1.025-1.376)	0.022		1.017 (0.871-1.188)	0.827
26	Tumor size (>4 cm vs. ≤4 cm)	0.988 (0.862-1.131)	0.858		0.967 (0.836-1.117)	0.647
27	Differentiation	_	_		0.999 (0.872-1.145)	0.989
28	Lauren type (diffuse or mixed vs. intestinal)	1.191 (1.027-1.381)	0.021		1.175 (0.997-1.386)	0.054
29	CEA (elevated vs. normal)	1.118 (0.951-1.313)	0.177		1.168 (0.987-1.383)	0.071
30	CA19-9 (elevated vs. normal)	1.331 (1.138-1.557)	<0.001		1.357 (1.152-1.600)	<0.001
31	TNM stage (IV vs. III vs. II vs. I)	1.802 (1.574-2.063)	<0.0001		1.919 (1.658-2.220)	<0.0001
32	Chemotherapy (yes vs. no)	0.890 (0.778-1.017)	0.086		0.702 (0.610-0.809)	<0.001
33	TME, tumor microenvironment.					

多变量分析结果进一步确认，在调整其他临床病理变量后，TME预测模型仍然是临床结局的独立预测因素（见表S12和S13）。

实验结果-TME预测模型预后价值

POSTN表达的预后价值

作者继续探究了TME对预后是否有价值（俩个子任务ImmunoScore得分/POSTN表达 和预后的关系）

附件

Table S14. Comparing the prediction accuracy of the integrated nomogram with predicted TME biomarkers and TNM stage in the training and validation cohorts.

	Disease-free survival		Overall survival	
Variable	**C-index (95% CI)**	***P***	**C-index (95% CI)**	***P***
Training cohort		<0.0001		<0.0001
Nomogram	0.726 (0.690-0.762)		0.723 (0.683-0.763)	
Predicted ImmunoScore	0.644 (0.604-0.684)		0.633 (0.588-0.678)	
Predicted POSTN	0.653 (0.609-0.697)		0.662 (0.617-0.707)	
TNM stage	0.639 (0.593-0.685)		0.640 (0.591-0.689)	
Internal validation cohort 1		<0.0001		<0.0001
Nomogram	0.675 (0.654-0.696)		0.667 (0.646-0.688)	
Predicted ImmunoScore	0.598 (0.575-0.621)		0.599 (0.575-0.623)	
Predicted POSTN	0.560 (0.537-0.583)		0.567 (0.543-0.591)	
TNM stage	0.660 (0.639-0.681)		0.653 (0.631-0.675)	
External validation cohort 1		<0.0001		<0.0001
Nomogram	0.647 (0.628-0.666)		0.652 (0.632-0.672)	
Predicted ImmunoScore	0.591 (0.571-0.611)		0.570 (0.549-0.591)	
Predicted POSTN	0.551 (0.530-0.572)		0.522 (0.500-0.544)	
TNM stage	0.623 (0.604-0.642)		0.635 (0.615-0.655)	

TME, tumor microenvironment.

将影像特征与肿瘤-淋巴结-转移（TNM）分期相结合的模型，在所有队列中相比单独使用TNM分期，均能显著提高预后预测的准确性（$p < 0.0001$）。

“传统分期 + AI模型” 比单用“传统分期”更能准确预测谁的病情更容易复发或恶化。

实验结果-与化疗获益的关联

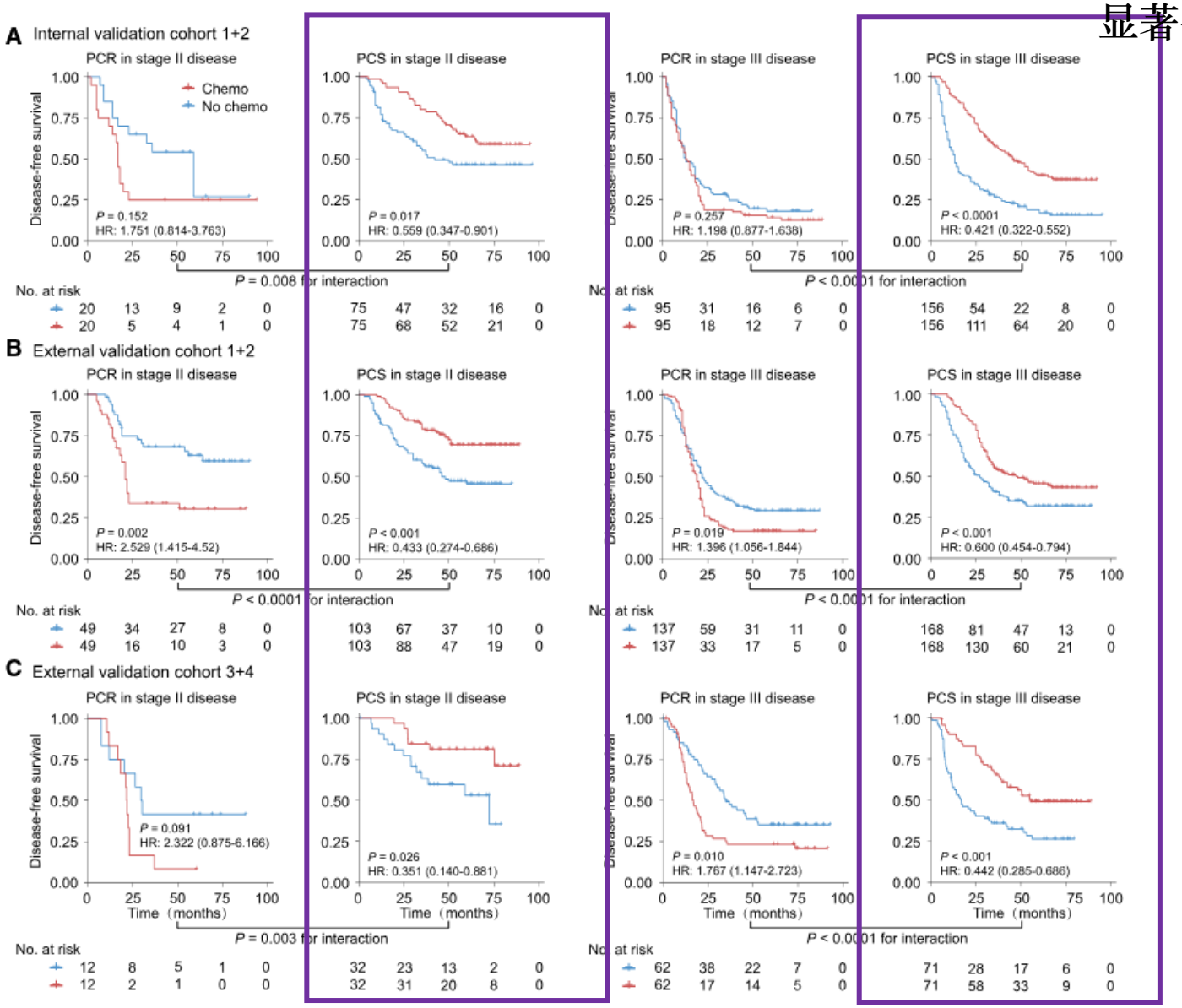
我们进一步评估了深度学习模型在预测II期和III期患者接受辅助化疗后的生存获益方面的价值。首先，我们进行了倾向评分匹配（PSM）分析，以平衡接受化疗和未接受化疗患者的特征。通过1:1的患者匹配策略，我们实现了两组之间的临床特征相似性。随后，我们比较了根据预测的化疗反应状态（预测化疗敏感（PCS）或预测化疗耐药（PCR））分组的患者，接受或未接受化疗的生存结果。

1. **倾向评分匹配（PSM）分析**
   - 为了让**化疗组和未化疗组**之间更具可比性，研究者使用了**倾向评分匹配**方法（PSM）。
   - 这种方法的目的是平衡两组患者的特征，比如年龄、性别、病情等，确保比较时两组在这些方面没有太大差异。
   - 通过1:1的匹配方式（每个化疗组的患者都找一个相似的未化疗组患者），最终确保了两组的**临床特征相似**，从而减少了干扰因素。
2. **根据化疗反应分组（PCS和PCR）**
   - 然后，研究人员根据深度学习模型预测的化疗反应分组：
     - **PCS（预测化疗敏感）组**：模型预测这些患者对化疗敏感。
     - **PCR（预测化疗耐药）组**：模型预测这些患者对化疗耐药。
   - 他们比较了这两组患者的生存情况，看是否化疗带来了生存期的改善。

实验结果-与化疗获益的关联

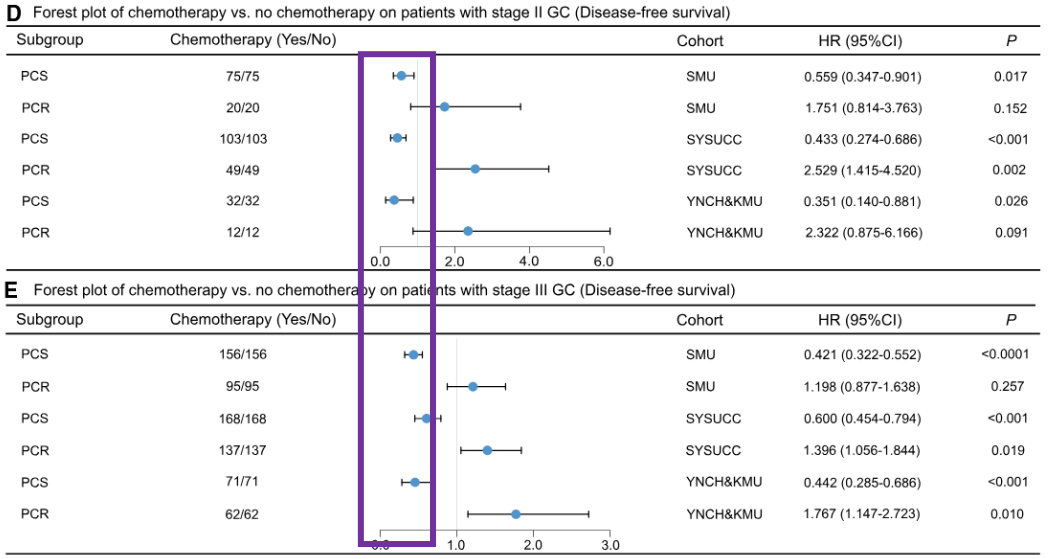
Kaplan-Meier生存曲线

显著分离



对于PCS组的患者，辅助化疗与无病生存期（DFS）（HR范围：0.351－0.600，$p < 0.05$）和总生存期（OS）（HR范围：0.186－0.553，$p < 0.05$）改善相关。

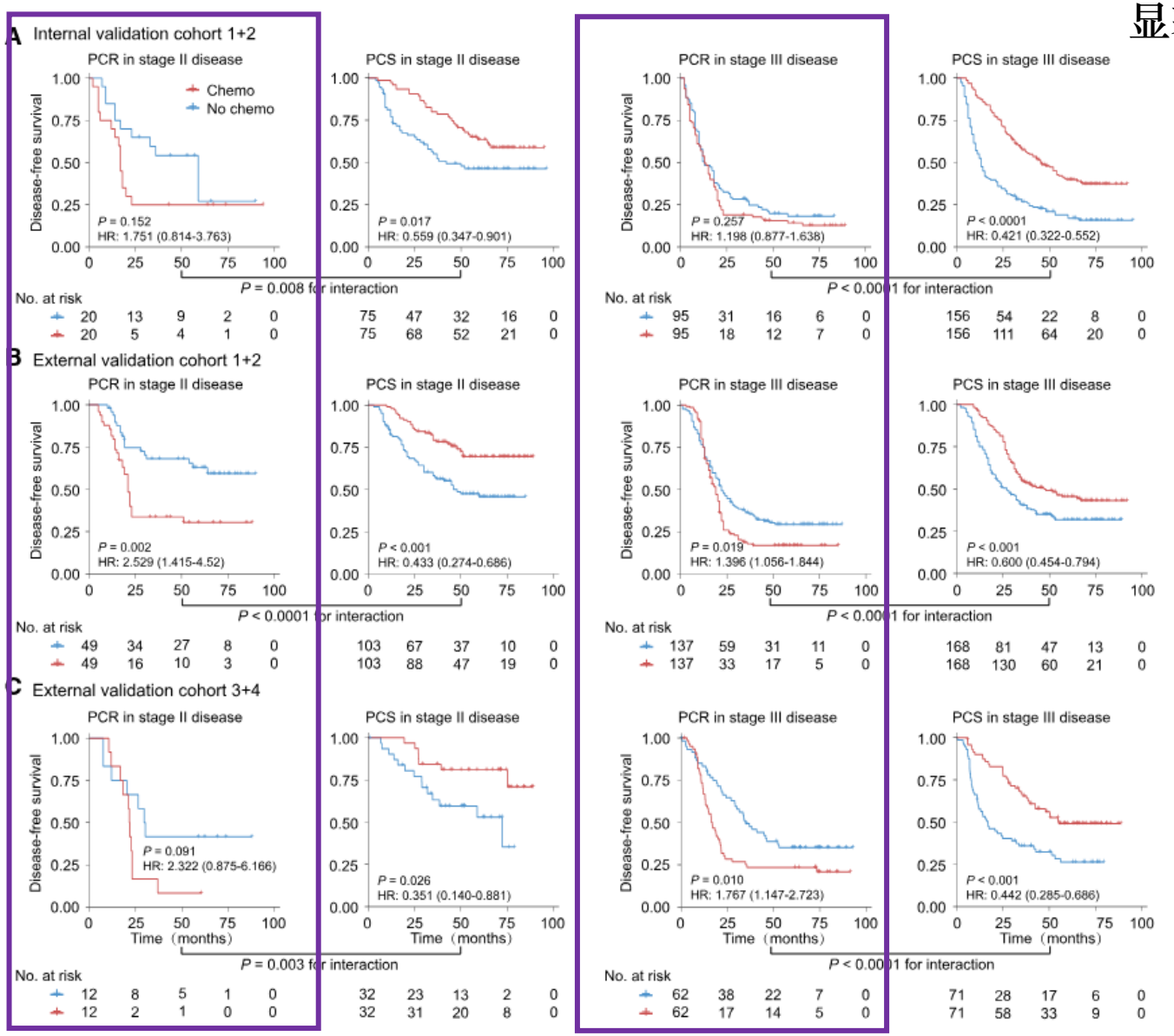
数据集中



实验结果-与化疗获益的关联

Kaplan-Meier生存曲线

显著交错

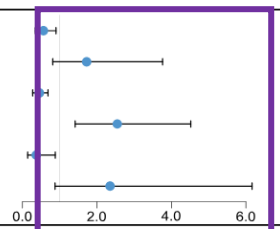


然而，对于PCR组的患者，辅助化疗与DFS和OS没有显著改善，甚至可能对生存产生不利影响。

数据分散

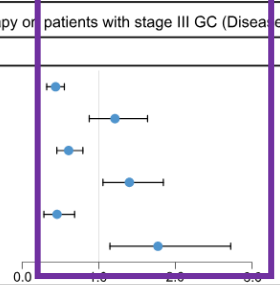
D Forest plot of chemotherapy vs. no chemotherapy on patients with stage II GC (Disease-free survival)

Subgroup	Chemotherapy (Yes/No)	Cohort	HR (95%CI)	P
PCS	75/75	SMU	0.559 (0.347-0.901)	0.017
PCR	20/20	SMU	1.751 (0.814-3.763)	0.152
PCS	103/103	SYSUCC	0.433 (0.274-0.686)	<0.001
PCR	49/49	SYSUCC	2.529 (1.415-4.520)	0.002
PCS	32/32	YNCH&KMU	0.351 (0.140-0.881)	0.026
PCR	12/12	YNCH&KMU	2.322 (0.875-6.166)	0.091



E Forest plot of chemotherapy vs. no chemotherapy on patients with stage III GC (Disease-free survival)

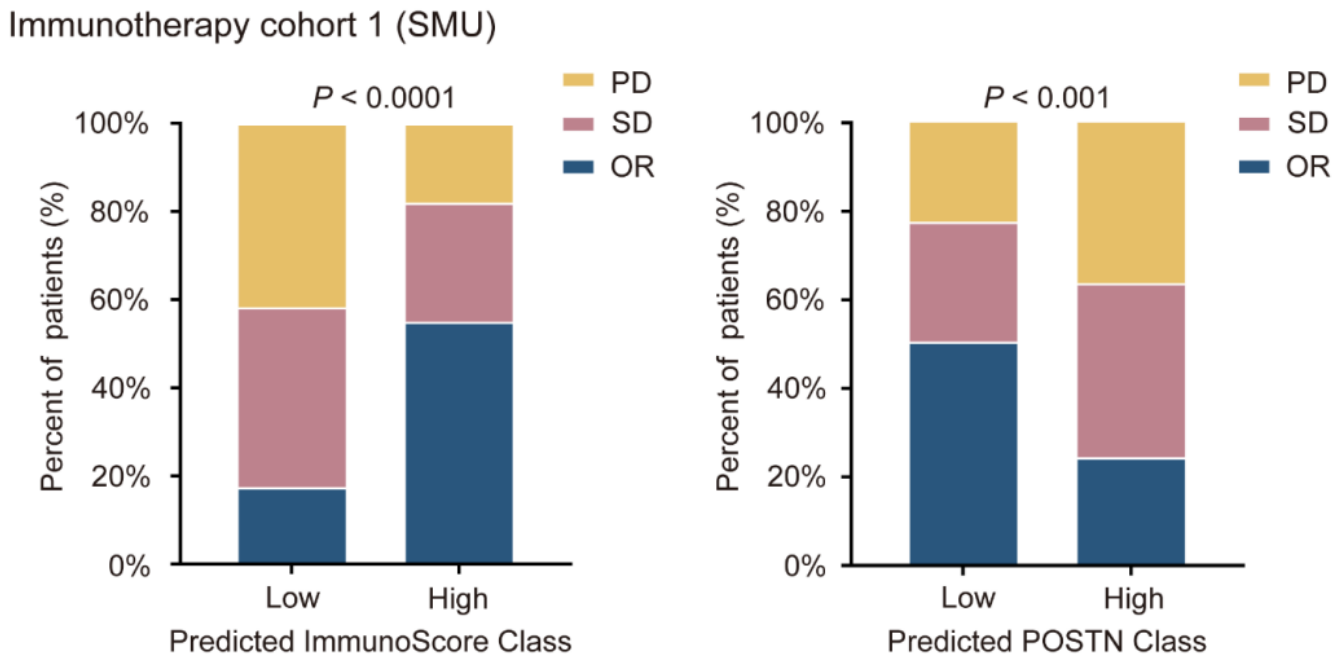
Subgroup	Chemotherapy (Yes/No)	Cohort	HR (95%CI)	P
PCS	156/156	SMU	0.421 (0.322-0.552)	<0.0001
PCR	95/95	SMU	1.198 (0.877-1.638)	0.257
PCS	168/168	SYSUCC	0.600 (0.454-0.794)	<0.001
PCR	137/137	SYSUCC	1.396 (1.056-1.844)	0.019
PCS	71/71	YNCH&KMU	0.442 (0.285-0.686)	<0.001
PCR	62/62	YNCH&KMU	1.767 (1.147-2.723)	0.010

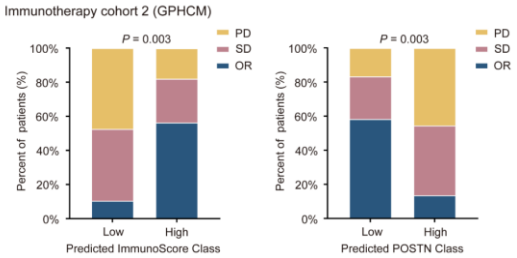


实验结果-TME与抗PD-1免疫疗法反应的关联

本研究随后评估了深度学习肿瘤微环境（TME）生物标志物与两组独立队列中免疫疗法反应的关系。

附件





疾病稳定	SD（Stable Disease）	肿瘤变化不大（没有明显缩小或增大）
疾病进展	PD（Progressive Disease）	肿瘤增大 ≥ 20% 或出现新病灶

有趣的是，我们发现，预测为高ImmunoScore组（55.0%和56.4%）或低POSTN组（50.6%和58.3%）的患者，其客观反应率（OR）高于预测为低ImmunoScore组（17.5%和10.5%）或高POSTN组（24.5%和13.6%）的患者。这个结果在整个免疫治疗队列中也得到了验证（$p < 0.0001$）

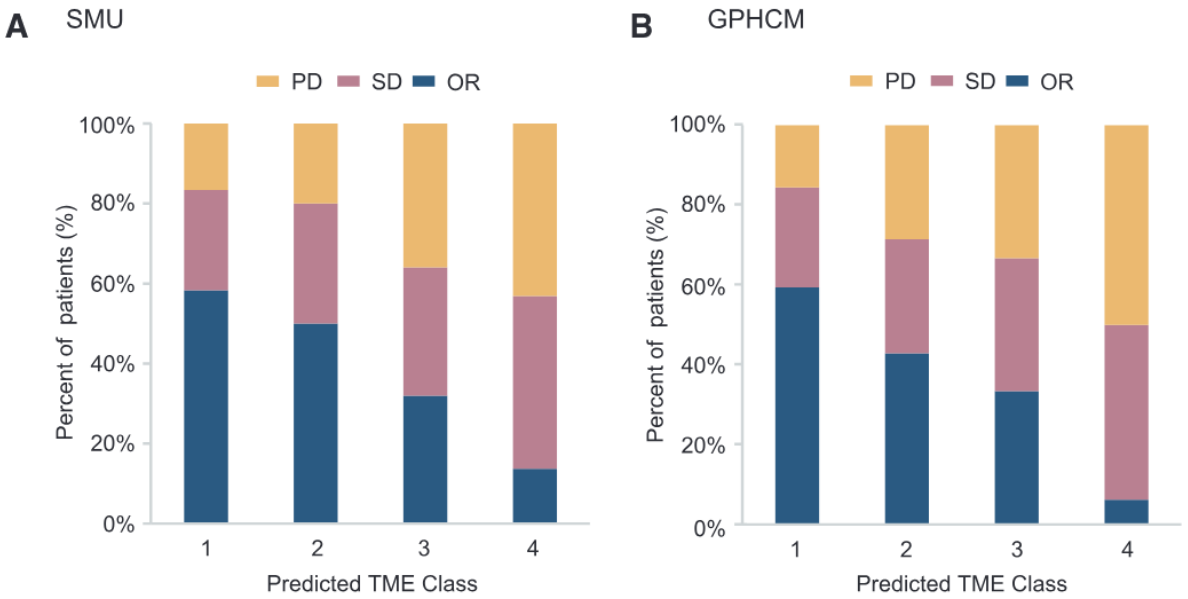
客观反应率（OR） 指的是：

在接受治疗的患者中，**肿瘤体积出现明显缩小（达到一定标准）** 的患者所占的比例。

实验结果-TME与抗PD-1免疫疗法反应的关联

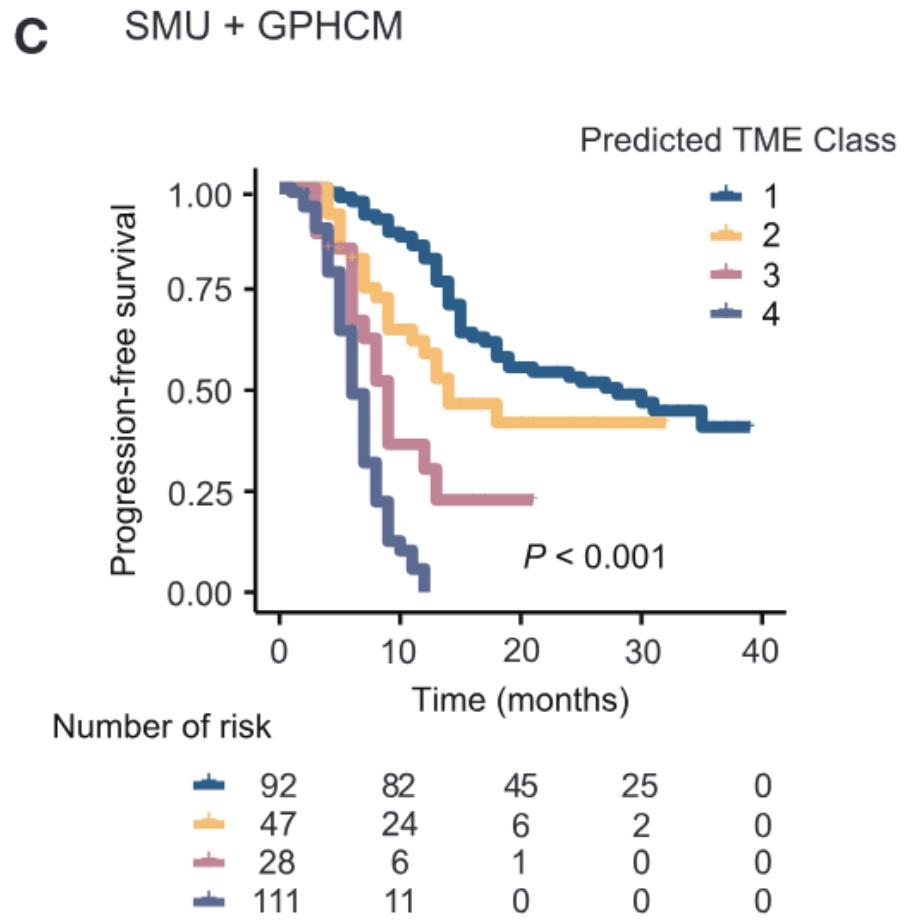
接下来，我们将预测的ImmunoScore和POSTN生物标志物整合到一个TME模型中。我们观察到不同TME类别之间的异质性客观反应率：

预测TME类别1（高ImmunoScore和低POSTN）：58.7%
预测TME类别2（高ImmunoScore和高POSTN）：48.9%
预测TME类别3（低ImmunoScore和低POSTN）：32.1%
预测TME类别4（低ImmunoScore和高POSTN）：12.6%
这些结果也在各免疫治疗队列中得到了确认（$p < 0.0001$）



实验结果-TME与抗PD-1免疫疗法反应的关联

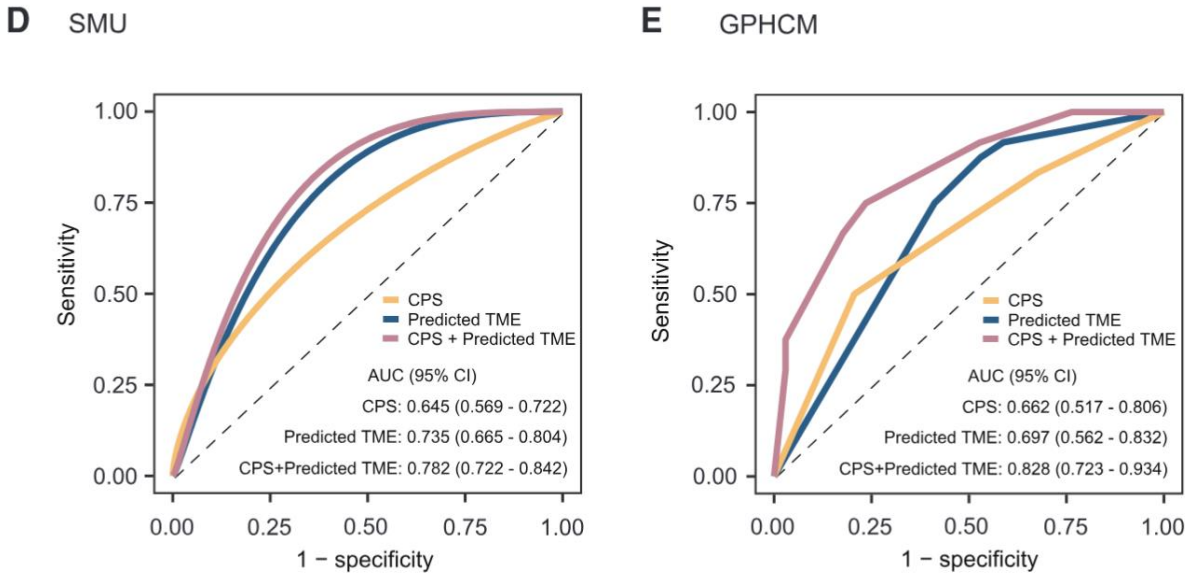
Kaplan-Meier生存曲线



Kaplan-Meier生存曲线验证了两种预测生物标志物（ImmunoScore或POSTN）以及预测TME模型的预后价值（$p < 0.001$）。

实验结果-TME与抗PD-1免疫疗法反应的关联

我们接着比较了预测TME模型和PD-L1在预测免疫疗法反应方面的表现。

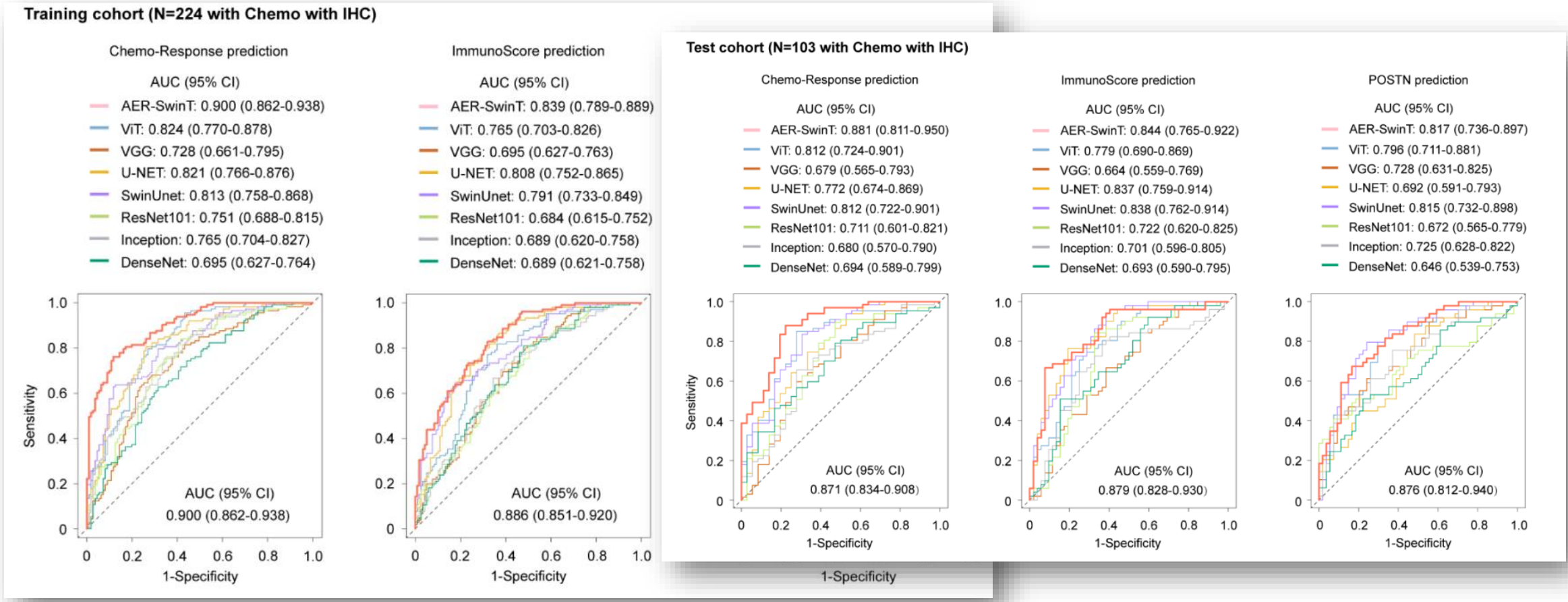


PD-L1表达的联合阳性评分（CPS）在预测免疫疗法反应方面的能力较为有限，在免疫治疗队列1和2中的AUC分别为0.645（95% CI：0.569－0.722）和0.662（95% CI：0.517－0.806）。然而，预测TME模型在免疫治疗队列中的AUC为0.735（95% CI：0.665－0.804）和0.697（95% CI：0.562－0.832），高于CPS。

重要的是，当将CPS与预测的TME模型结合成一个整合模型时，相比CPS，免疫疗法反应预测的准确性显著提高（AUC：0.782－0.828，$p < 0.05$）。

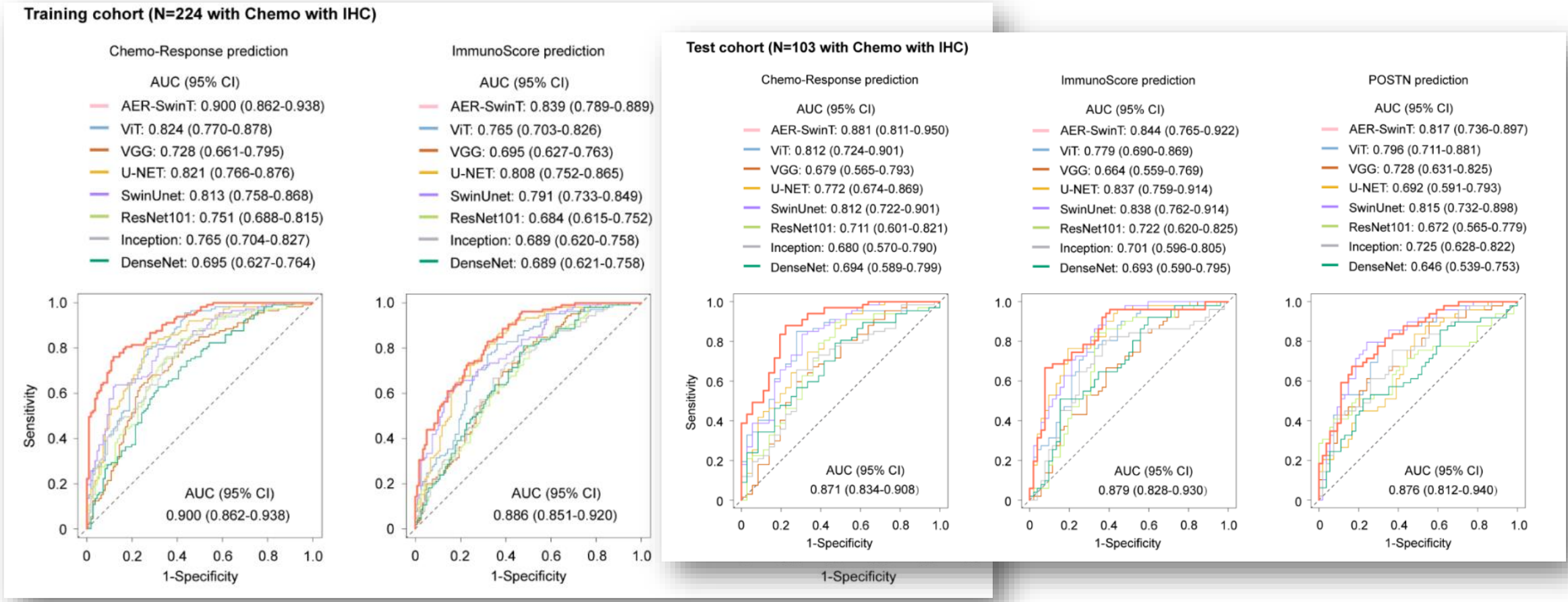
就是吹预测TME比CPS方法好，更有益于预测免疫疗法反应

实验结果-对比实验



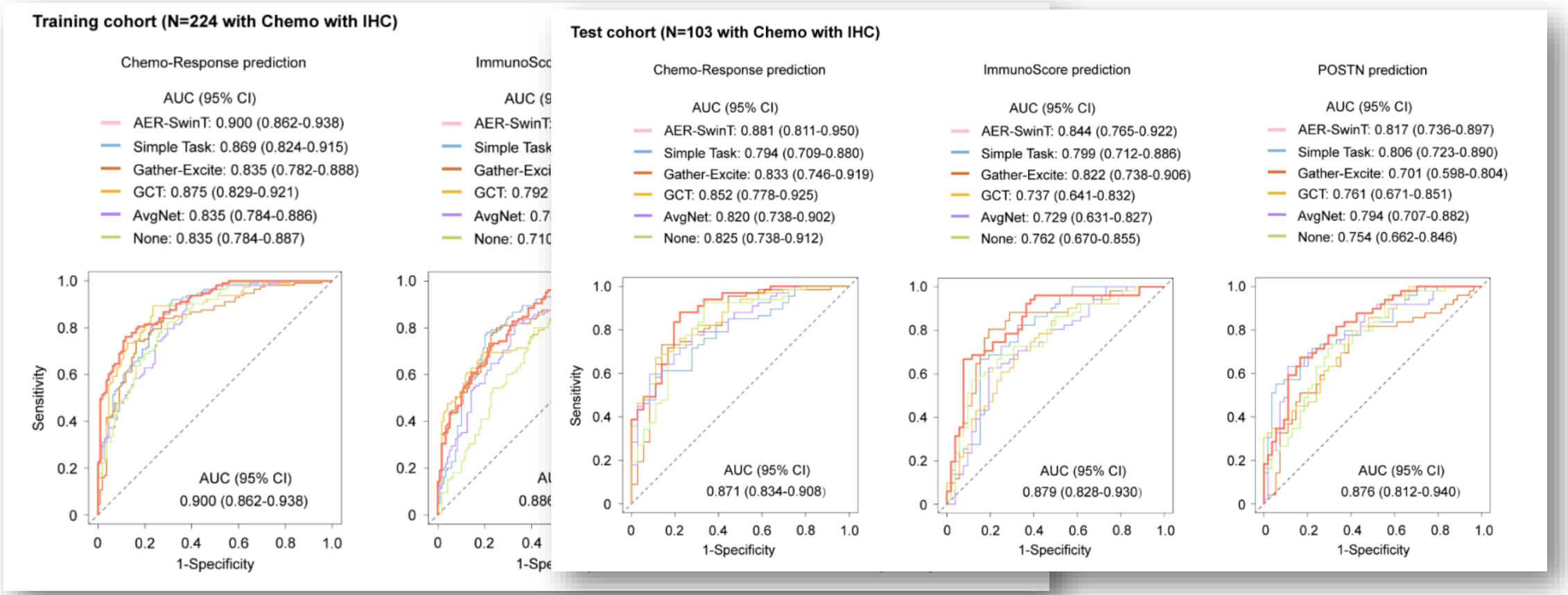
为了全面评估我们提出的AER-SwinT模型，我们进行了额外的对比实验，涉及使用相同数据集的基于CNN的方法（Unet、VGG、ResNet、Inception和DenseNet）以及基于Transformer的方法（ViT和SwinUnet）。

实验结果-对比实验



AER-SwinT在内部和外部数据集上始终优于基准模型。值得注意的是，基于Transformer的方法通常表现优于基于CNN的模型，突显了在数据中捕捉全局关系以增强分析能力的优势。

实验结果-消融实验



我们比较了仅用于化疗反应预测的单任务版本AER-SwinT与我们的多任务训练方法。图S17显示，采用多任务训练策略相比单任务训练显著提高了性能，强调了结合临床相关子任务有助于更具区分性的特征提取。另外，我们评估了基于通道平均的注意力图模块（SENet），该模块通过残差聚合实现通道注意力。我们还在模型中评估了其他聚合方法，如Gatherexcite、GCT和AvgNet。结果表明，基于SENet的设计达到了最高的性能。

Discussion

第一段：本文工作意义总结

Immune checkpoint blockade and adjuvant chemotherapy have shown quite durable anti-tumor responses, but the lack of effective biomarkers limits the therapeutic benefits.[4,14] These challenges highlight the importance of developing cost-effective and accurate prediction models for chemotherapy and immunotherapy. In this retrospective multi-center study of 3,095 patients, we developed a multitask deep learning model called AER-SwinT for the simultaneous prediction of chemotherapy response after radical surgery and the TME (ImmunoScore and POSTN expression) using preoperative CT images. By employing attention-enhanced hierarchical features and multitask learning, AER-SwinT accurately predicts chemotherapy response and TME characteristics. For patients predicted as chemo-sensitive through the proposed deep learning model, adjuvant chemotherapy was associated with improved survival, whereas those predicted as chemo-resistant did not. Furthermore, given the established clinical relevance of TME biomarkers, the multitask deep learning model showed predictive value for prognosis and could distinguish patients who might or might not benefit from anti-PD-1 immunotherapy.

- 提出了AER-SwinT的多任务深度学习模型。
- 该模型使用术前CT图像，同时预测化疗反应和肿瘤微环境（TME）特征（包括免疫评分和POSTN表达）。
- ER-SwinT采用了增强注意力的层次化特征和多任务学习，能够准确预测化疗反应和TME特征。
- 【化疗反应预测】对于化疗敏感的患者，模型预测结果表明辅助化疗与生存率提高相关。对于化疗耐药的患者，模型预测结果表明辅助化疗无显著生存效益。
- 【TME预测价值】模型考虑了TME生物标志物的临床相关性，显示了对预后的预测价值。免疫治疗预测：该模型能够区分哪些患者可能从抗PD-1免疫治疗中获益，哪些则可能不获益。

Discussion

第二段：强调化疗反应预测任务

Systemic chemotherapy is considered as a standard treatment for local advanced GC in Western or Eastern nations.[1,2] Adding chemotherapy to surgery has improved patient survival. However, larger variation in clinical outcomes and treatment responses is observed even among patients with same clinicopathologic features and similar treatment regimens.[3] Given these heterogeneous outcomes, accurate prediction of chemotherapy response after gastrectomy is crucial for making appropriate treatment decisions about adjuvant therapies.[4] To address these challenges, several biomarkers associated with chemotherapy response had been reported. Yong and colleagues developed a gene signature to predict response to 5-fluorouracil-based chemotherapy on a phase 2 clinical trial with 81 patients with GC.[53] Chen and colleagues found that pathomics signature was associated with chemotherapy response of GC in two cohorts with 480 patients.[54] However, these studies included a relatively small number of patients, and the methodology used was invasive and costly. By analyzing data for 3,095 patients, this study developed a non-invasive deep learning model based on preoperative CT images to directly predict response to 5-fluorouracil-based chemotherapy in resected GC. Our method can identify patients who can benefit from postoperative adjuvant chemotherapy in stage II and III GC. Patients identified as chemo-sensitive are likely to experience significant survival benefits from adjuvant chemotherapy, whereas those identified as chemo-resistant do not. The prediction model will allow the optimization of individual decision-making. By utilizing AER-SwinT, we can predict that sensitive patients would benefit from aggressive treatment regimens to improve survival, while resistant patients could be spared from the adverse effects of adjuvant therapies.

- 系统化疗被认为是局部晚期胃癌（GC）的标准治疗，但即使是具有相同临床特征的患者，治疗反应和临床结果也存在较大变异性。准确预测胃切除术后的化疗反应对于做出合理的辅助治疗决策至关重要。
- （现有方法概述）现有的研究尝试使用生物标志物（如基因签名和病理组学签名）预测化疗反应，但这些方法通常侵入性强、成本高且研究规模较小。
- （作者的方法）ER-SwinT该模型能够识别出可能从术后辅助化疗中受益的化疗敏感患者，以及那些从中获益较小的化疗耐药患者。
- （临床意义）通过使用AER-SwinT模型，医生可以优化个体化治疗决策，确保化疗敏感的患者接受更激进的治疗，而化疗耐药的患者则避免不必要的副作用。

Discussion

第三段：强调TME预测价值和免疫治疗反应预测

The recent success of immunotherapy across various cancer types highlights the need to better understand the mechanisms of effective anti-tumor immune responses and to identify the immunotherapy benefit for individual patient.[6–8,10,11,55] The TME is cumulatively recognized as the key regulator of all types of anti-cancer therapy.[19–21] Compared with chemotherapy non-responders, response was associated with on-treatment TME remodeling including natural killer cell recruitment, decreased tumor-associated macrophages, M1-macrophage repolarization, and increased effector T cell infiltration.[23] Thus, assessing the TME at treatment will assist decision-making of adjuvant chemotherapy, and introducing TME information in prediction model development for chemotherapy response is of great value. Additionally, increasing evidence established the role of TME as a determinant for predicting prognosis and immunotherapy responses in various cancer.[19–22] To dissect inter-tumor TME heterogeneity, we defined two biomarkers (ImmunoScore and POSTN) with established biological and clinical relevance.[21,24] The proposed deep learning model enables non-invasive, economical, accurate, and dynamic monitoring of the TME. Given the limitations in tissue access, as well as the high cost, time-consuming nature, and technical complexity of histological approaches, our method offers a viable alternative to overcome these challenges.

- （免疫反应识别）免疫治疗在多种癌症中的成功突显了对有效抗肿瘤免疫反应机制的深入理解的重要性，并且需要识别个体患者是否能从免疫治疗中获益。
- （TME与化疗反应关联）TME被认为是所有抗癌治疗的关键调节因子。治疗反应与TME的重塑相关，包括自然杀伤细胞的招募、巨噬细胞的变化以及T细胞的浸润增加。在治疗中评估TME有助于化疗决策，且将TME信息整合入预测模型能提高化疗反应的预测精度。
- （TME与免疫反应关联）TME不仅能帮助预测患者预后，还能预测免疫治疗反应。该研究通过定义免疫评分和POSTN作为生物标志物，进一步揭示肿瘤间TME的异质性。
- 提出的深度学习模型能够非侵入性、经济、准确且动态地监测TME，提供了克服组织学方法成本高、耗时长且技术复杂性的可行替代方案。

Discussion

第四段：解释为什么没有研究 围手术期 免疫反应预测

Although immunotherapy has been recommended as a first-line treatment option for advanced GC, its application in the perioperative setting is still in the early stages of exploration, and relevant data remain limited.[5,8–13] Therefore, this study only explored the application of the developed model in immunotherapy for advanced GC. Besides, 5-fluorouracil-based chemotherapy is considered the standard of care for stage II–III GC following radical surgery.[1,2] Postoperative adjuvant chemotherapy for advanced GC and immunotherapy for advanced GC are different treatment options targeting patients at different disease stages. Thus, it is worth noting that the predictive effect of the CT-based model developed in this study for chemotherapy and immunotherapy was explored using two distinct cohorts: the former comprising patients with GC receiving postoperative adjuvant therapy, and the latter comprising patients with advanced-stage GC. In the future, as more data become available, we will further investigate the model's predictive value in the context of perioperative immunotherapy.

- 尽管免疫治疗已被推荐作为晚期胃癌（GC）的一级治疗选择，但其在围手术期的应用仍处于探索的早期阶段，相关数据仍然有限。因此，本研究仅探讨了所开发模型在晚期胃癌免疫治疗中的应用。
- 术后辅助化疗和晚期胃癌的免疫治疗针对不同的疾病阶段，且它们是不同的治疗选择。
- 因此，值得注意的是，本研究开发的基于CT的预测模型在化疗和免疫治疗方面的预测效果，分别使用了两个不同的队列：前者是接受术后辅助治疗的胃癌患者，后者是晚期胃癌患者。
- 未来，随着更多数据的获得，我们将进一步探讨该模型在围手术期免疫治疗中的预测价值。

我的理解：作者针对化疗做了围手术期的探索，即探索了化疗受益分析，但是对于免疫治疗队列，并没有做类似“免疫治疗的受益分析”，即探究接受/不接受免疫治疗的病人存活情况。

Discussion

第五段：TME模型对于免疫治疗反应预测价值

Although immunotherapy has been recommended as a first-line treatment option for advanced GC, its application in the perioperative setting is still in the early stages of exploration, and relevant data remain limited.[5,8–13] Therefore, this study only explored the application of the developed model in immunotherapy for advanced GC. Besides, 5-fluorouracil-based chemotherapy is considered the standard of care for stage II–III GC following radical surgery.[1,2] Postoperative adjuvant chemotherapy for advanced GC and immunotherapy for advanced GC are different treatment options targeting patients at different disease stages. Thus, it is worth noting that the predictive effect of the CT-based model developed in this study for chemotherapy and immunotherapy was explored using two distinct cohorts: the former comprising patients with GC receiving postoperative adjuvant therapy, and the latter comprising patients with advanced-stage GC. In the future, as more data become available, we will further investigate the model's predictive value in the context of perioperative immunotherapy.

- 考虑到TME生物标志物与免疫治疗之间的已建立关联，我们观察到所提出的深度学习模型具有改善免疫治疗反应预测的潜力。
- 尽管像PD-L1这样的生物标志物被用于评估免疫治疗反应，但其灵敏度和特异性仍然不足，需要进一步改进。
- 通过整合TME模型，免疫治疗反应的预测准确性可以得到提升，从而优化个体化的治疗决策，特别是在辅助化疗和免疫治疗方面。

Discussion

第六段：算法层面动机

Using deep learning approach to automatically learn quantitative representation from medical images is the development trend of intelligent precision diagnosis and treatment research.[28,56] Most deep learning methods for clinical tasks rely on convolutional networks (CNNs) and transformer architectures.[36–40,57–59] CNN-based approaches are inherently constrained by their limited receptive field, hampering their ability to capture comprehensive global information from medical images. Transformer-based methods, while powerful, often struggle to extract multi-granular details effectively, which is critical for developing a thorough understanding of the data and achieving accurate analysis. This study addresses these limitations by introducing AER-SwinT, a deep learning model designed to predict chemotherapy response and assess the TME using preoperative CT images. Our approach leverages the hierarchical feature extraction capabilities of the Swin Transformer, enabling the model to process data from fine to coarse granularity. This allows AER-SwinT to capture both global and detailed information crucial for accurate clinical predictions. Additionally, our method adopts

- （多粒度特征提取）使用深度学习从医学影像中自动学习特征表示是精确诊断和治疗的未来发展趋势，传统的CNN和transformer模型在捕捉全局信息和细节方面存在局限。（基于CNN的方法固有地受到其有限感受野的限制，无法有效捕捉医学影像中的全面全局信息。虽然基于变换器的方法非常强大，但通常在有效提取多粒度细节方面存在困难，这对全面理解数据并实现准确分析至关重要。）
- （多粒度特征提取）AER-SwinT是一种旨在通过术前CT图像预测化疗反应并评估肿瘤微环境（TME）的深度学习模型。我们的方法利用Swin变换器的层次化特征提取能力，使模型能够处理从细粒度到粗粒度的数据。这使得AER-SwinT能够捕捉到对准确临床预测至关重要的全局信息和细节信息。

Discussion

第六段：算法层面动机

accurate clinical predictions. Additionally, our method adopts channel average-based attention maps at multiple stages. These attention maps dynamically highlight regions of interest within the CT images, allowing the model to focus on critical spatial locations through pointwise multiplication and residual connections. This mechanism ensures that the model emphasizes important features at each stage, significantly improving its ability to capture different details and enhance overall performance. Compared to

tures at each stage, significantly improving its ability to capture different details and enhance overall performance. Compared to traditional CNN and transformer models, our AER-SwinT integrates hierarchical feature extraction with dynamic attention mechanisms. Comparative experiments demonstrate that AER-SwinT outperforms both CNN-based methods and other transformer-based architectures, underscoring the importance of capturing global context for effective feature representation. Ablation analyses highlight the value of the multitask learning approach and channel attention residual aggregation. By incorporating clin-

- （基于通道平均的注意力增强残差）此外，我们的方法在多个阶段采用基于通道平均的注意力图。这些注意力图动态地突出CT图像中的兴趣区域，使得模型能够通过逐点乘法和残差连接，专注于关键空间位置。这一机制确保了模型在每个阶段都强调重要特征，显著提高了其捕捉不同细节的能力，提升了整体性能。
- 与传统的CNN和变换器模型相比，我们的AER-SwinT将层次化特征提取与动态注意力机制结合。比较实验表明，AER-SwinT优于基于CNN的方法和其他基于变换器的架构，强调了捕捉全局上下文对有效特征表示的重要性。

Discussion

第六段：算法层面动机

tion analyses highlight the value of the multitask learning approach and channel attention residual aggregation. By incorporating clinically relevant immune biomarkers, such as ImmunoScore and POSTN expression, as intermediate tasks, our model enhances predictive accuracy for chemotherapy and immunotherapy outcomes. This innovative approach overcomes the limitations of existing methods, offering a robust framework for comprehensive clinical analysis and marking a significant advancement in medical

- （多任务学习）消融分析突出了多任务学习方法和通道注意力残差聚合的价值。通过将临床相关的免疫生物标志物（如免疫评分和POSTN表达）作为中间任务，我们的模型提高了化疗和免疫治疗结果的预测准确性。

Discussion

第六段：算法层面动机

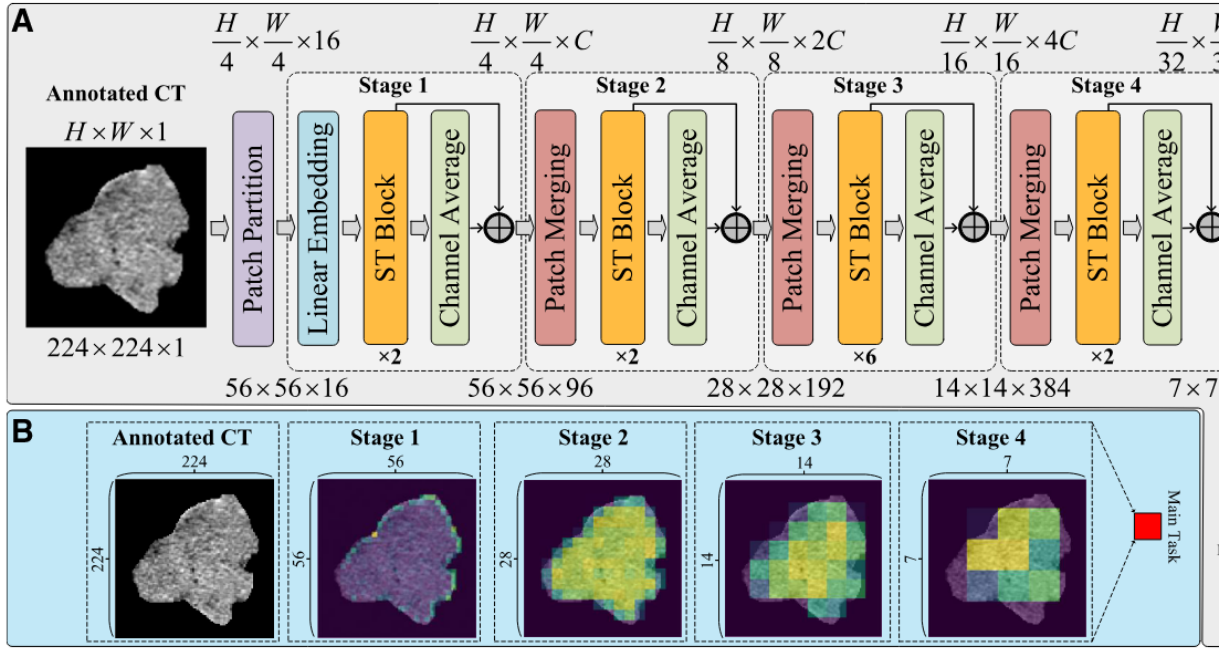
comes. This innovative approach overcomes the limitations of existing methods, offering a robust framework for comprehensive clinical analysis and marking a significant advancement in medical imaging and predictive analytics. Additionally, channel average-based attention maps effectively leverage global relationships within the data, significantly contributing to performance improvements. These findings suggest that integrating a channel attention mechanism helps isolate and amplify critical features, making it a valuable component in transformer-based medical image analysis.

- 这一创新方法克服了现有方法的局限，提供了一个强大的框架用于综合临床分析，标志着医学影像和预测分析的重大进展。此外，基于通道平均的注意力图有效地利用数据中的全局关系，显著推动了性能的提升。这些发现表明，整合通道注意力机制有助于隔离和放大关键特征，使其成为基于变换器的医学影像分析中的宝贵组成部分。

Discussion

第七段：可视化注意力图 描述

In addition, Figures 2B and S3 show the attention generated at each stage of our model. In the initial stage, our model primarily focuses on the edge information of the tumor within the CT images. As the model deepens, it progressively extracts coarser granularity information while simultaneously increasing attention on the internal regions of the tumor. These examples demonstrate that our method, through hierarchical feature extraction and guided attention mechanisms, enables the model to focus on different areas of the tumor at various granularities. Consequently, this approach allows for a more comprehensive analysis of the patient's image data.



- 此外，图展示了我们模型在每个阶段生成的注意力。在初始阶段，我们的模型主要关注CT图像中肿瘤的边缘信息。随着模型的深入，它逐渐提取更粗粒度的信息，同时增加对肿瘤内部区域的关注。这些示例表明，通过层次化特征提取和引导注意力机制，我们的方法使模型能够在不同的粒度下专注于肿瘤的不同区域。因此，这种方法使得患者图像数据的分析更加全面。

Discussion

第八段：再次强调引入多任务的优势

During the primary task of chemotherapy response prediction, we adopted training strategies that introduced TME and routine clinical features, which further improved the model's performance and outperformed a traditional single-task deep learning model or single-scale convolution neural network. Compared with traditional machine-learning-based radiomics models and deep-learning-based single-task imaging models, the proposed multitask AER-SwinT model optimized the training procedure and achieved cost-effective, accurate predictions. Future research should appropriately adopt deep learning techniques that offer cost and time advantages and ensure the development of robust models that meet clinical requirements.

个人理解：特征中加入常规临床特征和TME因素不需要算力成本，且多任务的框架 只需要用共享的特征做多个任务，用更低的成本达到了更高的准确率。

- 化疗反应预测的主要任务中，我们采用了引入TME和常规临床特征的训练策略，这进一步提高了模型的性能，超越了传统的单任务深度学习模型或单尺度卷积神经网络。与传统的基于机器学习的放射组学模型和基于深度学习的单任务影像模型相比，所提出的多任务AER-SwinT模型优化了训练过程，实现了具有**成本效益**和**准确性**的预测。未来的研究应适当采用具有成本和时间优势的深度学习技术，并确保开发出符合临床需求的稳健模型。

Discussion

第九段：CT图像异质性

Although some methods were adored to reduce the heterogeneity of CT images, the heterogeneity of these data should not be ignored. In the future, several strategies should be addressed to reduce the impact of CT scanner heterogeneity on model performance. First, we aim to standardize the models and parameters of CT machines across different centers to reduce variability in imaging data. Second, we plan to develop and incorporate advanced normalization algorithms that dynamically adjust for scanner-specific differences during preprocessing, ensuring consistent feature extraction across datasets. Third, we will leverage data augmentation techniques, such as adding noise simulated from different scanning protocols and domain adaptation methods, to enhance the model's adaptability to data from diverse scanners. These efforts will help minimize scanner-induced variability and improve the model's reliability in clinical applications.

尽管有一些方法被用来减少CT图像的异质性，但这些数据的异质性不应被忽视。未来，应采取几种策略来减少CT扫描仪异质性对模型性能的影响。首先，我们的目标是标准化不同中心的CT机器和参数，以减少影像数据的变异性。其次，我们计划开发并整合先进的标准化算法，在预处理过程中动态调整扫描仪特定的差异，确保在不同数据集之间提取一致的特征。第三，我们将利用数据增强技术，如添加来自不同扫描协议模拟的噪声和领域适应方法，以增强模型对来自不同扫描仪数据的适应性。这些努力将有助于最小化扫描仪引起的变异性，并提高模型在临床应用中的可靠性。

Discussion - Limitations of the study

Limitations of the study

This study has some limitations. The primary limitation of this study is its retrospective nature. The second limitation is that the data were obtained from patients of East Asian origin only, and the pathological subtype and distribution of clinical characteristics might be different in patients from other geographical regions, necessitating validation by other diverse populations and ethnic groups in a large cohort. The third point is that CT images were achieved from various scanners in different institutions, which may increase the heterogeneity of the data. Fourthly, this study only explored the application of the developed model in immunotherapy for advanced GC. Further investigation in the context of perioperative immunotherapy is need. Additionally, gender was included as a demographic variable in the dataset, but no gender-specific analyses were conducted. We recognize this as a limitation of the current study, and further investigations are warranted to determine whether gender may affect the study outcomes. Finally, the patients enrolled in the study were treated over decades and not within a randomized trial setting. Although the treatment was 5-fluorouracil based, the treatment drugs and cycles are not uniform, and the decision to treat patients with chemotherapy or not was made by the clinicians or patients, or both. As such, a rigorously designed randomized controlled trial to validate the generalizability and reproducibility is required.

首先，研究的主要局限性在于其回顾性特点。第二个局限性是数据仅来自东亚背景的患者，其他地区患者的病理亚型和临床特征分布可能不同，因此需要在其他多样化人群和不同族群的大型队列中进行验证。第三，CT图像来自不同机构的各种扫描仪，这可能增加数据的异质性。第四，本研究仅探讨了所开发模型在晚期胃癌免疫治疗中的应用，未来在围手术期免疫治疗的背景下仍需进一步研究。此外，性别作为一个人口学变量被包含在数据集中，但未进行性别特定的分析。我们认识到这是当前研究的一个局限性，未来需要进一步研究性别是否会影响研究结果。最后，研究中纳入的患者接受了数十年的治疗，而不是在随机试验设定下进行的。尽管治疗基于5-氟尿嘧啶，但治疗药物和周期并不统一，是否接受化疗的决策是由临床医生、患者或二者共同决定的。

好词好句

- 1) However, despite the survival benefits of adjuvant chemotherapy, the 5-year overall survival rate for advanced GC remains below 40%, highlighting a significant risk of unnecessary or delayed treatment for a substantial number of patients
  - highlighting a significant risk of：可以用来强调现有问题遇到的困境
- 2) These conflicting results suggest an urgent clinical need for predictive biomarkers to identify which patients will benefit from adjuvant chemotherapy.
  - suggest an urgent clinical need for xxx
  - 表明了 xxxx 的紧急需求

好词好句

- 3) Recently, the emergence of immunotherapies, including immune checkpoint inhibitors (ICIs), cancer vaccines, oncolytic viruses, and cell therapies, has revolutionized cancer treatment
  - has revolutionized：彻底改变了、变革型的变化
  - 比较有气势的表达

- 4) Therefore, accurately identifying patients who will benefit from immunotherapy to maximize therapeutic outcomes is a critical issue that needs to be addressed.
  - is a critical issue that needs to be addressed
  - 是一个需要解决的关键问题

好词好句

- 5) Therefore, developing a non-invasive and cost-effective model to assess chemotherapy response while integrating TME status to enhance model performance is imperative
  - Imperative： 紧急的，迫切的

- 6) Currently, there is a lack of effective biomarkers for predicting the benefit of 5-fluorouracil-based chemotherapy, making the identification of predictive biomarkers to personalize treatment for patients with GC crucial and long overdue.
  - Crucial： 关键的，至关重要的
  - Long overdue : 早该进行

好词好句

- 7) Our previous studies have shown that ImmunoScore and periostin (POSTN) expression are significantly associated with chemotherapy benefit.
  - 描述先前工作
  - Significantly：显著的
- 8) Recently, deep learning has emerged as a transformative methodology for automatically learning representative features from annotated tumor images for disease evaluation.
  - XX 方法的浮现 : emerged
  - 变革性的方法： transformative methodology

好词好句

- 9) The TME is cumulatively recognized as the key regulator of all types of anti-cancer therapy.
  - Be cumulatively recognized： 被公认为是xxx
  - 感觉也比较有气势，但使用这句话需要加参考文献

- 10) Given the limitations in tissue access, as well as the high cost, time-consuming nature, and technical complexity of histological approaches, our method offers a viable alternative to overcome these challenges.
  - 鉴于xxx局限性： Given the limitations
  - 我们的方法提供了可行的选项： our method offers a viable alternative

观点论据

- 1) Accurate evaluation of the TME can improve the assessment of immunotherapy effectiveness in GC. Additionally, the tumor immune status of patients is closely correlated with chemotherapy response, highlighting the value of incorporating TME evaluation into chemotherapy response assessment.
- **探究肿瘤微环境的价值（免疫疗法和化疗效果评估）**
  - 对TME的准确评估可改善对胃癌免疫疗法疗效的评估。此外，患者的肿瘤免疫状态与化疗反应密切相关，这凸显了将 TME 评估纳入化疗反应评估的价值。

观点论据

- 2) Currently, there is a lack of effective biomarkers for predicting the benefit of 5-fluorouracil-based chemotherapy, making the identification of predictive biomarkers to personalize treatment for patients with GC crucial and long overdue.
- **探索化疗有效生物标志物至关重要**
  - 目前，缺乏有效的生物标志物来预测以 5 氟尿嘧啶为基础的化疗的疗效，因此，鉴定预测性生物标志物以对 GC 患者进行个性化治疗至关重要，而且早该进行。

观点论据

- 4) However, the assessment of the TME relying on histopathological staining has limitations, such as invasiveness, sample heterogeneity, high cost, time consuming nature, and technical complexity.
- **传统TME评估具有侵入性、异质性、高成本、耗时性和复杂性**
  - 然而，依靠组织病理学染色对肿瘤组织进行评估有其局限性，如侵入性、样本异质性、高成本、耗时和技术复杂性。

观点论据

- 5) In contrast, multitask deep learning enables the simultaneous analysis of different tasks within a single model. By sharing feature representations and interactions among related tasks, multitask learning is more data efficient and has been shown to reduce overfitting and improve model generalization across various applications, including computer vision, disease diagnosis, and drug discovery.
- **多任务学习的好处**
  - 相比之下，多任务深度学习可在单个模型中同时分析不同的任务。 通过共享相关任务之间的特征表征和交互，多任务学习的数据效率更高，并已证明可在计算机视觉、疾病诊断和药物发现等各种应用中减少过度拟合，提高模型泛化能力。