JNM
The Journal of Nuclear Medicine

Meeting Report | PIDS: Data Sciences & Imaging Informatics

Longitudinal Multi-Instance Learning of Baseline and Post-Therapy 18F-FDG PET/CT for Prognostic Stratification in Esophageal Cancer

Liang Guo, Ayah Nawwar, Iain Lyburn, Julie Searle, Vijay Shah, Guenther Platsch, Daniel Darian, Bin Lou, Bruce Spottiswoode and Lauren Partin

Journal of Nuclear Medicine May 2026, 67 (supplement 1) 261352;

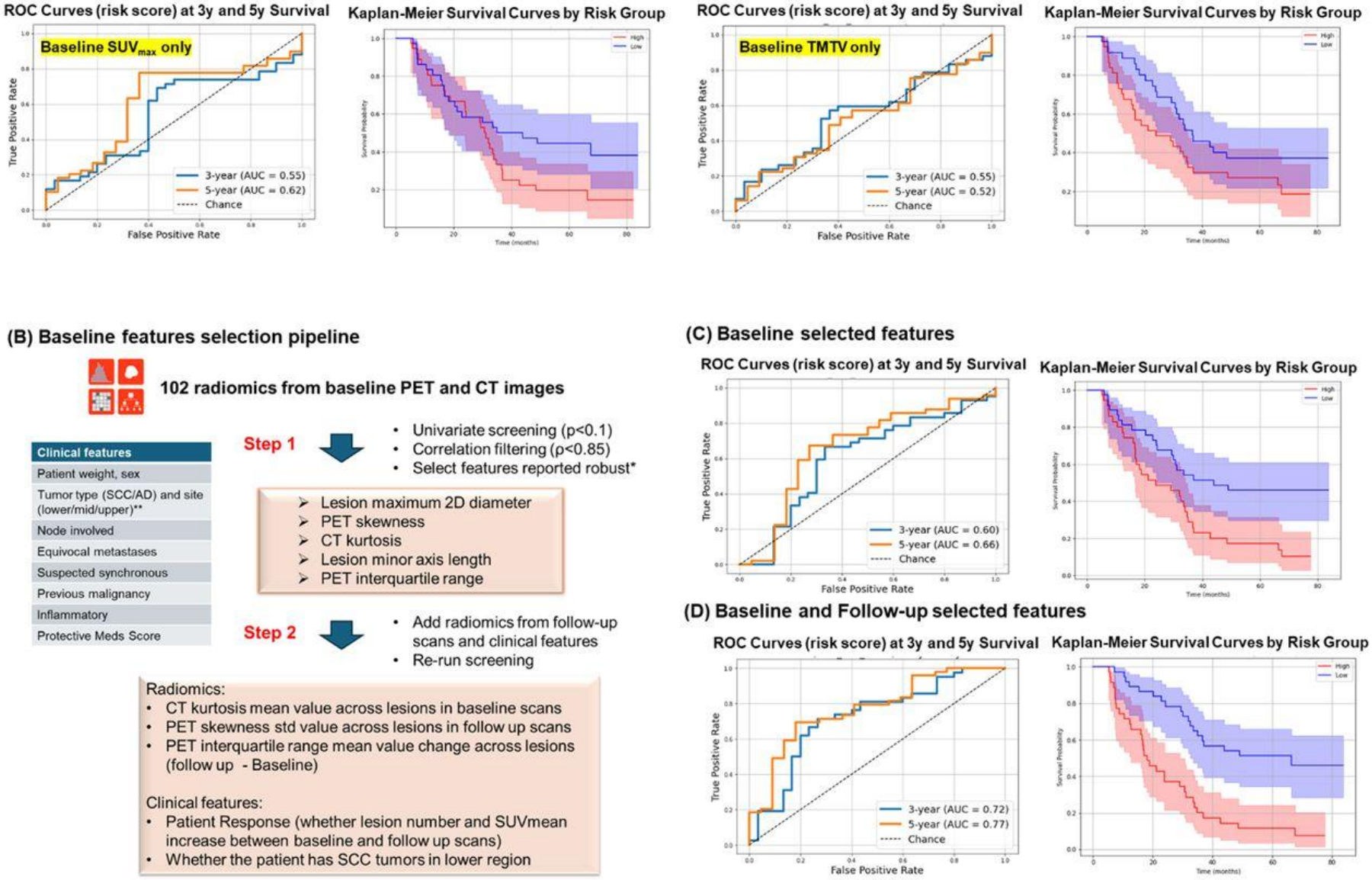
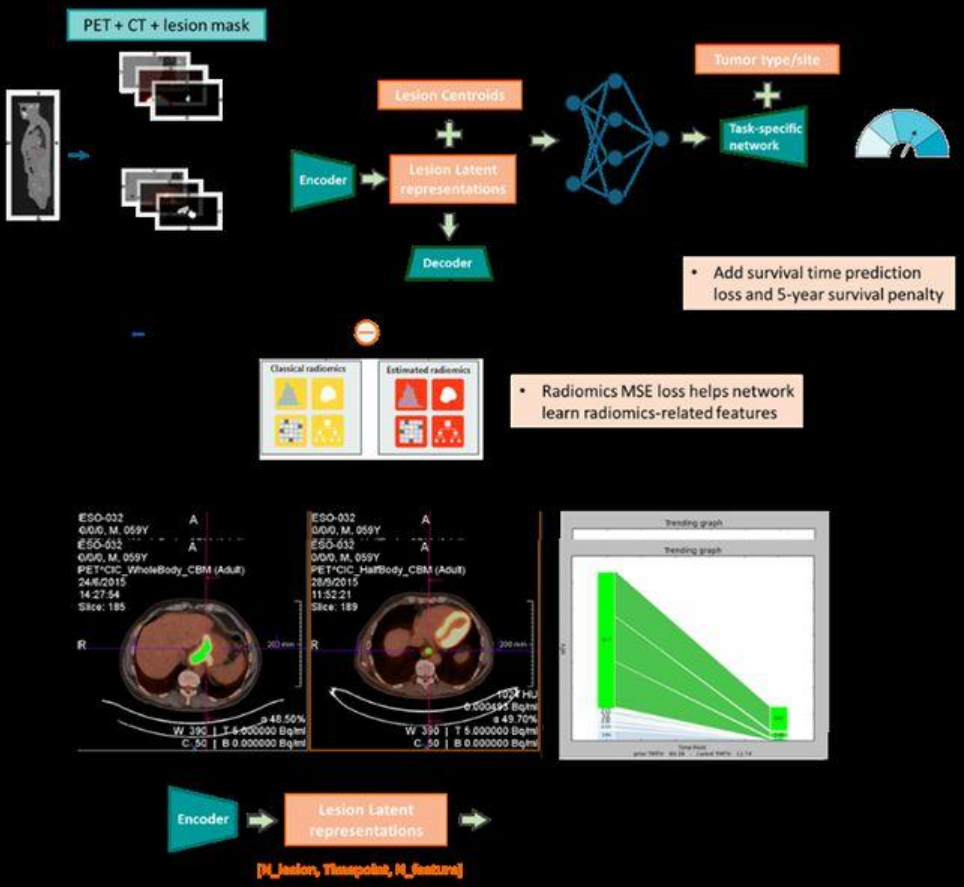
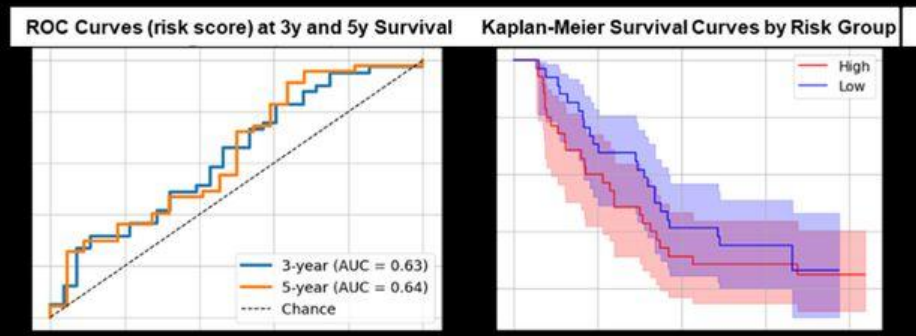
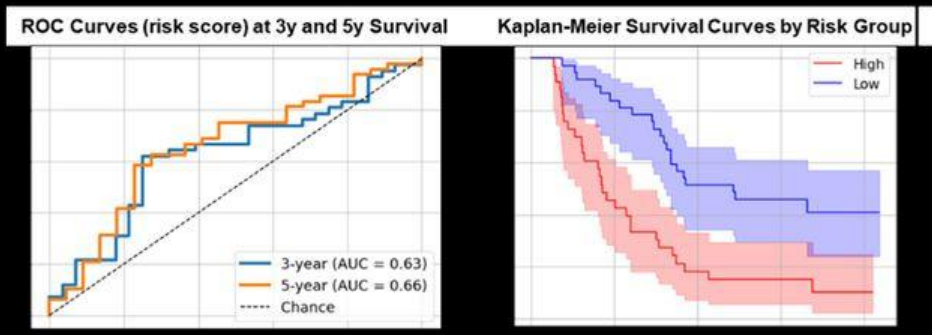


Fig 1. Radiomics workflow and survival prediction results. (A) Univariable model with baseline SUV_{max} only and baseline TMTV only. (B) Two-step feature screening pipeline: baseline radiomics selection (Step 1) followed by addition of follow-up radiomics and clinical features (Step 2). (C) Baseline selected features: ROC and Kaplan-Meier curves show moderate discrimination. (D) Adding follow-up radiomics and clinical features improves 3- and 5-year survival prediction and risk stratification.

*Thomas et al, 2023. "Reproducibility in Radiomics: A Comparison of Feature Extraction Methods and Two Independent Datasets
**SCC: squamous cell carcinoma, AD: adenocarcinoma


PET + CT + lesion mask
Lesion Centroids
Encoder
Lesion Latent representations
Decoder
Tumor type/site
Task-specific network
• Add survival time prediction loss and 5-year survival penalty
Classical radiomics
Estimated radiomics
• Radiomics MSE loss helps network learn radiomics-related features
Encoder
Lesion Latent representations
[N_lesion, Timepoint, N_feature]


ROC Curves (risk score) at 3y and 5y Survival
3-year (AUC = 0.63)
5-year (AUC = 0.64)
Chance
Kaplan-Meier Survival Curves by Risk Group
High
Low


ROC Curves (risk score) at 3y and 5y Survival
3-year (AUC = 0.63)
5-year (AUC = 0.66)
Chance
Kaplan-Meier Survival Curves by Risk Group
High
Low

Abstract
261352

Introduction: F-FDG PET/CT is routinely used for staging and response assessment in esophageal cancer (EC) following chemoradiotherapy. However, validated quantitative biomarkers that integrate longitudinal changes in metabolic and morphological patterns remain limited. Conventional metrics such as SUVmax and total metabolic tumor volume (TMTV) only capture partial aspects of tumor biology. This study aimed to develop a longitudinal multi-instance learning (MIL) framework that integrates radiomics and clinical features from baseline and post-therapy PET/CT to improve individualized risk stratification and survival prediction in EC patients.

Methods: A total of 74 EC patients (median age = 66±12 y, 56 male) who underwent baseline and 2-week post-therapy follow-up F-FDG PET/CT were retrospectively analyzed. All patients received standard chemoradiotherapy treatment between the two scans. Primary and metastatic lesions were segmented by two experienced board-certified PET-CT readers according to PERCIST criteria, using a 42% SUVmax threshold. For each lesion, 102 radiomic features in each modality (first-order, shape, and texture of PET and CT images) were extracted from both timepoints using Pyradiomics. Since each patient could have multiple lesions, patient-level feature representations were summarized using the mean and standard deviation of lesion-wise features. Clinical variables (shown in Fig 1B) were also included. Feature preprocessing involved z-score normalization, univariate Cox screening ($p < 0.1$), and removal of highly correlated features ($|\rho| > 0.85$). The feature list was further refined to retain only the subset of those previously reported as robust. Using these, three conventional models – Cox proportional hazards, Random Survival Forest (RSF), and Support Vector Machine (SVM) - were evaluated by time-dependent AUC and Kaplan-Meier curves using leave-one-out cross validation (LOOCV). In addition, as a clinically familiar baseline comparator, we evaluated an univariable model based solely on baseline lesion SUVmax and TMTV to compare prognostic information provided by multivariable features. For deep learning, each patient was represented as a set of lesion-level feature vectors from baseline and follow-up scans. An attention-based MIL network aggregated lesion embeddings into a patient-level risk vector using a combined survival time prediction mean-squared-error loss and 5-year survival penalty. Auxiliary radiomics prediction loss was included to reduce overfitting. Models were trained separately for baseline-only and combined baseline and follow-up inputs. For combined input, a Multi Time Point Analysis (MTPA) software prototype was used to link lesions between two scans. The input and network structure are shown in Fig 2A. LOOCV was used for evaluation, and Kaplan-Meier curves were generated using the median-split risk scores.

Results: Feature screening using baseline PET/CT alone identified a subset of prognostic radiomic descriptors, shown in Fig 1B. A traditional Cox model using these baseline-only features achieved a 5-yr survival AUC of 0.66, compared to 0.62 with SUVmax-only and 0.52 with TMTV-only. Incorporating follow-up PET/CT features and clinical variables further improved performance to 0.77 as shown in Fig 1D. Among conventional machine-learning approaches, the Cox model outperformed Random Survival Forest (RSF: 0.70) and Support Vector Machine (SVM: 0.72). For deep learning–MIL, the baseline-only model achieved a 5-yr survival AUC of 0.64, and the longitudinal model achieved 0.66, both lower than the Cox model. However, the MIL longitudinal-model demonstrated a clear separation of high- and low-risk groups (log-rank $p < 0.01$), indicating superior risk stratification when integrating both baseline and post-therapy lesion information.

Conclusions: This study demonstrates that integrating baseline and post-therapy 18F-FDG PET/CT using both traditional Cox modeling and a longitudinal multi-instance learning framework enhances prognostic assessment in esophageal cancer. Although the MIL models did not surpass the Cox model in 5-yr survival AUC, likely due to the limited dataset size, they provided strong separation of high- and low-risk cohorts, highlighting the value of capturing lesion-level heterogeneity in treatment response assessment. These results suggest that longitudinal lesion-informed modeling, particularly with larger training cohorts, has potential to support adaptive treatment strategies and personalized follow-up in esophageal cancer management.