Fusing Data from CT Deep Learning, CT Radiomics and Peripheral Blood Immune profiles to Diagnose Lung Cancer in Symptomatic Patients

Rami Mustapha[1]*, Balaji Ganeshan[2]*, Sam Ellis[3]*, Luigi Dolcetti[1]*, Mukunthan Tharmakulasingam[1], Karen DeSouza[1], Xiaolan Jiang[1], Courtney Savage[1], Sheena Lim[4], Emily Chan[3], Andrew Thornton[2], Luke Hoy[2], Raymond Endozo[2], Rob Shortman[2], Darren Walls[2], Shih-hsin Chen[2], Ashley M Groves[2], Julia A Schnabel[1,3], Thida Win[4]^, Paul R Barber[1^$], Tony Ng[1^$]

Author affiliations:

[1] Comprehensive Cancer Centre, King's College London, London UK
[2] University College London, London UK
[3] School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK
[4] East and North Hertfordshire NHS Trust, UK

* = Joint first author
^ = Joint last author
$ = Joint corresponding author

Abstract

Background: Lung cancer is the leading cause of cancer-related deaths. Diagnosis at late stages is common due to the largely non-specific nature of presenting symptoms contributing to high mortality. There is a lack of specific, minimally invasive low-cost tests to screen patients ahead of the diagnostic biopsy.

Patients and Methods: 344 symptomatic patients from the lung clinic of Lister hospital suspected of lung cancer were recruited. Predictive covariates were successfully generated on 170 patients from Computed Tomography (CT) scans using CT Texture Analysis (CTTA) and Deep Learning Autoencoders (DLA) as well as from peripheral blood data for immunity using high depth flow-cytometry and for exosome protein components. Predictive signatures were formed by combining covariates using Bayesian regression on a randomly chosen 128-patient training set and validated on a 42-patient held-out set. Final signatures were generated by fusing the data sources at different levels.

Results: Immune and DLA were the best single modality signatures with test set AUCs of 0.76 (95% CI: 0.61 – 0.91) and 0.75 (95% CI: 0.60 - 0.90) respectively. The final combined signature had a ROC AUC of 0.86 (95% CI: 0.73 - 0.99) on the withheld test set. The overall sensitivity and specificity were 0.722 and 0.901 respectively.

Conclusions: Combining immune monitoring with CT scan data is an effective approach to improving sensitivity and specificity of Lung cancer screening even in symptomatic patients.

Keywords: Lung cancer; Early Diagnosis; CT Scan; Blood test; Immune

Highlights:

- Combining immune monitoring in peripheral blood with CT scan data improves lung cancer screening sensitivity and specificity.
- Elevated levels of KIR3DL1+ CD8 T cells may be indicative of cancer.
- A cancer biomarker that combines a deep learning autoencoder with peripheral immune profiling achieved a 0.86 ORC AUC.

Introduction

Lung cancer and importance of early detection

Lung cancer is by far the leading cause of cancer-related deaths worldwide with an estimated 1.8 million death in 2020 [1]. In the U.S.A., lung cancer accounts for about two and a half times the number of cancer deaths than the second leading cause, colorectal cancer. This is despite the fact that lung cancer only contributes to about 12% of newly diagnosed cancer cases, which is far less than prostate cancer and breast cancer, the leading cancer diagnoses for men and women, respectively [2]. This high mortality of lung cancer has been linked to diagnosis at late stages. 48% of lung cancer cases are diagnosed with a distant metastasis and an expected 5-year survival rate of 8% compared to the ≥60% rate when diagnosis happens at a localized stage [3]. In England, 5-year survival for a stage I diagnosis is 63%, which drops down to 4.3% for a stage IV diagnosis [4]. Delayed diagnosis can be attributed to the largely asymptomatic nature of early-stage lung cancer [5]. Other data have shown that patients delaying their first general practitioner (GP) visit after manifestation of early symptoms is another obstacle to a timely diagnosis [6]. Beyond that, a study from 2013 showed that a third of lung cancer patients have 3 or more visits with their GP with lung cancer associated symptoms before special referral, which forms a sharp contrast to the 3% observed in breast cancer patients [7]. Similarly, the UK's National Cancer Diagnosis Audit reported primary care delays of 60 and 90 days experienced by 17.9% and 10.8% of suspected lung cancer patients, respectively [8]. This highlights the need for lung cancer screening, which to a certain extent has been implemented in the form of low dose computed tomography (LDCT) for high-risk patients. Despite this practice leading to about a 20% reduction in mortality across multiple clinical trials, there is still room for improvement [9, 10]. LDCT screening excludes younger patients and those at lower risk as it has been shown to be not reliable as a screening tool for the broader population, owing to its high chance of false positives [11, 12]. Lung cancer screening also heavily relied on smoking history, which excludes the early detection of lung cancer in non-smokers which represent around 25% of cases [13]. There is a clear need for a minimally invasive robust screening test for lung cancer with similar sensitivity and specificity to a mammography for breast cancer, or the faecal occult blood test currently in use for colorectal cancer. Furthermore, this needs to be extended beyond screening, as numerous healthcare systems, including the NHS in the UK, have put in place a rapid diagnostic pathway for already symptomatic patients suspected of lung cancer [14].

Role of radiologist and importance of alternative image analysis

The diagnostic pathway for lung cancer is heavily reliant on radiologist assessment at multiple stages. According to the UK NHS National Optimal Lung Cancer Pathway (NOLCP) [15] and National Institute for Health and Care Excellence (NICE) guidelines, the typical pathway may include a chest radiograph (CXR), a CT scan, a PET-CT, a

CT guided biopsy, all of which require a radiologists intervention, bronchoscopy (EBUS-TBNA, EUS-FNA or navigational) and pleural procedures. This reliance is further exacerbated by the large amount of CT imaging data from lung cancer screening programmes. This workload combined with the shortage of radiologists in the UK and other European countries is manifesting in delays in examinations as recently reported by the Royal College of Radiologists [16, 17]. Implementing automated image analysis approaches like radiomics-based texture and deep learning analyses is a viable approach, especially given the availability of large screening datasets for training purposes. The low-dose radiation used in LDCT, to minimize patient exposure, does negatively impact the visibility of small or low-contrast lesions due to increased noise and reduced spatial resolution and artificial intelligence can alleviate this in many areas of CT screening [18,19]. In this study we exploit the latent vector of a deep autoencoder as a low dimensional representation of the lesion image characteristics and evaluate its predictive performance.

Liquid biopsy

Using a liquid biopsy for early detection and diagnosis of lung cancer has been heavily assayed to varying degrees of success [20]. Of note, Galleri® which is a circulating tumoral DNA (ctDNA) methylation-based multicancer early detection (MCED) test is seeing adoption across clinics. Early results seem very promising in the symptomatic setting, though these are not comparable across all cancer types. Of note, sensitivity for stage I lung cancer was 8.7% and 0% in the symptomatic and asymptomatic settings, respectively [21, 22]. Combining LDCT with minimally invasive blood test is one approach that can improve the shortcoming of either approach. This has proven useful in the BioMilk study that showed that double positivity LDCT and microRNA signature doubles the sensitivity of LDCT alone [23]. Similarly, other studies combined circulating proteins with LDCT achieving improved specificity [24]. However, there is still room for improvement using other blood biomarkers which include circulating tumour cells, microRNA, exosomes, tumour educated platelets, metabolites, tumour associated proteins, and autoantibodies.

Immunology in lung cancer

Little research exists on the feasibility of immune profiling of circulating immune cells as a biomarker for early lung cancer. Peripheral blood mononuclear cells (PBMCs) are a heterogenous population of immune cells usually classified into subsets. The numbers, relative frequencies and functional status these subsets provide information on the state of the immune response and there have been recent advents in technology beyond classical multiparametric flow cytometry. Using patient PBMCs, single cell RNA sequencing was used to identify circulating anti-tumor CD8 cells, and Cytometry by time of flight (CyTOF) was used to correlate the frequency of HLADR expressing monocytes with response to ICB therapy [25, 26]. Numerous research groups, including ours, have shown that changes at the tissue level of a solid cancer immune response are reflected as systemic changes in the immune response that can be

detected at the peripheral level [25-27]. Lung cancer promotion after initial mutagenesis has been shown to be inflammation driven, with a recent link established to air pollution [28], further highlighting the potential of an immune based biomarker for early detection. Cancer specific T cell receptor repertoire has been shown to be effective at identifying early-stage lung cancer in an LDCT screened cohort. This proves that cancer specific immune cells can be reflected in the periphery [29].

Methods

Study design and data collection

The study entitled “Improving the Early Detection of Lung Cancer by Combining Exosomal Analysis of Hypoxia with Standard of Care Imaging (LungExoDETECT)” (https://clinicaltrials.gov/ct2/show/NCT04629079) is a prospective cohort study of patients referred to secondary care for the investigation of clinical symptoms or signs suspicious of lung cancer. The study analysis will determine whether the assay can detect clinical lung cancer at the time of imaging, and interval cancers during subsequent follow up. The study aimed to establish preliminary ROC AUC and sensitivity/specificity data for the "combined CT/blood risk stratification marker" and provide initial data on the potential association of the "combined CT/blood risk score" with the subsequent cancer progression and treatment response.

The study included patients who have been referred to the Lung Cancer Clinic and Multi-Disciplinary Team (MDT) at The Lister, Hertford County, UK and New QEII Hospitals, Welwyn Garden City, UK for investigation of suspected lung cancer. Participants were recruited between October 2020 and November 2021 (344 patients). As part of the standard of care pathway they received a CT scan, at that time peripheral blood samples were collected. 174 cases were excluded, 162 due to lack of visible lesion or lesion <6mm and 12 cases with non-lung cancer. The patient pathway and study design is illustrated in Figure 1 and patient characteristics are in Table 1.

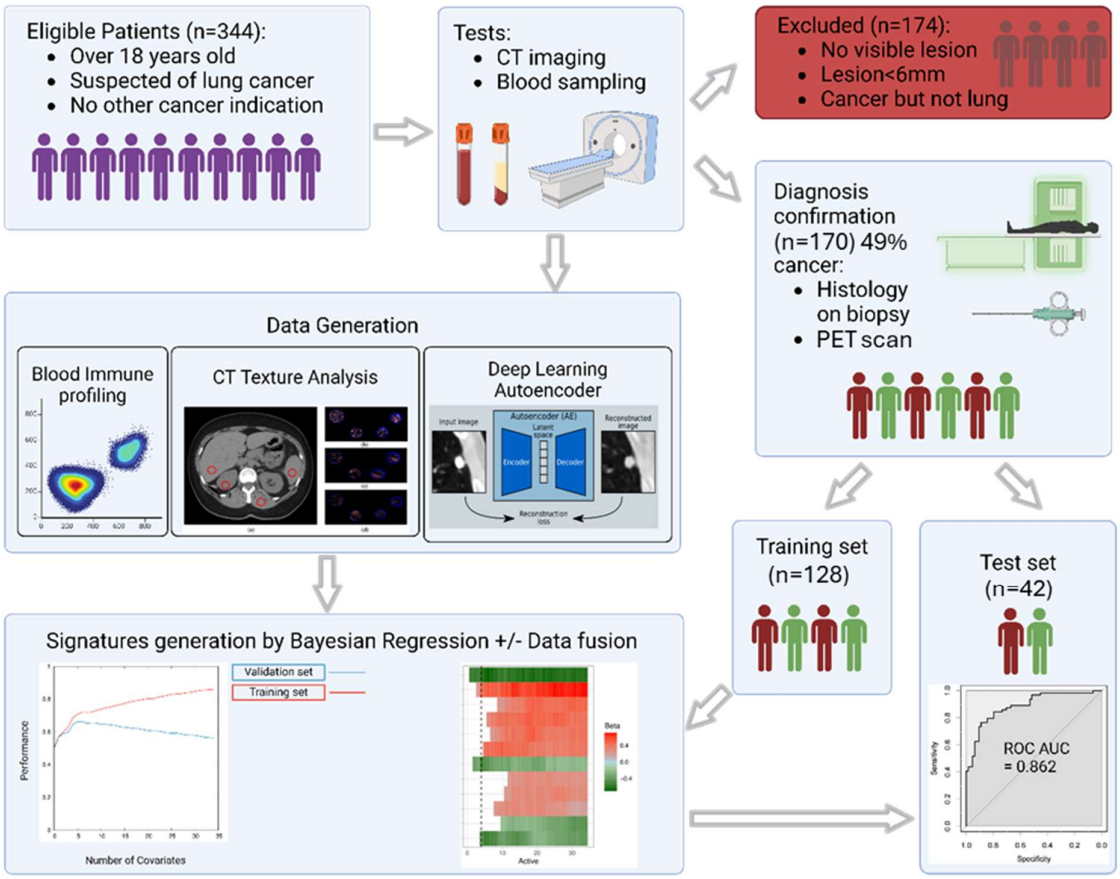

Figure 1. Showing the patient pathway and study design. 344 patients referred to the clinic were recruited to the study. As part the standard of care pathway they received a CT scan, at which time we collected the peripheral blood samples. 174 cases were excluded from the analysis set due to lack of visible lesion, the lesion was smaller than 6mm or the diagnosis was a non-lung cancer diagnosis. The analysis was performed on the remaining patients.

Patient Characteristic	Non-cancer patients *n* = 91	Lung cancer patients *n* = 79	p-value*
Age			
Mean	71.9	74.9	0.152
Range	27.0 to 94.0	33.0 to 97.0	
Sex			
Male	43	38	0.646
Female	48	41	
Smoking status			
Ex-Smoker	33	31	0.07
Non-Smoker	29	12	
Smoker	17	21	
Unknown	12	15	
Cohort			
Training	64	64	0.1
Test	27	15	
Staging			
No Cancer	91		
Stage I		23	
Stage II		9	
Stage III		18	
Stage IV		29	

Table 1. Patient characteristics and distribution. The table shows the difference between cancer and non-cancer as well as staging for cancer patients. *P-values were calculated to compare data from individuals with and without lung cancer for the following variables: mean ages using Student's unpaired two-tailed t-test whereas for sex distribution, smoking status and cohort using a $\chi 2$ test.

Sample and Data Processing

The details of CT scan acquisition, CT Texture Analysis (CTTA), the Deep Learning Autoencoder (DLA), peripheral blood sampling and processing, flow cytometry and model (aka signature) generation are given in the Supplementary Material. The total number of features generated by each modality are: 42 from CTTA, 32 from DLA, 45 from flow cytometry and 68 from exosome dot-blots.

Model Performance Analysis

Of all the study samples, one quarter was held out of the model training and served as an internal test set. The performance of both univariate biomarkers and model predicted outcomes from the training and test sets were judged by the area under the receiver operator characteristic curve (ROC AUC) with the cancer diagnosis result from the clinic as the gold standard using the R 'pROC' package and the 'ci.auc' function. Sensitivity and specificity were calculated at the Youden point of the ROC. A

Risk Score can be generated by combining covariate raw values in a linear model according to their given weights and intercept.

Results

Patient Characteristics

The characteristics of the patients suspected of lung cancer and included in the final analysis set are shown in Table 1. Of those suspected of having lung cancer 46.5% (79 patients) were diagnosed with lung cancer over the follow up period. There were no significant differences between the cancer and not cancer patients with regards to sex, age, smoking status nor assignment to training or test set. The division of patients between training (3/4) and test set (1/4) was retrospective and at random, with the condition of having an equal balance of cancer and non-cancer patients in the training set for signature generation (Supplementary material). Amongst the lung cancer patients, the division across stage is shown in Table 1. Of note, 23 of the lung cancer patients (29.1%) were stage I, signifying that this is a good cohort for the potential development of a screening signature. 12 patients of the total recruited were diagnosed with a cancer that is not lung cancer (characteristics in Supplementary Table S4).

Generation of a Predictive Model from CTTA and DLA

The imaging data were analysed using CTTA and DLA, as explained in the methodology. Separately, using the data from each analysis, a risk signature was generated using Bayesian multi-variate regression (BMR) which also performed covariate selection (supplementary material). The signatures are presented in Figure 2A. The DLA signature performed better than the texture analysis signature on both the training and test cohorts. On the training set DLA had an AUC of 0.78 (95% CI: 0.70 - 0.86), whereas CTTA had an AUC of 0.75 (95% CI: 0.67 - 0.84). On the test set DLA had an AUC of 0.75 (95% CI: 0.6 - 0.9), whereas CTTA had a much lower AUC of 0.64 (95% CI: 0.47 - 0.82) (Figure 2B). The CTTA and DLA risk signatures could significantly segregate cancer from non-cancer patients in the training cohort (Figure 2B, $p<0.001$), whereas only the DLA risk signature reached significance in the test cohort (Figure 2C, $p=0.004$). Figure 2D presents the result of combining CTTA and DLA analyses.

The CTTA features selected by BMR were the mean intensity at spatial scale filter (SSF) = 0 (Mean_0) – which relates to the CT attenuation / density of the tumour, the mean intensity of positive pixels at scale SSF=2 (mpp_2) – which relates to the average brightness at fine texture scale (corresponding to features of 2mm in size) within the tumour, and the histogram skewness at SSF=4 – which relates to bright or dark objects at medium texture scale (corresponding to features of 4mm in size) within the tumour. The CTTA and DLA combined model uses a combination of

features from the single analysis models, except the standard deviation of pixel values at scale SSF=3 (sd_3).

The DLA latent space features selected by BMR (numbered 11, 15, 18, 20 and 26 from 32 total latent variables) can be interpreted visually by varying these values and observing the effect on the image reconstruction of the decoder (Supplementary Figure 3). The significant parameters relate to lesion size and the morphology of the lesion and surrounding tissue.

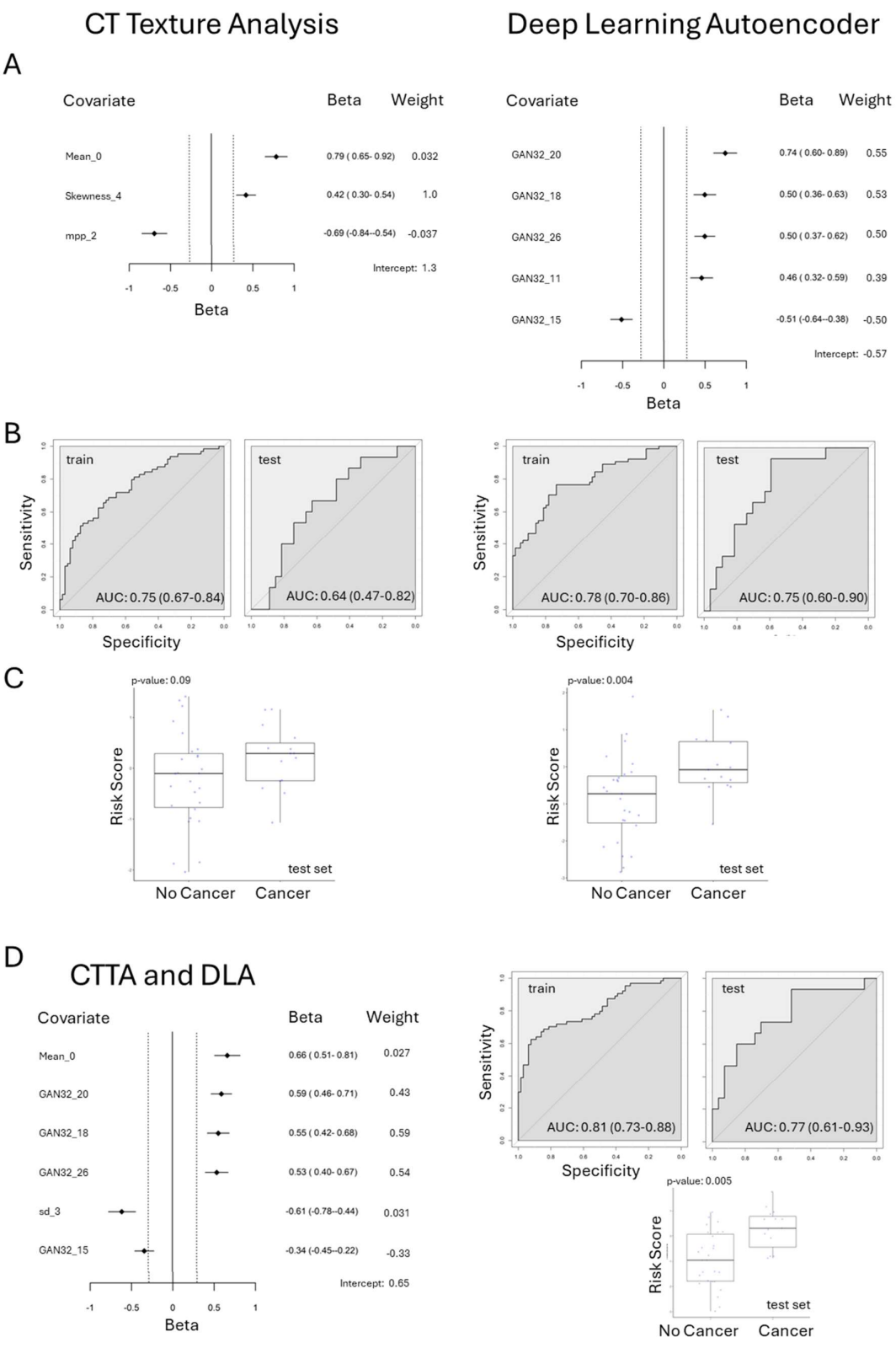

Figure 2. Risk generation by BMR from CT images comparing CT texture analysis to deep learning autoencoder. (A) Forest plot of the covariates within each of the risk score for each CTTA (left) and DLA (right) with dotted line indicating the performance range of randomised data, covariate model Beta value (with 95% CI) and weight in signature. (B) ROC AUC curves of the risk signature for the training set and test set, with AUC (and 95% CI). (C) Box plots with overlaid scatter points, showing the distribution of each risk score CTTA (left) and DLA (right) across non-cancer and cancer patients. (D) Results, as above, for model generation from features of CTTA and DLA combined.

Generation of a Predictive Model from Peripheral Blood

The peripheral blood samples were analysed for their immune and their exosome protein content (dot-blot) as indicated in the methods. Separately, using the data from each analysis modality, a risk signature was generated (Figure 3A). The flow cytometry signature performed far better on the training and test cohort than the exosome protein content signature. On the test set, the immune signature had an AUC of 0.76 (95% CI: 0.61 – 0.91), whereas the exosome derived signature had an AUC of 0.58 (95% CI: 0.41 - 0.76) (Figure 3B). Of note, the flow cytometry signature was driven by two significant features: a high proportion of type 2 dendritic cells indicated a non-cancer pathology, whereas a high proportion KIR3DL1 expressing CD8 T lymphocytes indicated lung cancer (Figure 3A). The flow cytometry risk score was significantly different between cancer and non-cancer, whereas the exosome dot-blot score was not and hence that signature was not used for further combined risk signature generation (Figure 3C).

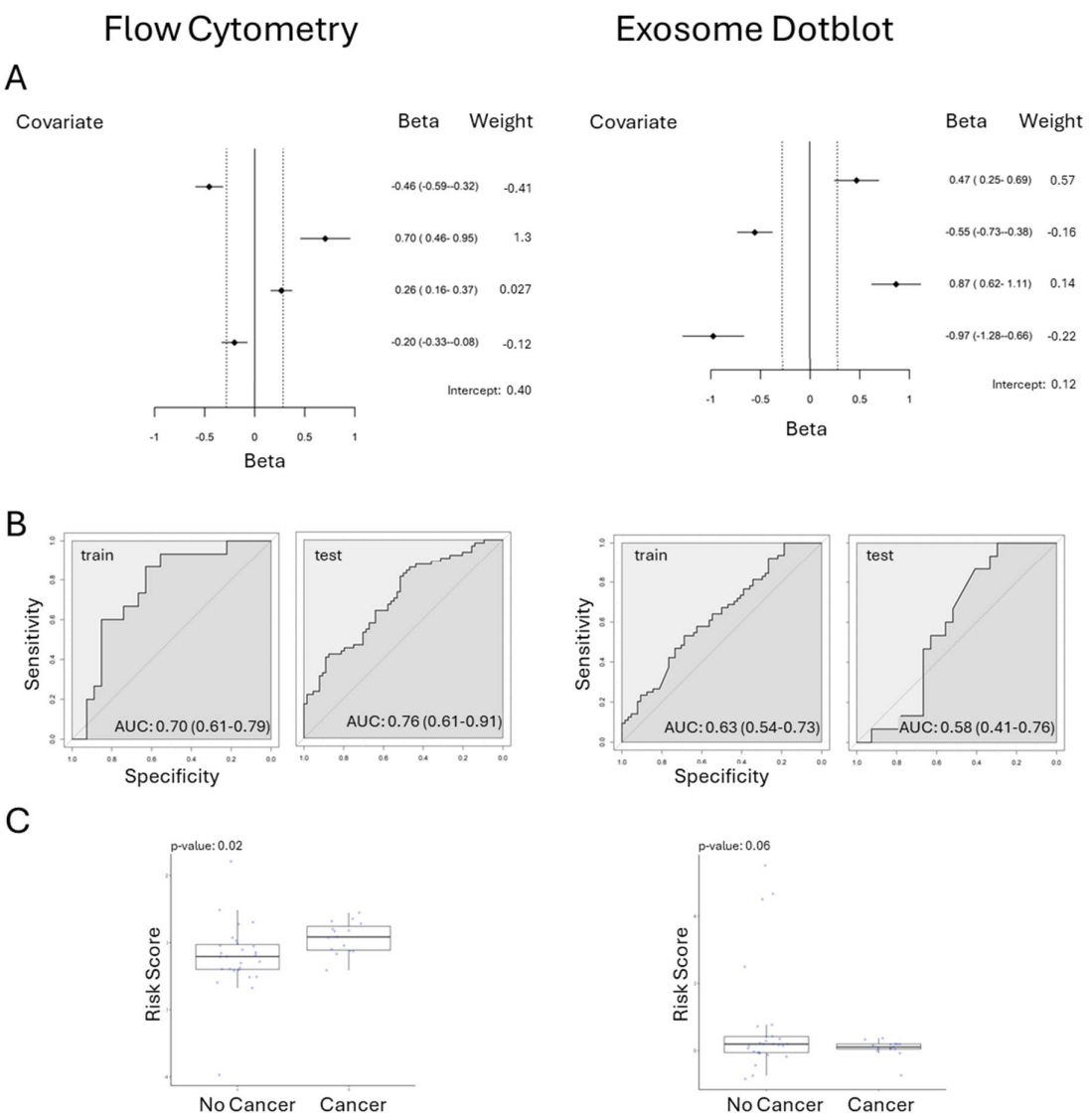

Figure 3. Risk generation by BMR from peripheral blood samples comparing flow cytometry to plasma exosome dotblot analysis. (A) Forest plot of the covariates within each of the risk score for each flow cytometry (left) and exosome dotblot (right) with dotted line indicating the performance range of randomised data, covariate model Beta value (with 95% CI) and weight in signature. (B) ROC AUC curves of the risk signature for the training set and test set, with AUC (and 95% CI). (C) Box plots with overlaid scatter points, showing the distribution of each risk score flow cytometry (left) and exosome dotblot (right) across non-cancer and cancer patients.

Combining CT Analyses with Peripheral Blood Data

We proceeded to combine the data from the blood sample with the CTTA and DLA individually, at the covariate level (early fusion). Both combined signatures from BMR featured covariates from the immune and imaging data with the upregulation of KIR3DL1 CD8 T cells being the strongest predictor of a lung cancer (Figure 4A). The risk signature generated from the combination of flow cytometry and DLA performed

far better on the training set with an AUC of 0.85 (95% CI: 0.78 - 0.91) (Figure 4B). However, these combinations did not generate a signature that outperformed the immune signature alone, as shown by the AUC on the testing set where the Immune-DLA signature had an AUC of 0.77 (95% CI: 0.61 - 0.92). The risk signature generated from the combination of peripheral immunity and CTTA on the training set had an AUC of 0.78 (95% CI: 0.70 - 0.86), whereas for the testing set the AUC was 0.67 (95% CI: 0.51 - 0.84). The immune-CTTA signature performed far worse than the immune alone signature.

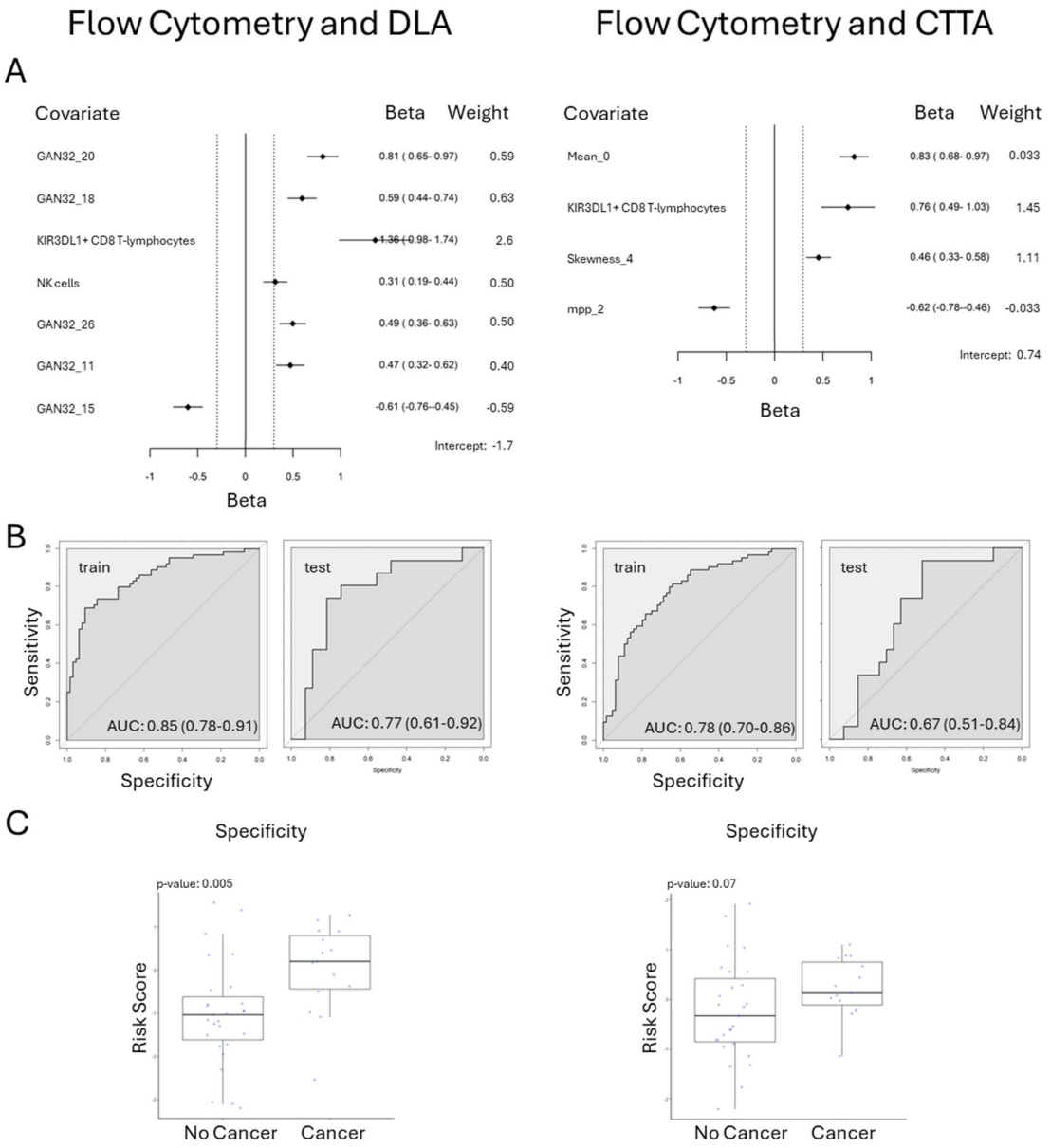

Figure 4. Risk generation by BMR from CT image analysis combined with flow cytometry, comparing deep learning and texture analysis. (A) Forest plot of the covariates within each of the risk score for each flow cytometry with DLA (left) and flow cytometry with CTTA (right) with dotted line indicating the performance range of randomised data, covariate model Beta value (with 95% CI) and

weight in signature. (B) ROC AUC curves of the risk signature for the training set and test set, with AUC (and 95% CI). (C) Box plots with overlaid scatter points, showing the distribution of each risk score of flow cytometry with DLA (left) and flow cytometry with CTTA across non-cancer and cancer patients.

Generation of a Combined Model by Data Fusion

Given that combination of the peripheral immune signature with either CT analysis modality (CTTA or DLA) did not impact the predictive power of the signature, we proceeded to combine all methods of CT and blood analyses using data fusion at different levels. Combination at the covariate level (early fusion) did not show a major improvement of the risk signature with AUC 0.76 (95% CI: 0.61 - 0.91) on the test set (Figure 5A, “All Combined early fusion”). The signature generation process included co-variates from all three modalities, but no CTTA covariates were selected in the final signature (Figure 5B).

We also generated predictive signatures by fusing the features included in individual signatures (feature level or intermediate fusion) and by fusing the individual signatures themselves (signature level or late fusion). With feature level fusion (intermediate) the covariate selection again only maintained covariates from the immune and DLA analyses, although CTTA was included in the generation process, with the signature still being driven by the upregulation of KIR3DL1+ CD8 T lymphocytes as a strong predictor of lung cancer (Figure 5C). The ROC AUC for feature level combination was 0.79 (95% CI: 0.65 - 0.93) on the test set (Figure 5A, “All Combined Int. Fusion”). Finally, the best signature was obtained when fusing the individual CTTA, DLA and immune signatures (late fusion), previously generated from the three separate analyses. The signature was almost equally driven by the immune and DLA signatures (Figure 5D), and to a lesser degree by the CTTA signature. This final signature had the highest ROC AUC of 0.86 (95% CI: 0.73 - 0.99) on the hold-out test set (Figure 5A, “All Combined late fusion”). All of the fusion models could significantly split the training set between cancer and non-cancer patients (Figure 5 B, C, D), with late fusion being more robust against overfitting to the training set.

An equivalent approach to data combination and signature generation was performed using an Elastic Net algorithm (glmnet R package) [30, 31] which produced better signatures on the training set, but struggled with overfitting which impacted the ROC AUCs on the test set. Of note, the final combined signature from Elastic Net only had a ROC AUC of 0.67 (95% CI: 0.5 - 0.84) on the test set. Because of this, we are focusing on the results from BMR. A summary of all the Elastic net signatures is in Supplementary Figure 4.

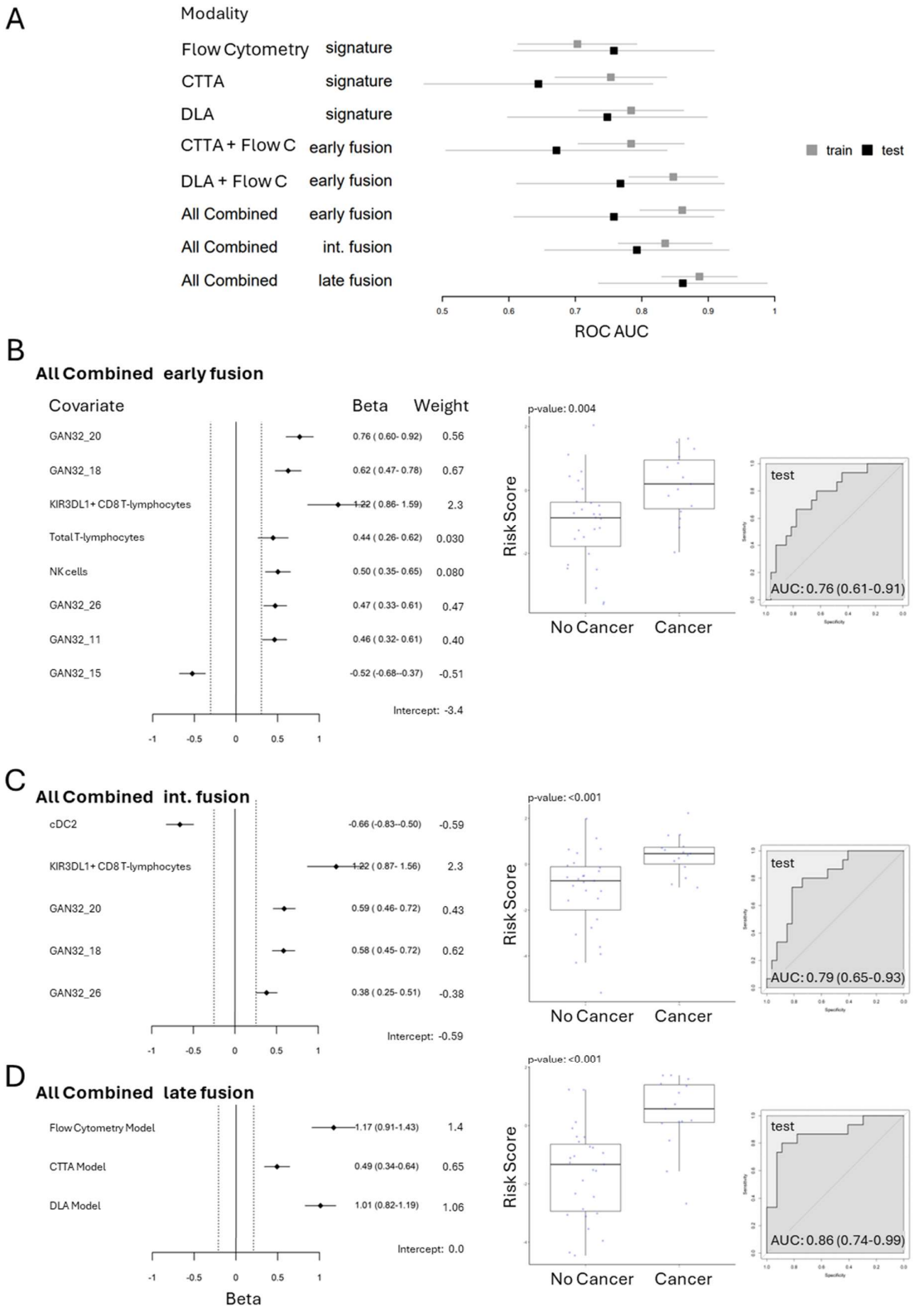

Figure 5 Risk signature generation from fusion models. (A) Summary of ROC AUC values on the training set (grey) and test set (black) for all the generated signatures using BMR, from individual signatures from immune, CTTA and DLA analyses, as well as early fusion (covariate level fusion)

models from immune and CTTA, and immune and DLA, and all three combined, and all three combined using intermediate and late fusion methods. (B) Covariate level (early fusion) of all three analyses showing a forest plot of the covariates within the risk signature using covariate level fusion model with dotted line indicating the range, test set cancer prediction box plot and ROC curve. (C) Feature level (intermediate fusion) of all three analyses showing a forest plot of the covariates within the risk signature using the feature level fusion model, box plot and ROC curve. (D) Signature level (late fusion) of all three analyses showing a forest plot of the covariates within the signature level fusion model, box plot and ROC curve.

Combined Signature Specificity and Sensitivity to Detect Lung Cancer at Different Stages and Across Other Cancers

Disease staging was intentionally excluded from the data sets for both signature generation and testing. As current methodologies struggle at detecting early-stage lung cancer [21, 22] where current available therapies are most effective, we assessed the ability of the final signature to detect cancer across all stages for the training and test sets (Figure 6A). As expected, test set sensitivity (53%) was lower than that of the training set (76%). In the test set, Stage I sensitivity was far lower at 25% compared to all the other stages. Interestingly, the specificity was high at 92% for the training set despite the fact that these are heavily symptomatic patients (Figure 6B) .

Data were analysed for the 12 patients who received a non-lung cancer diagnosis. As the primary disease was not a lung cancer, the efficacy of the lung imaging data was unclear, and hence we used the immune signature alone. The sensitivity of this signature to detect cancer in these cohorts was 82%, with 9/11 cancers detected (Figure 6C). The two missed cancers were both breast cancer.

A

Lung cancer	Total Cases	True Positive	False negative	Total Sensitivity	Training Set Sensitivity	Test Set Sensitivity
All Cases	79	57	22	0.722	0.76	0.533
Stage I	23	14	9	0.609	0.684	0.25
Stage II	9	5	4	0.556	0.5	1
Stage III	18	15	3	0.833	0.867	0.667
Stage IV	29	23	6	0.793	0.864	0.571

B

	Non-Cancer Patients	True Negative	False Positive	Total Specificity	Training Set Specificity	Test Set Specificity
Combined Late-fusion Signature	91	82	9	0.901	0.891	0.926

C

Immune Signature: Other Cancers	True Positive	False negative	Sensitivity
Colorectal	2	0	1
Lymphoma	2	0	1
Breast cancer	0	2	0
Thymoma	1	0	1
Chronic Lymphocytic Leukemia	2	0	1
Kidney cancer with Lung metastasis	1	0	1
Prostate	1	0	1
Melanoma	1	0	1

Figure 6 Tables of the performance of the final signature (the combined late-fusion model) when thresholding the risk score according to Youden's criterion: sensitivity (A) and specificity (B). Table C shows the performance of the immune signature with those patients who were diagnosed with a non-lung cancer; the lung CT scan therefore not being relevant as a predictor of non-lung cancer.

Discussion

Immune Subpopulations

We have previously shown that early on-treatment changes in peripheral levels of memory CD8 T-cells can predict response in head and neck cancer patients [27], and that pre-treatment levels of a subpopulation antigen presenting capable myeloid cells can predict treatment outcome. We therefore postulated that that the systemic immune response may also be indicative of diagnosis of cancer in this study and highlight the fact that the signature obtained from a simple blood sample had a stronger predictive power than a chest CT scan. The strongest predictor in the signatures was the upregulation of KIR+ CD8 T lymphocytes. KIR3DL1 is part of the

inhibitory KIR family. Its expression on CD8 T cells has been linked to effector CD8 T cells with reduced proliferative capabilities and poor IFNg production following TCR engagement [32, 33]. The inhibitory effect of KIR binding to HLA has been shown to not directly impact the T lymphocyte's ability to degranulate but rather to compromise activation induced transcription which is required for clonal expansion and cytokine production [34]. KIR expression has been linked to demethylation following chronic TCR stimulation which is consistent with cancer immunosurveillance [35, 36]. In our data set, the KIR+ CD8 population exhibited a low level of expression of CD107a (Lamp1) indicating that those cells did not engage in degranulation which is even more indicative of their exhausted functional state (Supp. Fig 5).

Of particular interest is that this signature could distinguish between a cancer and a non-cancer pathology. As most of the other pathologies were linked to an infection or an allergic reaction, the upregulation of cDC2 in those patients is only logical. cDC2 are a subtype of DCs that play a key role for antigen presentation for a type 2 helper response (TH2) [37] which is key in controlling humoral immunity. cDC2 have also been correlated with their ability to activate a TH17 response, which has been shown to be upregulated in chronic allergy [38, 39]. The cDC2 cells expressed high levels of HLADR and CD38, indicating that they were indeed functional and pro-inflammatory. This subpopulation of myeloid cells did not express PDL2 indicating that it was not functionally suppressed (Supp. Fig. 6).

Deep Learning Radiomics and Combined Model Generation

There has been significant recent activity in the development of techniques to extract novel information from medical images and other patient sample data and exploit it for clinical use[40]. Links between imaging and the immune system have also been made[41-45]. Grossman et al.**Error! Bookmark not defined.** evaluated radiomic texture and shape features in NSCLC and found correlations with immune response, inflammation and survival, and a combined signature that combined clinical, genetic and radiomics achieved a test set concordance index of 0.73 against OS.

Deep Learning networks have shown great promise in several fields, including cancer diagnosis, for extracting significant features from complex data, but their often black-box nature is one barrier to clinical use[46]. Autoencoders are able to learn semantically meaningful representations of images in an unsupervised manner and offer some advantages: training on an external data set and transparency through the exploration of the latent space parameters. Examples have been published in using epigenetic data to determine subtypes of lung cancer[47], and image-based lesion detection and segmentation[48].

The method for combining radiomics with blood data, needs consideration (i.e. data fusion[49]) as do the algorithms used to generate the predictive models. A model formed from many covariates can easily have exceptional performance on the training set, but this often does not translate to test and validation data sets (the problem of overfitting). Many studies use parameter regularisation such as LASSO or Elastic Net algorithms

for signature generation and covariate selection. In particular, a study investigated the feasibility of preoperative 18F-fluorodeoxyglucose (FDG) PET/CT radiomics with machine learning (LASSO) to predict microsatellite instability (MSI) status in colorectal cancer (CRC) patients with a test set performance (ROC AUC) of 0.867 from a published set of radiomics features[50]. Methods more advanced than regularisation offer a greater potential for generalisation to test and validation sets. For instance, neural networks and XGBoost models have achieved high ROC AUC of 0.75-0.8 in multimodal data sets[51-53]. Bayesian methods offer robust models that are generalisable and more interpretable and can have automatic feature selection that aid interpretability[54, 55].

A ROC AUC of 0.8 is within the realms of what may be clinically useful depending on the context. Choosing the optimal point on the ROC curve, to define operational sensitivity and specificity, may not be an obvious one for the clinical situation. Indeed, the multi-cancer Grail tests[21, 56] have a similar AUC. The Galleri test, currently under clinical trial in the UK, has a specificity of around 99% but a sensitivity as low as 17% (depending on cancer and stage)[57].

This present study has data collected in a prospective and standardised manner and used Bayesian machine learning techniques that emphasise reproducibility. Furthermore, ours is a “white box” method that generates an optimal set of parameters and weights that are available for interpretation. We explore a combination of novel deep learning radiomics and pre-determined standardized features in the form of the well-published filtration-histogram based texture analysis approach (using the commercially available TexRAD software) for feature selection from CT images.

Acknowledgements

We are grateful to all participating staff at The Lister, Hertford County, UK and New QEII Hospitals, Welwyn Garden City, UK who were responsible for patient recruitment and sample collection.

Funding

This work was funded by the CRUK grant [C1519/A27375]. SE, EC and JAS were further supported by core funding from the Wellcome Trust/EPSRC Centre for Medical Engineering [WT203148/Z/16/Z] and by the National Institute for Health and Care Research (NIHR) Clinical Research Facility at Guy's and St Thomas' NHS Foundation Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

This work was undertaken at UCLH/UCL, which received a proportion of the funding from the UK's Department of Health's NIHR Biomedical Research Centre's funding scheme.

Conflicts of Interest

BG is the co-founder/co-inventor of TexRAD texture analysis software used in this study for CT Texture analysis and a shareholder (not an employee) of Feedback Plc., a UK-based company which owns the TexRAD texture analysis software. PB and TN are both currently employees of GlaxoSmithKline (GSK) but this study has not received any funding support from GSK.

All remaining authors have declared no conflicts of interest.

Ethics

All procedures were performed in compliance with relevant laws and institutional guidelines and have been approved by the appropriate institutional committee(s), and informed consent obtained from the subjects (REC reference: 19/EE/0357 20th Feb 2020).

Code Availability

Model generation was performed using the MsaWrapper R package (https://github.com/paulbarber/msaWrapper) and Saddle Point Signature (commercial software distributed by Saddle Point Science Ltd).

Data Availability

Anonymised data used for this analysis is available at reasonable request from: kcl.figshare.com/account/articles/28070390

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 Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. Ca-a Cancer Journal for Clinicians 2023; 73 (1): 17-48.

3 Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. Nat Rev Clin Oncol 2023; 20 (9): 624-639.

4 Cancer Research UK, Lung cancer statistics, https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer#heading-Zero accessed May. 2024.

5 Malalasekera A, Nahm S, Blinman PL et al. How long is too long? A scoping review of health system delays in lung cancer. European Respiratory Review 2018; 27 (149): 180045.

6 Walter FM, Rubin G, Bankhead C et al. Symptoms and other factors associated with time to diagnosis and stage of lung cancer: a prospective cohort study. Br J Cancer 2015; 112 Suppl 1 (Suppl 1): S6-13.

7 Lyratzopoulos G, Abel GA, Mcphail S et al. Measures of promptness of cancer diagnosis in primary care: secondary analysis of national audit data on patients with 18 common and rarer cancers. British Journal of Cancer 2013; 108 (3): 686-690.

8 Swann R, McPhail S, Witt J et al. Diagnosing cancer in primary care: results from the National Cancer Diagnosis Audit. Br J Gen Pract 2018; 68 (666): e63-e72.

9 National Lung Screening Trial Research T, Aberle DR, Berg CD et al. The National Lung Screening Trial: overview and study design. Radiology 2011; 258 (1): 243-253.

10 de Koning HJ, van der Aalst CM, de Jong PA et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. N Engl J Med 2020; 382 (6): 503-513.

11 Wood DE, Kazerooni EA, Aberle D et al. NCCN Guidelines(R) Insights: Lung Cancer Screening, Version 1.2022. J Natl Compr Canc Netw 2022; 20 (7): 754-764.

12 Dingillo G, Bassiri A, Badrinathan A et al. Lung Cancer in Young Patients is Associated With More Advanced Disease but Better Overall Survival. J Surg Res 2023; 292: 307-316.

13 Woodard GA, Jablons DM. Shades of Gray: Do Never Smokers Benefit From Lung Cancer Screening Programs? J Thorac Oncol 2024; 19 (8): 1135-1137.

14 Dolly SO, Jones G, Allchorne P et al. The effectiveness of the Guy's Rapid Diagnostic Clinic (RDC) in detecting cancer and serious conditions in vague symptom patients. Br J Cancer 2021; 124 (6): 1079-1087.
15 National Optimal Lung Cancer Pathway (NOLCP). https://rmpartners.nhs.uk/wp-content/uploads/2024/09/national-optimal-lung-cancer-pathway_v4_01jan2024.pdf: NHS England; 2024.
16 The Royal College of Radiologists, Clinical oncology census reports 2022. https://www.rcr.ac.uk/news-policy/policy-reports-initiatives/clinical-oncology-census-reports/ accessed April. 2024.
17 Lung Cancer Europe: 2020 Position Paper. Disparities and Challenges in Access to Lung Cancer Diagnosis and Treatment Across Europe. Accessed: April.
18 Cellina M, Cacioppa LM, Ce M et al. Artificial Intelligence in Lung Cancer Screening: The Future Is Now. Cancers (Basel) 2023; 15 (17).
19 Kang E, Min J, Ye JC. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. Med Phys 2017; 44 (10): e360-e375.
20 Casagrande GMS, Silva MO, Reis RM, Leal LF. Liquid Biopsy for Lung Cancer: Up-to-Date and Perspectives for Screening Programs. Int J Mol Sci 2023; 24 (3).
21 Nicholson BD, Oke J, Virdee PS et al. Multi-cancer early detection test in symptomatic patients referred for cancer investigation in England and Wales (SYMPLIFY): a large-scale, observational cohort study. Lancet Oncol 2023; 24 (7): 733-743.
22 Schrag D, Beer TM, McDonnell CH, 3rd et al. Blood-based tests for multicancer early detection (PATHFINDER): a prospective cohort study. Lancet 2023; 402 (10409): 1251-1260.
23 Pastorino U, Boeri M, Sestini S et al. Baseline computed tomography screening and blood microRNA predict lung cancer risk and define adequate intervals in the BioMILD trial. Ann Oncol 2022; 33 (4): 395-405.
24 Ostrin EJ, Bantis LE, Wilson DO et al. Contribution of a Blood-Based Protein Biomarker Panel to the Classification of Indeterminate Pulmonary Nodules. J Thorac Oncol 2021; 16 (2): 228-236.
25 Krieg C, Nowicka M, Guglietta S et al. High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy. Nature Medicine 2018; 24 (2): 144-+.
26 Pauken KE, Shahid O, Lagattuta KA et al. Single-cell analyses identify circulating anti-tumor CD8 T cells and markers for their enrichment. J Exp Med 2021; 218 (4).
27 Barber PR, Mustapha R, Flores-Borja F et al. Predicting progression-free survival after systemic therapy in advanced head and neck cancer: Bayesian regression and model development. Elife 2022; 11.
28 Hill W, Lim EL, Weeden CE et al. Lung adenocarcinoma promotion by air pollutants. Nature 2023; 616 (7955): 159-167.
29 Li M, Zhang C, Deng S et al. Lung cancer-associated T cell repertoire as potential biomarker for early detection of stage I lung cancer. Lung Cancer 2021; 162: 16-22.
30 Friedman J, Tibshirani R, Hastie T. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software 2010; 33 (1).

31 Simon N, Friedman J, Tibshirani R, Hastie T. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. Journal of Statistical Software 2011; 39 (5).
32 Anfossi N, Doisne JM, Peyrat MA et al. Coordinated expression of Ig-like inhibitory MHC class I receptors and acquisition of cytotoxic function in human CD8+ T cells. J Immunol 2004; 173 (12): 7223-7229.
33 Arlettaz L, Degermann S, De Rham C et al. Expression of inhibitory KIR is confined to CD8+ effector T cells and limits their proliferative capacity. Eur J Immunol 2004; 34 (12): 3413-3422.
34 Henel G, Singh K, Cui D et al. Uncoupling of T-cell effector functions by inhibitory killer immunoglobulin-like receptors. Blood 2006; 107 (11): 4449-4457.
35 McLane LM, Abdel-Hakeem MS, Wherry EJ. CD8 T Cell Exhaustion During Chronic Viral Infection and Cancer. Annu Rev Immunol 2019; 37: 457-495.
36 Li G, Yu M, Weyand CM, Goronzy JJ. Epigenetic regulation of killer immunoglobulin-like receptor expression in T cells. Blood 2009; 114 (16): 3422-3430.
37 Kumar S, Jeong Y, Ashraf MU, Bae YS. Dendritic Cell-Mediated Th2 Immunity and Immune Disorders. Int J Mol Sci 2019; 20 (9).
38 Izumi G, Nakano H, Nakano K et al. CD11b(+) lung dendritic cells at different stages of maturation induce Th17 or Th2 differentiation. Nat Commun 2021; 12 (1): 5029.
39 Zhao J, Lloyd CM, Noble A. Th17 responses in chronic allergic airway inflammation abrogate regulatory T-cell-mediated tolerance and contribute to airway remodeling. Mucosal Immunol 2013; 6 (2): 335-346.
40 Chetan MR, Gleeson FV. Radiomics in predicting treatment response in non-small-cell lung cancer: current status, challenges and future perspectives. Eur Radiol 2021; 31 (2): 1049-1058.
41 Grossmann P, Stringfield O, El-Hachem N et al. Defining the biological basis of radiomic phenotypes in lung cancer. Elife 2017; 6.
42 Yang Q, Gong H, Liu J et al. A 13-gene signature to predict the prognosis and immunotherapy responses of lung squamous cell carcinoma. Sci Rep 2022; 12 (1): 13646.
43 Wiesweg M, Mairinger F, Reis H et al. Machine learning reveals a PD-L1-independent prediction of response to immunotherapy of non-small cell lung cancer by gene expression context. Eur J Cancer 2020; 140: 76-85.
44 Nunez NG, Berner F, Friebel E et al. Immune signatures predict development of autoimmune toxicity in patients with cancer treated with immune checkpoint inhibitors. Med 2023; 4 (2): 113-129 e117.
45 Duruisseaux M, Martinez-Cardus A, Calleja-Cervantes ME et al. Epigenetic prediction of response to anti-PD-1 treatment in non-small-cell lung cancer: a multicentre, retrospective analysis. Lancet Respir Med 2018; 6 (10): 771-781.
46 Singh A, Sengupta S, Lakshminarayanan V. Explainable Deep Learning Models in Medical Image Analysis. J Imaging 2020; 6 (6).
47 Wang Z, Wang Y. Extracting a biologically latent space of lung cancer epigenetics with variational autoencoders. BMC Bioinformatics 2019; 20 (Suppl 18): 568.

48 Astaraki M, Toma-Dasu I, Smedby Ö, Wang C. Normal Appearance Autoencoder for Lung Cancer Detection and Segmentation. Cham: Springer International Publishing; 2019:249-256.

49 Lipkova J, Chen RJ, Chen B et al. Artificial intelligence for multimodal data integration in oncology. Cancer Cell 2022; 40 (10): 1095-1110.

50 Kim S, Lee JH, Park EJ et al. Prediction of Microsatellite Instability in Colorectal Cancer Using a Machine Learning Model Based on PET/CT Radiomics. Yonsei Med J 2023; 64 (5): 320-326.

51 Jiang J, Jin Z, Zhang Y et al. Robust Prediction of Immune Checkpoint Inhibition Therapy for Non-Small Cell Lung Cancer. Front Immunol 2021; 12: 646874.

52 Tian P, He B, Mu W et al. Assessing PD-L1 expression in non-small cell lung cancer and predicting responses to immune checkpoint inhibitors using deep learning on computed tomography images. Theranostics 2021; 11 (5): 2098-2107.

53 Ahn BC, So JW, Synn CB et al. Clinical decision support algorithm based on machine learning to assess the clinical response to anti-programmed death-1 therapy in patients with non-small-cell lung cancer. Eur J Cancer 2021; 153: 179-189.

54 Shalabi A, Inoue M, Watkins J et al. Bayesian clinical classification from high-dimensional data: Signatures versus variability. Stat Methods Med Res 2018; 27 (2): 336-351.

55 Hozumi H, Shimizu H. Bayesian network enables interpretable and state-of-the-art prediction of immunotherapy responses in cancer patients. PNAS Nexus 2023; 2 (5): pgad133.

56 Klein EA, Beer TM, Seiden M. The Promise of Multicancer Early Detection. Comment on Pons-Belda et al. Can Circulating Tumor DNA Support a Successful Screening Test for Early Cancer Detection? The Grail Paradigm. Diagnostics 2021, 11, 2171. Diagnostics 2022; 12 (5): 1243.

57 Pons-Belda OD, Fernandez-Uriarte A, Diamandis EP. Multi Cancer Early Detection by Using Circulating Tumor DNA—The Galleri Test. Reply to Klein et al. The Promise of Multicancer Early Detection. Comment on “Pons-Belda et al. Can Circulating Tumor DNA Support a Successful Screening Test for Early Cancer Detection? The Grail Paradigm. Diagnostics 2021, 11, 2171”. Diagnostics 2022; 12 (5): 1244.