Imperial College London

162P - A CT Radiomics Model to Predict Overall Survival following Curative-Intent Radiotherapy for Stage I-III Non-Small Cell Lung Cancer

Sumeet Hindocha[1,2,3,4], Thomas G. Charlton[5], Kristofer Linton-Reid[4], Benjamin Hunter[1,3,4], Charleen Chan[3], Merina Ahmed[1], Emily J. Greenlay[1], Catey Bunce[1], Jason Lunn[3], Simon Doran[3], Matthew Orton[1], Shahreen Ahmad[5], Fiona McDonald[1], Imogen Locke[1], Danielle Power[6], Matthew Blackledge[3], Richard W. Lee[1,3*], Eric O. Aboagye[4*]

1=The Royal Marsden NHS Foundation Trust, 2=AI for Healthcare Centre for Doctoral Training, Imperial College London, 3=Institute of Cancer Research, 4= Department of Surgery & Cancer, Imperial College London, 5=Guy's and St Thomas' NHS Foundation Trust, 6=Imperial College Healthcare NHS Trust.

Introduction:

Recurrence occurs in up to 36% of patients treated with curative-intent radiotherapy for NSCLC[1]. Identifying patients at higher risk of recurrence for more intensive surveillance may facilitate earlier introduction of the next line of treatment. High-quality evidence to inform optimal surveillance strategies is lacking[2].

We present a radiomics model to predict overall survival (OS) 2 years post-treatment, using routinely available radiotherapy planning CTs and the gross tumour volume (GTV), contoured for radiotherapy purposes as the region of interest for feature extraction. Such a model could be used to risk-stratify patients for tailored surveillance.

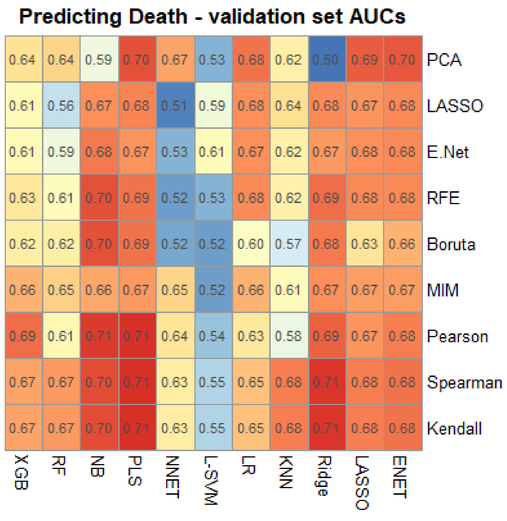


Figure 1. A Heatmap illustrating the performance of each machine learning algorithm (rows) with each feature reduction technique (columns), measured by validation set AUC. PCA: Principle Component Analysis, LASSO: Least Absolute Shrinkage and Selection Operator, E Net: Elastic-Net, RFE: Recursive Feature Elimination, MIM: Mutual Information, XGB: Extreme Gradient Boosting machine, RF: Random Forest, NB: Naïve-Bayes, PLS: Partial Least Squares, NNET: Neural Network, L-SVM: Linear Support Vector Machine, LR: Logistic regression, KNN: K-Nearest Neighbours, Ridge: Ridge regression.

Methods:

A retrospective multi-centre study of patients receiving stereotactic or conventional (chemo)radiotherapy for stage I-III NSCLC was undertaken. Cases with a GTV encompassing the primary tumour were included from 5 hospitals and divided into a training, validation, and external test set. Radiomic features were extracted from GTVs using our in-house software, TexLAB 2.0.

Survival status (alive/dead) at 2 years from first fraction of radiotherapy was binarized for classification purposes – cases scored "1" if there was death within 2 years, and "0" otherwise. Data from 4 hospitals were combined and cases randomly assigned to training or validation sets with an 80:20 ratio, stratified by the binarized outcome. Data from the 5th hospital was used as a geographically external test set.

Feature reduction (FR) may be required prior to machine learning (ML) with radiomic features to increase prediction accuracy. We explored a combination of 9 feature reduction techniques with 11 machine learning classifiers.

ROC curves were created for the validation set results of each FR – ML classifier combination and the AUC calculated. Ensemble prediction models were explored by averaging predictions of the 3 algorithms with the highest AUC in the validation set. Where the ensemble model was superior, it was selected as the final model. Otherwise, the single algorithm with the highest AUC was selected as the final model and deployed on the external test set.

The model was benchmarked against a model built on TNM-stage data. Youden Index was derived from the validation set ROC curve to distinguish high and low risk groups. Kaplan-Meier analysis was performed.

	Validation Set Results			External Test Set Result		
	AUC	95% CI	AUC p-value	AUC	95% CI	AUC p-value
Our model	0.712	0.592 – 0.832	0.013	0.685	0.585 – 0.784	0.621
TNM Model	0.573	0.442 – 0.704	-	0.663	0.579 – 0.77	-

Table 1: Validation and external test set AUC values with 95% confidence intervals for our radiomic model and a model built on TNM stage alone. AUC p-values compare our model to the TNM model.

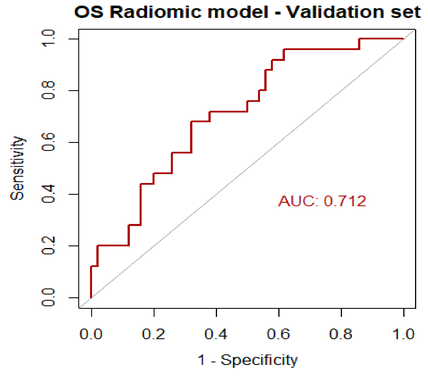
Results:

509 patients were included, training=302, validation=75, test=132. Median follow-up was 762 days. Train-validation and external test set mean age was 74 and 71 respectively. Death rates at 2 years were 33.7% vs 32.6%.

The final survival prediction model was a Partial Least Squares classifier with Spearman correlation (Figure 1).

Validation and test set AUCs and 95% confidence intervals are shown in Figure 2 and Table 1 together with a comparison against the TNM stage benchmarking model.

Our model had superior AUC to the TNM model in both validation and test sets. Kaplan Meier curves showed marked separation with significant log-rank tests (Figure 3).



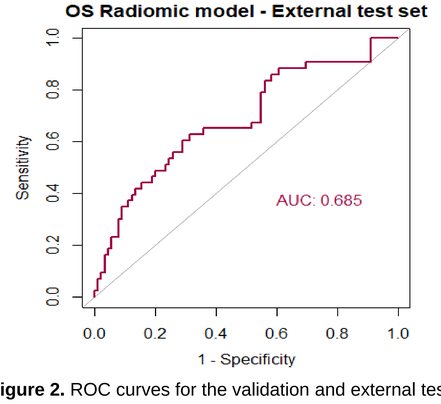


Figure 2. ROC curves for the validation and external test set.

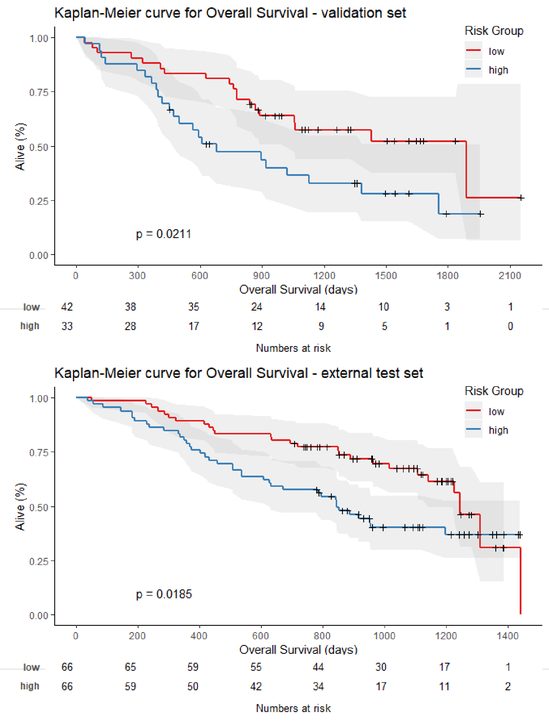


Figure 3. Kaplan Meier survival curves for low vs. high-risk groups in validation and external test sets. P-values correspond to log-rank tests.

Conclusion:

We present a validated and externally tested survival-stratification model that uses routinely available radiotherapy data. This model could be integrated into the radiotherapy workflow to inform personalised surveillance at the point of treatment for patients with NSCLC.

This work was funded by the NIHR Biomedical Research Centre at The Royal Marsden NHS Foundation Trust and The Institute of Cancer Research and the UKRI Centre for Doctoral Training in AI for Healthcare, grant number: P/S023283/1. Dr Sumeet Hindocha has no conflicts of interest to declare.

Contact: **sh806@ic.ac.uk** Twitter: **@shindocha87**

The ROYAL MARSDEN NHS Foundation Trust | ICR The Institute of Cancer Research | FHealth