# Mining whole-lung information by artificial intelligence for predicting EGFR genotype and targeted therapy response in lung cancer: a multicohort study

*Shuo Wang*, He Yu*, Yuncui Gan*, Zhangjie Wu, Encheng Li, Xiaohu Li, Jingxue Cao, Yongbei Zhu, Liusu Wang, Hui Deng, Mei Xie, Yuanyong Wang, Xidong Ma, Dan Liu, Bojiang Chen, Panwen Tian, Zhixin Qiu, Jinghong Xian, Jing Ren, Kun Wang, Wei Wei, Fei Xie†, Zhenhui Li†, Qi Wang†, Xinying Xue†, Zaiyi Liu†, Jingyun Shi†, Weimin Li†, Jie Tian†*

## Summary

**Background** Epidermal growth factor receptor (*EGFR*) genotype is crucial for treatment decision making in lung cancer, but it can be affected by tumour heterogeneity and invasive biopsy during gene sequencing. Importantly, not all patients with an *EGFR* mutation have good prognosis with EGFR-tyrosine kinase inhibitors (TKIs), indicating the necessity of stratifying for *EGFR*-mutant genotype. In this study, we proposed a fully automated artificial intelligence system (FAIS) that mines whole-lung information from CT images to predict *EGFR* genotype and prognosis with EGFR-TKI treatment.

**Methods** We included 18 232 patients with lung cancer with CT imaging and *EGFR* gene sequencing from nine cohorts in China and the USA, including a prospective cohort in an Asian population (n=891) and The Cancer Imaging Archive cohort in a White population. These cohorts were divided into thick CT group and thin CT group. The FAIS was built for predicting *EGFR* genotype and progression-free survival of patients receiving EGFR-TKIs, and it was evaluated by area under the curve (AUC) and Kaplan-Meier analysis. We further built two tumour-based deep learning models as comparison with the FAIS, and we explored the value of combining FAIS and clinical factors (the FAIS-C model). Additionally, we included 891 patients with 56-panel next-generation sequencing and 87 patients with RNA sequencing data to explore the biological mechanisms of FAIS.

**Findings** FAIS achieved AUCs ranging from 0·748 to 0·813 in the six retrospective and prospective testing cohorts, outperforming the commonly used tumour-based deep learning model. Genotype predicted by the FAIS-C model was significantly associated with prognosis to EGFR-TKIs treatment (log-rank p<0·05), an important complement to gene sequencing. Moreover, we found 29 prognostic deep learning features in FAIS that were able to identify patients with an *EGFR* mutation at high risk of TKI resistance. These features showed strong associations with multiple genotypes (p<0·05, *t* test or Wilcoxon test) and gene pathways linked to drug resistance and cancer progression mechanisms.

**Interpretation** FAIS provides a non-invasive method to detect *EGFR* genotype and identify patients with an *EGFR* mutation at high risk of TKI resistance. The superior performance of FAIS over tumour-based deep learning methods suggests that genotype and prognostic information could be obtained from the whole lung instead of only tumour tissues.

**Funding** National Natural Science Foundation of China.



## Introduction

Targeted therapy to epidermal growth factor receptor (EGFR) has revolutionised the treatment of lung cancer.[1,2] Presently, the administration of EGFR-targeted therapy is determined by *EGFR* genotype, assessed through the gene sequencing of biopsied tumour tissues, which faces the difficulties of locating suitable tumour tissues because of the high genetic heterogeneity of lung cancer,[3,4] the changeable mutation status over time,[5,6] and poor DNA quality.[7,8] Most importantly, many patients with an *EGFR* mutation develop disease progression within 9–15 months after receiving EGFR-tyrosine kinase inhibitors (TKIs).[9] This suggests the need for stratifying *EGFR*-mutant genotypes according to their prognosis to targeted therapy, which cannot be assessed through gene sequencing. Consequently, a non-invasive method that can detect *EGFR* genotype and stratify patients with an *EGFR* mutation according to their therapeutic response to EGFR-TKIs is needed to assist decision making in targeted therapy, which is a great supplement to gene sequencing.

Artificial intelligence combined with CT imaging has shown promising results for non-invasively analysing lung cancer,[3,10,11] providing an alternative to analysing the whole tumour and its microenvironment with no additional cost.

Previous studies adopted tumour-based methods to extract tumour features in CT imaging for predicting

*Lancet Digit Health* 2022; 4: e309–19

Published Online March 24, 2022 https://doi.org/10.1016/S2589-7500(22)00024-3

*Contributed equally

†Joint last authors

**Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Engineering Medicine, Beihang University, Beijing, China** (Prof S Wang PhD, Z Wu BE, Y Zhu PhD, L Wang ME, Prof J Tian PhD); **Key Laboratory of Big Data-Based Precision Medicine, Beihang University, Ministry of Industry and Information Technology, People's Republic of China, Beijing, China** (Prof S Wang, Z Wu, Y Zhu, L Wang, Prof J Tian); **Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, China** (Prof J Tian); **CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, Beijing, China** (Prof S Wang, Z Wu, Y Zhu, L Wang, Prof K Wang PhD, Prof W Wei PhD, Prof J Tian); **Department of Respiratory and Critical Care Medicine, West China Hospital, Sichuan University, Chengdu, China** (H Yu MD, Y Gan MD, D Liu MD, B Chen MD, P Tian MD, Z Qiu MD, W Li MD); **Department of Respiratory and Critical Care Medicine, Second Affiliated Hospital, Dalian Medical University, Dalian, China** (E Li MD, Q Wang MD); **Department of Radiology, First Affiliated Hospital, Anhui Medical University, Hefei, China** (X Li MD); **Department of Radiology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China** (J Cao MD,

J Shi MD); **Department of Respiratory and Critical Care, Beijing Shijitan Hospital, Capital Medical University, Peking University Ninth School of Clinical Medicine, Beijing, China** (H Deng MA, X Xue MD); **Department of Respiratory and Critical Care** (M Xie MD, X Ma MD) **and College of Pulmonary and Critical Care Medicine** (F Xie MD), **Chinese PLA General Hospital, Beijing, China; Department of Thoracic Surgery, Tangdu Hospital of Air Force Military Medical University, Xi'an, China** (Y Wang MD, X Xue); **Department of Thoracic Surgery, Affiliated Hospital of Qingdao University, Qingdao, China** (Y Wang); **Department of Clinical Research Management** (J Xian BS) **and Integrated Care Management Center** (J Ren BS), **West China Hospital, Sichuan University, Chengdu, China; Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China** (Z Li MD, Z Liu MD); **Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China** (Z Li, Z Liu); **Department of Radiology, Third Affiliated Hospital, Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, China** (Z Li); **Department of Respiratory Medicine, Weifang Medical university, Weifang, China** (X Xue)

Correspondence to: Prof Jie Tian, Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Engineering Medicine, Beihang University, Beijing 100191, China
**jie.tian@ia.ac.cn**

## Research in context

### Evidence before this study

We searched PubMed on Aug 9, 2021, for research articles that contained the terms "EGFR mutation" AND "whole-lung analysis" AND ("artificial intelligence" OR "deep learning") AND "lung cancer", without date or language restrictions. We found no studies using a whole-lung analysis method to predict epidermal growth factor receptor (EGFR) genotype in lung cancer. We further searched the terms "EGFR mutation" AND ("artificial intelligence" OR "deep learning") AND "lung cancer" AND "CT". We found 14 studies that used deep learning and CT imaging to predict EGFR mutation status in patients with lung cancer. The largest dataset obtained so far included 914 patients in a retrospective cohort from an Asian population, without a large-scale, prospective, and global dataset for a thorough validation. Moreover, the published studies used tumour-based methods, which require manual tumour annotation in CT images, making the methods inconvenient to use in clinical practice. Some global lung characteristics have been reported to be associated with EGFR genotype and can affect treatment response; however, this information is ignored in tumour-based analyses. Additionally, these studies predicted EGFR mutation status alone, and did not identify patients with an EGFR mutation who are at high risk of having resistance to tyrosine kinase inhibitors (TKIs). The biological mechanisms behind the prognostic value of CT features also needs further exploration and explanation.

### Added value of this study

We proposed a novel fully automated artificial intelligence system (FAIS) to mine whole-lung information for genotype analysis, which outperformed commonly used tumour-based deep learning methods. This system was trained and validated in large-scale datasets, which are nearly 20-times larger than previous studies. Moreover, FAIS learned to identify patients with an *EGFR* mutation who are at high risk of having TKI resistance, which is a good supplement to gene sequencing. Furthermore, we assessed global lung features using gene pathway analysis in RNA-sequencing data, and we observed associations between the CT imaging phenotype mined in the whole lung and drug resistance or cancer progression mechanisms.

### Implications of all the available evidence

FAIS provides a very convenient method to non-invasively detect *EGFR* genotype and identify patients with an *EGFR* mutation who are at high risk of having TKI resistance. The superior performance of the whole-lung analysis over the commonly used tumour-based methods suggests that genotype and prognostic information correlates with macro-level changes in the whole lung instead of only in tumour tissues. Further gene pathway analysis identified several biological mechanisms involved in the association between global lung characteristics and genetic activities.

*EGFR* mutation status, using either the quantitative radiomic method or a deep learning model. Radiomic methods first extract hand-crafted image features of the tumour and then select some important features to train a machine learning model to predict *EGFR* genotype.[7,12] Deep learning methods integrate feature extraction and model building processes into a unified convolutional neural network[13,14] and can learn more effective features from tumour image automatically by changing the convolutional neural network architecture.[15] Most existing methods have only focused on analysing tumour tissues in CT imaging; however, many tissues and the environment outside the tumour can also reflect genotype information and might affect therapeutic efficacy. For instance, previous studies have shown that the characteristics of lungs (eg, pleural retraction and convergence of surrounding vessels) are associated with *EGFR* genotype.[16–19] Additionally, lung functional changes or abnormalities might also affect treatment efficacy, which are ignored in tumour-based methods. Most importantly, previous studies only predicted *EGFR* mutation status, but they had weak ability to stratify *EGFR*-mutant genotype according to their prognosis for targeted therapy. Moreover, previous tumour-based methods relied on time-consuming image annotation and were validated in small retrospective datasets (fewer than 1000 patients); a fully automated method that is validated in large, prospective datasets including different ethnicities is more desired for clinical practice. Lastly, the associations between high-dimensional CT features and genetic-level mechanisms remain unclear to clinicians.

In this study, we propose that genotype and prognostic information can be mined from the whole lung and thus, we aimed to build a novel, fully automated artificial intelligence system (FAIS) that mines whole-lung information to predict *EGFR* genotype and progression-free survival (PFS) of patients after receiving EGFR-TKIs, focused on providing a non-invasive and convenient method that is validated in a large cohort to assist targeted therapy planning. We also aimed to explore biological mechanisms of artificial intelligence methods in inferencing genetic activities.

## Methods

### Study design and participants

The overall study design is captured in figure 1. With approval of the ethics committee of the hospitals, we included 18 232 patients with lung cancer from nine cohorts (table 1), including seven retrospective cohorts collected from eight provinces in China, the public The Cancer Imaging Archive (TCIA) cohort comprising a White population in the USA, and a prospective cohort from China. *EGFR* gene sequencing results and CT images at diagnosis time (20 319 CT sequences) were

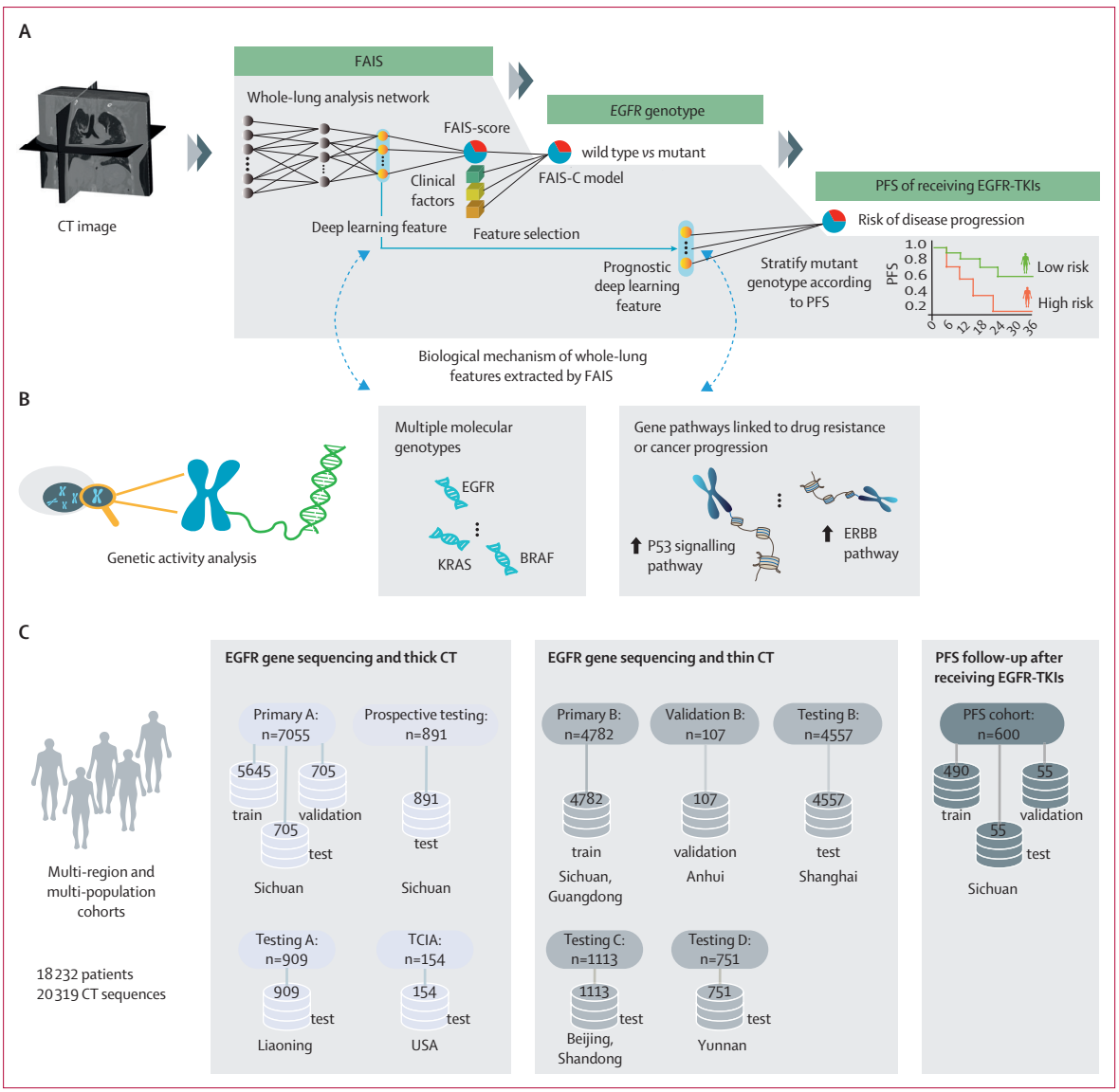


***Figure 1:*** **Workflow of the proposed FAIS and study design**
(A) Inference process of FAIS in predicting EGFR genotype and PFS in patients after receiving EGFR-TKIs. (B) Mining associations between genetic activities and whole-lung features extracted by FAIS. (C) Cohorts used in this study. 2087 patients were included in both the primary cohort A and primary cohort B since they had both thick CT and thin CT. PFS cohort is a subset of primary cohort A. EGFR=epidermal growth factor receptor. FAIS=fully automated artificial intelligence system. PFS=progression-free survival. TKIs=tyrosine kinase inhibitors.

obtained for all patients. Each cohort was collected from a different centre. The primary cohort A, prospective testing cohort, testing cohort A, and TCIA cohort included only thick CT images. The primary cohort B, validation cohort B, and testing cohorts B to D included only thin CT images. Detailed inclusion criteria and data sources are presented in the appendix (p 2).

In clinical practice, CT scanning protocols are highly variable in different hospitals, and CT slice thickness has the largest effect on artificial intelligence-based methods.[20] Consequently, we divided CT images into two categories, thick CT (≥3·75 mm) and thin CT (<3·75 mm), and we trained and tested the proposed model using both thick CT and thin CT. In thick CT images, primary cohort A was randomly split into training, internal validation, and internal testing sets at a ratio of 8:1:1 for model training, hyper-parameter tuning, and internal testing. We used the testing cohort A and the TCIA cohort for independent testing in both Asian and White populations; the prospective testing cohort was used for prospective testing (figure 1). In thin CT images, we used primary cohort B for model training and validation cohort B for model hyper-parameter tuning; testing cohorts B to D were used for multicentre independent testing. We used

See Online for appendix

| | Primary A (n=7055) | Testing A (n=909) | Prospective testing (n=891) | TCIA (n=154) | Primary B (n=4782) | Validation B (n=107) | Testing B (n=4557) | Testing C (n=1113) | Testing D (n=751) | PFS cohort (n=600) |
|---|---|---|---|---|---|---|---|---|---|---|
| CT thickness | Thick CT | Thick CT | Thick CT | Thick CT | Thin CT | Thin CT | Thin CT | Thin CT | Thin CT | Thick CT |
| Data source | Sichuan | Liaoning | Sichuan | USA | Sichuan, Guangdong | Anhui | Shanghai | Beijing, Shandong | Yunnan | Sichuan |
| Age, years | 59·00 (51·00–66·00) | 63·00 (55·00–69·00) | 59·00 (51·50–67·00) | 68·00 (64·00–74·75) | 60·00 (52·00–67·00) | 62·00 (54·50–66·00) | 63·00 (55·00–68·00) | 61·00 (53·00–67·00) | 56·00 (49·00–63·00) | 59·00 (50·00–67·00) |
| Sex | | | | | | | | | | |
| Female | 3250 (46·1%) | 560 (61·6%) | 428 (48·0%) | 52 (33·8%) | 2231 (46·7%) | 49 (45·8%) | 2319 (50·9%) | 595 (53·5%) | 375 (49·9%) | 349 (58·2%) |
| Male | 3805 (53·9%) | 349 (38·4%) | 463 (52·0%) | 102 (66·2%) | 2551 (53·3%) | 58 (54·2%) | 2238 (49·1%) | 518 (46·5%) | 376 (50·1%) | 251 (41·8%) |
| Stage | | | | | | | | | | |
| I–III | 3988 (56·5%) | 570 (62·7%) | 579 (65·0%) | 104 (67·5%) | 3725 (77·9%) | 70 (65·4%) | 4315 (94·7%) | 772 (69·4%) | 453 (60·3%) | 57 (9·5%) |
| IV | 3027 (42·9%) | 313 (34·4%) | 312 (35·0%) | 4 (2·6%) | 1000 (20·9%) | 36 (33·6%) | 1 (<0·1%) | 237 (21·3%) | 293 (39·0%) | 540 (90·0%) |
| NA | 40 (0·6%) | 26 (2·9%) | 0 | 46 (29·9%) | 57 (1·2%) | 1 (0·9%) | 241 (5·3%) | 104(9·3%) | 5 (0·7%) | 3 (0·5%) |
| Smoking status | | | | | | | | | | |
| Former | 2758 (39·1%) | 155 (17·1%) | 203 (22·8%) | 115 (74·7%) | 1525 (31·9%) | 25 (23·4%) | 798 (17·5%) | 331 (29·7%) | 279 (37·2%) | 150 (25·0%) |
| Never | 3834 (54·3%) | 738 (81·2%) | 688 (77·2%) | 39 (25·3%) | 3004 (62·8%) | 81 (75·7%) | 3591 (78·8%) | 687 (61·7%) | 472 (62·8%) | 407 (67·8%) |
| NA | 463 (6·6%) | 16 (1·8%) | 0 | 0 | 253 (5·3%) | 1 (0·9%) | 168 (3·7%) | 95 (8·5%) | 0 | 43 (7·2%) |
| Histology | | | | | | | | | | |
| Adenocarcinoma | 5979 (84·7%) | 885 (97·4%) | 756 (84·8%) | 135 (87·7%) | 3632 (76·0%) | 80 (74·8%) | 3709 (81·4%) | 1038 (93·3%) | 671 (89·3%) | 574 (95·7%) |
| SCC | 723 (10·2%) | 9 (1·0%) | 97 (10·9%) | 16 (10·4%) | 293 (6·1%) | 12 (11·2%) | 558 (12·2%) | 51 (4·6%) | 50 (6·7%) | 16 (2·7%) |
| Others | 353 (5·0%) | 15 (1·7%) | 38 (4·3%) | 3 (1·9%) | 857 (17·9%) | 15 (14·0%) | 290 (6·4%) | 24 (2·2%) | 30 (4·0%) | 10 (1·7%) |
| Tumour location | | | | | | | | | | |
| Left | 2920 (41·4%) | 350 (38·5%) | 377 (42·3%) | NA | 1952 (40·8%) | 38 (35·5%) | 1868 (41·0%) | 496 (44·6%) | 296 (39·4%) | 246 (41·0%) |
| Right | 4088 (57·9%) | 541 (59·5%) | 511 (57·4%) | NA | 2812 (58·8%) | 61 (57·0%) | 2625 (57·6%) | 611 (54·9%) | 451 (60·1%) | 350 (58·3%) |
| Bilateral | 47 (0·7%) | 18 (2·0%) | 3 (0·3%) | NA | 18 (0·4%) | 8 (7·5%) | 64 (1·4%) | 6 (0·5%) | 4 (0·5%) | 4 (0·7%) |
| Tumour family history | | | | | | | | | | |
| Yes | 1138 (16·1%) | 98 (10·8%) | 116 (13·0%) | NA | 550 (11·5%) | 4 (3·7%) | 21 (0·5%) | 129 (11·6%) | 46 (6·1%) | 119 (19·8%) |
| No | 5457 (77·3%) | 795 (87·5%) | 775 (87·0%) | NA | 3981 (83·2%) | 102 (95·3%) | 4361 (95·7%) | 890 (80·0%) | 701 (93·3%) | 438 (73·0%) |
| NA | 460 (6·5%) | 16 (1·8%) | 0 | NA | 251 (5·2%) | 1 (0·9%) | 175 (3·8%) | 94 (8·4%) | 4 (0·5%) | 43 (7·2%) |
| EGFR genotype | | | | | | | | | | |
| Mutant | 3481 (49·3%) | 367 (40·4%) | 439 (49·3%) | 39 (25·3%) | 2917 (61·0%) | 55 (51·4%) | 2650 (58·2%) | 741 (66·6%) | 364 (48·5%) | 600 (100%) |
| Wild type | 3574 (50·7%) | 542 (59·6%) | 452 (50·7%) | 115 (74·7%) | 1865 (39·0%) | 52 (48·6%) | 1907 (41·8%) | 372 (33·4%) | 387 (51·5%) | 0 |

Data are n (%) or mean (SD). EGFR=epidermal growth factor receptor. NA=not applicable. PFS=progression-free survival. SCC=squamous cell carcinoma. TCIA=The Cancer Imaging Archive.

***Table 1:*** **Characteristics of patients in the ten cohorts**

the PFS cohort to assess performance of the proposed model in predicting personalised PFS in patients receiving EGFR-TKIs; this cohort is a subset of the primary cohort A and also includes training, validation, and testing sets (appendix p 2).

### Development of the FAIS

The FAIS included two main components: automated lung segmentation and *EGFR* genotype prediction (appendix p 10). Without manual annotation, FAIS imported original chest CT images of a patient, and directly predicted the corresponding *EGFR*-mutant probability; for a patient with an EGFR mutation, FAIS also predicted their PFS after receiving EGFR-TKIs.

We first used the DenseNet121-FPN model[21–23] to segment lung areas in chest CT images automatically (appendix pp 3–4). To guarantee the acquisition of the complete lung field, we used a three-dimensional bounding box of the segmented lung mask to crop the whole lung area within the CT image, which was defined as a lung region of interest. Afterward, this region of interest was standardised with Z scores and resized to $48 \times 240 \times 360$ voxels for further analysis by our proposed EGFRNet. EGFRNet consists of four dense blocks, where each block is a stack of convolution, batch normalisation, and activation layer. After the last convolutional layer, we used average pooling to generate 768-dimensional deep learning features (appendix pp 10, 13), which were fully connected to the output neuron for predicting the probability (FAIS score) of a given patient having an *EGFR* mutation.

Notably, in the lung region of interest, some other organs (eg, spine) were included. To enable EGFRNet to focus on lung areas, we proposed a lung mask-guided attention mechanism, which enhanced the response of the lung area and suppressed the non-lung areas

(appendix p 5). During model training, we proposed an area under the curve (AUC) loss function that calculated the approximated AUC error in a training batch for auxiliary training (appendix pp 6–7).[24] After model training, we used the Grad-CAM[25] algorithm to visualise the *EGFR* mutation-relevant lung characteristics discovered by FAIS.

To compare the advantage of FAIS over the commonly used tumour-based methods, we built two tumour-based deep learning models as comparisons (appendix p 8), and we compared their performance in the primary cohort A and the TCIA cohort.

### Development of the FAIS-C model

Some clinical factors have been reported to be associated with *EGFR* mutations;[7,8] consequently, we incorporated clinical factors and the FAIS score to build the FAIS-C model. First, we used multivariate least absolute shrinkage and selection operator (LASSO) to select diagnostic clinical factors from a list of clinical factors collected including age, sex, tumour stage, smoking status, histology, tumour location, and tumour family history (table 1). Second, the selected clinical factors and the FAIS score were sent into a support vector machine to build the FAIS-C model for *EGFR* genotype prediction. All these procedures were done in the training set in primary cohort A. For some patients with missing clinical values, we used the mean value or the categorical value with the highest frequency in the training set for imputation.

### PFS prediction using FAIS

Among the 768-dimensional deep learning features extracted by FAIS, we used LASSO-Cox regression to select prognostic features for predicting personalised PFS in patients after receiving EGFR-TKIs, using the training set in the PFS cohort. Finally, the LASSO-Cox model generated a risk score for each patient by use of the selected prognostic features. A larger risk score indicates increased risk of disease progression and corresponds to a shorter PFS. To stratify patients into high-risk and low-risk groups, we used the X-tile software (version 3.6.1) to select the optimal cutoff risk score using the training set data; this cutoff value was subsequently used in the validation and testing sets for model performance evaluation.

### Exploring biological mechanisms of the whole-lung features

We included 891 patients who underwent eight-panel or 56-panel next generation sequencing in the primary cohort A to explore the associations between the 768-dimensional deep learning features and 56 common genotypes. For each genotype, we used independent sample *t* tests or Wilcoxon tests to examine whether a deep learning feature had significant difference between the mutant and wild-type groups. This process was repeated for each deep learning feature and each of the 56 genes.

To explore the biological mechanisms of the prognostic value of FAIS, we included 87 patients in the TCIA cohort that had both RNA-sequencing data and CT imaging for analysis. We first used FAIS to extract the prognostic deep learning features of the patients from the whole lung, and then we did gene pathway analysis to examine the associations between the prognostic deep learning features and important gene pathways (appendix p 9).

### Statistical analysis

Statistical analysis was done with Python, version 3.7. To evaluate performance of FAIS, the FAIS-C model, and tumour-based deep learning models, we used AUC, accuracy, F1 score, precision, and recall as measures. When evaluating prognostic performance of FAIS, we used Kaplan-Meier analysis and log-rank test to assess whether PFS of the high-risk and low-risk groups identified by FAIS had significant differences. All statistical results were considered significant at $p<0·05$.

### Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

## Results

In thick CT images, FAIS achieved AUC values of 0·779 (95% CI 0·761–0·795) in the internal validation set, 0·759 (0·742–0·777) in the internal testing set, 0·770 (0·753–0·786) in the testing cohort A, 0·756 (0·742–0·772) in the prospective testing cohort, and 0·755 (0·709–0·798) in the TCIA cohort (table 2). Notably, FAIS with the proposed lung mask-guided attention mechanism and AUC loss function showed improved performance in the TCIA cohort (appendix p 11). As adenocarcinoma accounts for 80% of lung cancer cases and responds well to EGFR-targeted therapy, we did a stratified analysis to examine performance of FAIS in lung adenocarcinoma and achieved similar results (appendix p 14).

In thin CT images, FAIS achieved AUC values of 0·813 (0·773–0·853) in the validation cohort B, 0·761 (0·755–0·768) in the testing cohort B, 0·797 (0·784–0·812) in the testing cohort C, and 0·748 (0·732–0·765) in the testing cohort D (table 2). In all the centres including thick and thin CT images, FAIS achieved stable performance, indicating the robustness of FAIS regarding various CT scanning protocols.

In the FAIS-C model, three clinical factors (smoking status, sex, and histological subtype) were ultimately selected as important diagnostic factors. By combining the FAIS score and the three clinical factors, the FAIS-C model achieved AUC values of 0·834 (0·795–0·872) in the validation cohort B, 0·788 (0·773–0·802) in the prospective testing cohort, 0·776 (0·738–0·816) in the TCIA cohort, 0·764 (0·747–0·780) in the testing cohort A, 0·800 (0·793–0·807) in the testing cohort B, 0·812

| | AUC (95% CI) | Accuracy (95% CI) | F1 score (95% CI) | Precision (95% CI) | Recall (95% CI) |
|---|---|---|---|---|---|
| **FAIS** | | | | | |
| Primary A—validation | 0·779 (0·761–0·795) | 0·711 (0·693–0·727) | 0·711 (0·693–0·727) | 0·712 (0·696–0·729) | 0·711 (0·693–0·727) |
| Primary A—testing | 0·759 (0·742–0·777) | 0·685 (0·668–0·703) | 0·685 (0·668–0·703) | 0·687 (0·671–0·706) | 0·685 (0·668–0·703) |
| Testing A | 0·770 (0·753–0·786) | 0·723 (0·709–0·737) | 0·724 (0·710–0·739) | 0·728 (0·714–0·743) | 0·723 (0·709–0·737) |
| Prospective testing | 0·756 (0·742–0·772) | 0·697 (0·683–0·711) | 0·697 (0·682–0·711) | 0·697 (0·683–0·712) | 0·697 (0·683–0·711) |
| TCIA | 0·755 (0·709–0·798) | 0·688 (0·650–0·725) | 0·708 (0·673–0·742) | 0·778 (0·744–0·813) | 0·688 (0·650–0·725) |
| Validation B | 0·813 (0·773–0·853) | 0·729 (0·685–0·770) | 0·725 (0·680–0·767) | 0·738 (0·699–0·782) | 0·729 (0·685–0·770) |
| Testing B | 0·761 (0·755–0·768) | 0·704 (0·698–0·711) | 0·702 (0·696–0·709) | 0·702 (0·696–0·709) | 0·704 (0·698–0·711) |
| Testing C | 0·797 (0·784–0·812) | 0·731 (0·719–0·745) | 0·735 (0·723–0·748) | 0·743 (0·731–0·756) | 0·731 (0·719–0·745) |
| Testing D | 0·748 (0·732–0·765) | 0·674 (0·658–0·690) | 0·673 (0·656–0·689) | 0·679 (0·663–0·696) | 0·674 (0·658–0·690) |
| **FAIS-C model** | | | | | |
| Primary A—validation | 0·803 (0·787–0·819) | 0·708 (0·690–0·724) | 0·707 (0·690–0·724) | 0·712 (0·695–0·729) | 0·708 (0·690–0·724) |
| Primary A—testing | 0·797 (0·780–0·814) | 0·725 (0·708–0·742) | 0·725 (0·708–0·742) | 0·730 (0·715–0·748) | 0·725 (0·708–0·742) |
| Testing A | 0·764 (0·747–0·780) | 0·729 (0·714–0·743) | 0·732 (0·716–0·745) | 0·741 (0·727–0·756) | 0·729 (0·714–0·743) |
| Prospective testing | 0·788 (0·773–0·802) | 0·709 (0·694–0·724) | 0·709 (0·694–0·724) | 0·711 (0·697–0·727) | 0·709 (0·694–0·724) |
| TCIA | 0·776 (0·738–0·816) | 0·669 (0·632–0·706) | 0·690 (0·657–0·725) | 0·770 (0·740–0·806) | 0·669 (0·632–0·706) |
| Validation B | 0·834 (0·795–0·872) | 0·757 (0·717–0·797) | 0·756 (0·717–0·797) | 0·758 (0·722–0·802) | 0·757 (0·717–0·797) |
| Testing B | 0·800 (0·793–0·807) | 0·747 (0·740–0·753) | 0·744 (0·737–0·750) | 0·745 (0·738–0·751) | 0·747 (0·740–0·753) |
| Testing C | 0·812 (0·799–0·827) | 0·763 (0·751–0·776) | 0·764 (0·752–0·777) | 0·765 (0·754–0·779) | 0·763 (0·751–0·776) |
| Testing D | 0·763 (0·746–0·779) | 0·676 (0·660–0·694) | 0·674 (0·658–0·692) | 0·685 (0·669–0·703) | 0·676 (0·660–0·694) |
| **Tumour-based deep learning model (2D)** | | | | | |
| Primary A—validation | 0·708 (0·689–0·728) | 0·652 (0·634–0·671) | 0·652 (0·634–0·670) | 0·657 (0·639–0·676) | 0·652 (0·634–0·671) |
| Primary A—testing | 0·758 (0·740–0·775) | 0·697 (0·680–0·715) | 0·696 (0·679–0·714) | 0·707 (0·690–0·725) | 0·697 (0·680–0·715) |
| TCIA | 0·637 (0·589–0·686) | 0·608 (0·569–0·647) | 0·633 (0·597–0·669) | 0·738 (0·703–0·778) | 0·608 (0·569–0·647) |
| **Tumour-based deep learning model (3D)** | | | | | |
| Primary A—validation | 0·712 (0·693–0·729) | 0·662 (0·645–0·678) | 0·662 (0·644–0·678) | 0·667 (0·649–0·683) | 0·662 (0·645–0·678) |
| Primary A—testing | 0·728 (0·710–0·748) | 0·678 (0·661–0·696) | 0·678 (0·661–0·695) | 0·683 (0·667–0·702) | 0·678 (0·661–0·696) |
| TCIA | 0·653 (0·607–0·698) | 0·581 (0·542–0·622) | 0·599 (0·560–0·637) | 0·709 (0·670–0·748) | 0·581 (0·542–0·622) |

2D=two-dimensional. 3D=three-dimensional. AUC=area under the curve. EGFR=epidermal growth factor receptor. FAIS=fully automated artificial intelligence system. TCIA=The Cancer Imaging Archive.

*Table 2:* **The performance of different models in predicting EGFR genotype**

(0·799–0·827) in the testing cohort C, and 0·763 (0·746–0·779) in the testing cohort D (table 2).

As a comparison, the commonly used tumour-based deep learning models (two-dimensional or three-dimensional) showed a somewhat poorer performance (table 2). We further compared performance of FAIS with that of the tumour-based deep learning model (three-dimensional) regarding different training data amounts. Both models performed better as the training set size increased (appendix p 11). However, regardless of the training set size, FAIS always outperformed the tumour-based deep learning model, indicating that the presence of an *EGFR* mutation was correlated with macro-level changes in the whole lung instead of in tumour tissues alone.

Consistent with our hypothesis, in many patients, regions identified as *EGFR* mutation-relevant by FAIS were within the tumour area (figure 2A, B) because gene mutation occurs mainly in tumour cells. However, our results suggest that some functional or structural changes in non-tumour tissues of the lungs might be correlated with *EGFR* mutation. In some cases, FAIS discovered that the interaction between a tumour and its surrounding tissues was important for inferencing *EGFR* genotype. For instance, vascular convergence around the tumour (figure 2C, D), pleural retraction (figure 2E, F), and invasion (figure 2G, H) were associated with *EGFR* genotype inferenced by FAIS.

To better understand the advantages of whole-lung analysis over the commonly used tumour-based analysis, we visualised the response of several neuron layers (structures comprising small computational nodes to extract features) in FAIS (defined as deep learning features) to CT images of different patients (appendix p 12). On the one hand, in each neuron layer, some deep learning features automatically focused on the tumour area to mine high-dimensional information from the tumour and its microenvironment. On the other hand, other features extracted complementary information focusing on global lung appearance and functional change. In the last FAIS layer, the deep learning features showed strong associations with *EGFR* genotype, in which patients in the same class remained clustered and clearly separated from those in the other class (appendix p 12).

Although all patients in the PFS cohort were confirmed to have an *EGFR* mutation, they had a large variance in PFS after receiving EGFR-TKIs (median 11·42 months, IQR 6·16–18·20), indicating the necessity of stratifying *EGFR*-mutant genotype according to patients' therapeutic response to targeted therapy. Among patients with an *EGFR* mutation confirmed by gene sequencing, we found that patients predicted to have wild-type *EGFR* by the FAIS-C model had a shorter PFS than those predicted to have an *EGFR* mutation (median 8·40 months, 4·63–15·87, *vs* 12·03 months, 7·03–18·9, log-rank p=0·0019, figure 3A). Consequently, the FAIS-C model learned to stratify the *EGFR*-mutant genotype according

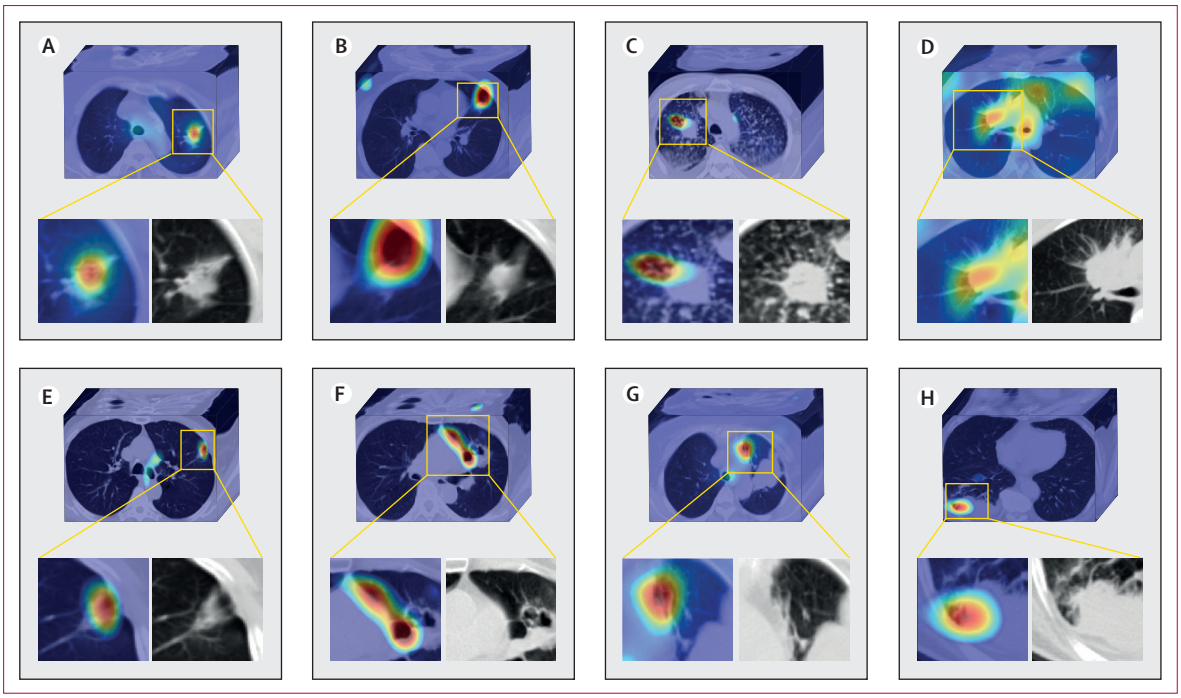


***Figure 2:*** **EGFR mutation-relevant lung characteristics discovered by FAIS**
The figure presents CT images from eight patients. Bright red colour represents high association with EGFR genotype, and dark blue colour represents weak association. EGFR=epidermal growth factor receptor. FAIS=fully automated artificial intelligence system.

to their prognosis to targeted therapy, which is a valuable supplement to gene sequencing.

The prognostic value of the FAIS-C model indicates that some of the 768-dimensional deep learning features in FAIS are able to convey prognostic information about EGFR-targeted therapy. We found 29 prognostic features that showed strong associations with the PFS of patients receiving EGFR-TKIs (appendix p 15). Relying only on the 29 prognostic deep learning features, the model was able to stratify patients with an *EGFR* mutation into high-risk and low-risk groups indicating their risk of disease progression after treatment (log-rank p<0·0001, p=0·035, and p=0·023; figure 3B–D).

Previous studies suggested the use of *EGFR* mutation subtypes to stratify *EGFR*-mutant genotype. For example, patients with exon 19 deletions ($EGFR^{19del}$) have the best prognosis with EGFR-TKI treatment, followed by L858R point mutation in exon 21 ($EGFR^{L858R}$), whereas other uncommon subtypes (eg, $EGFR^{G719X}$, $EGFR^{L861Q}$, and $EGFR^{S768I}$) were associated with a poorer prognosis.[26,27] However, we found no significant difference in patients' PFS between these three mutation subtypes (figure 3E). Moreover, we found that each mutation subtype contained patients with poor prognosis and good prognosis, and FAIS was able to stratify patients with different prognosis in each mutation subtype. For instance, although patients with $EGFR^{19del}$ or $EGFR^{L858R}$ showed good prognosis to EGFR-TKI treatment in previous studies, FAIS identified high-risk groups with poor PFS in these two subtypes (median 8·50 months, 5·07–12·53, with $EGFR^{19del}$ and 8·43 months, 4·35–11·92, with $EGFR^{L858R}$, figure 3F). In uncommon *EGFR* subtypes (denoted as $EGFR^{uncommon}$), FAIS stratified patients into two groups, with the high-risk group showing a short PFS (5·87 months, 3·85–8·50) whereas the low-risk group showed a better prognosis (21·20 months, 13·30–33·08) than that with the $EGFR^{19del}$ and $EGFR^{L858R}$ subtypes.

In the 891 patients in primary cohort A who underwent eight-panel or 56-panel next generation sequencing, we found that many deep learning features extracted from the whole lung were associated with multiple genes (figure 4A). Moreover, 75 deep learning features were significantly associated with more than one gene, suggesting that interactions between genes might exist and can be captured by FAIS.

Notably, we found that the 29 prognostic deep learning features in FAIS were associated with multiple genes that can affect the prognosis of patients receiving EGFR-TKIs (*EGFR*, *KRAS*, *ALK*, *BRAF*, and *ROS1*; figure 4B). Moreover, gene pathway analysis in the TCIA cohort revealed that four prognostic deep learning features were strongly associated with gene pathways linked to drug resistance or cancer progression. This finding might explain why deep learning features were able to predict a personalised prognosis to EGFR-TKI treatment. For example, the 341 deep learning feature corresponded to the upregulated P53 pathway, and the 716 deep learning feature corresponded to the upregulated ERBB pathway.

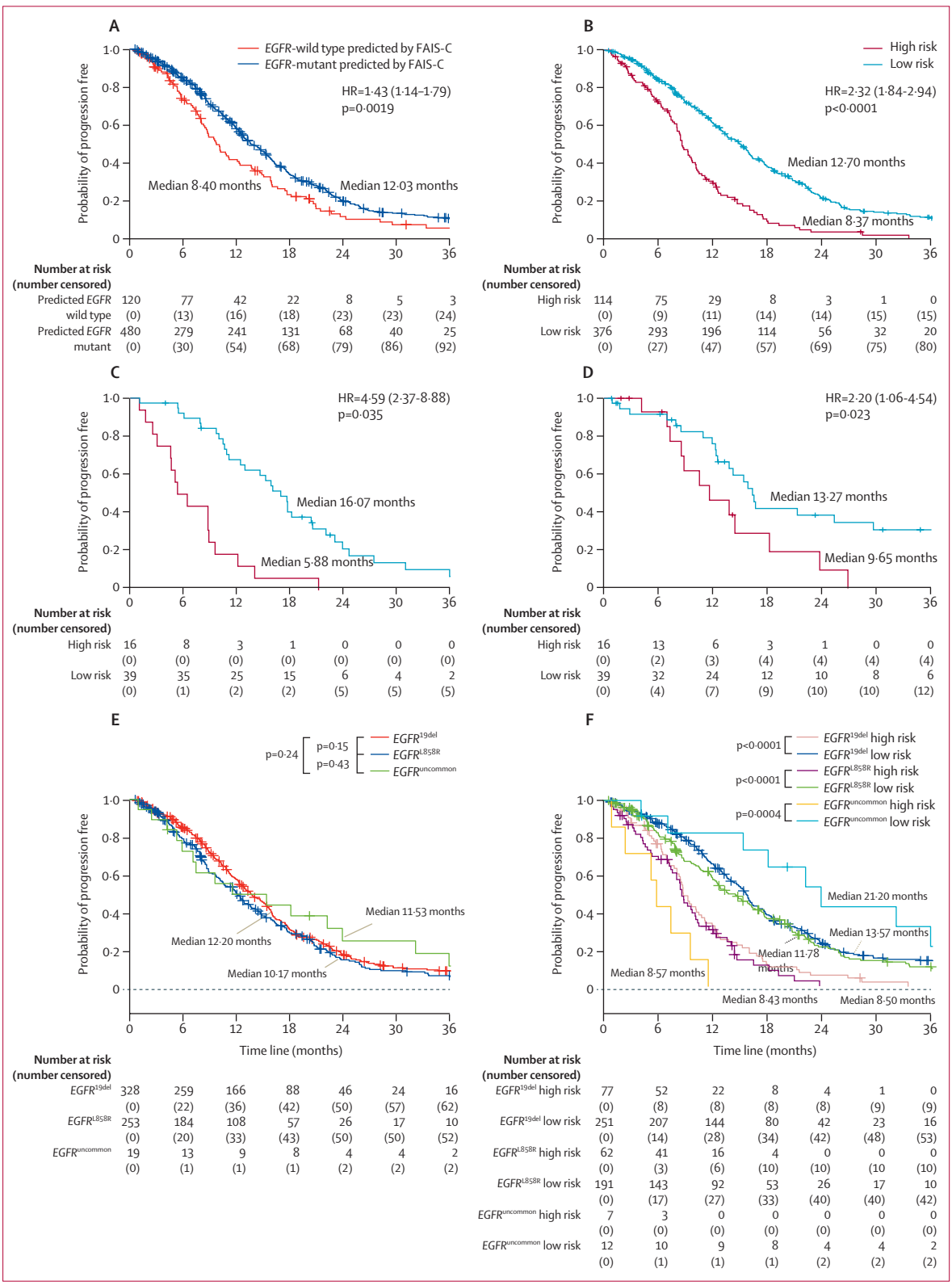


***Figure 3:*** **Kaplan-Meier analysis in the PFS cohort**
The figure presents Kaplan-Meier curves of the EGFR-mutant and EGFR-wild type groups predicted by the FAIS-C model (A); of the high-risk and low-risk groups in the training set (B), validation set (C), and testing set (D); of $EGFR^{19del}$, $EGFR^{L858R}$, and $EGFR^{uncommon}$ subtypes (E); and of the high-risk and low-risk groups stratified by FAIS in each EGFR mutation subtype (F). Vertical lines represent censored data. Log-rank test was used to assess the difference of PFS between different groups. EGFR=epidermal growth factor receptor. FAIS=fully automated artificial intelligence system. HR=hazard ratio. PFS=progression-free survival.

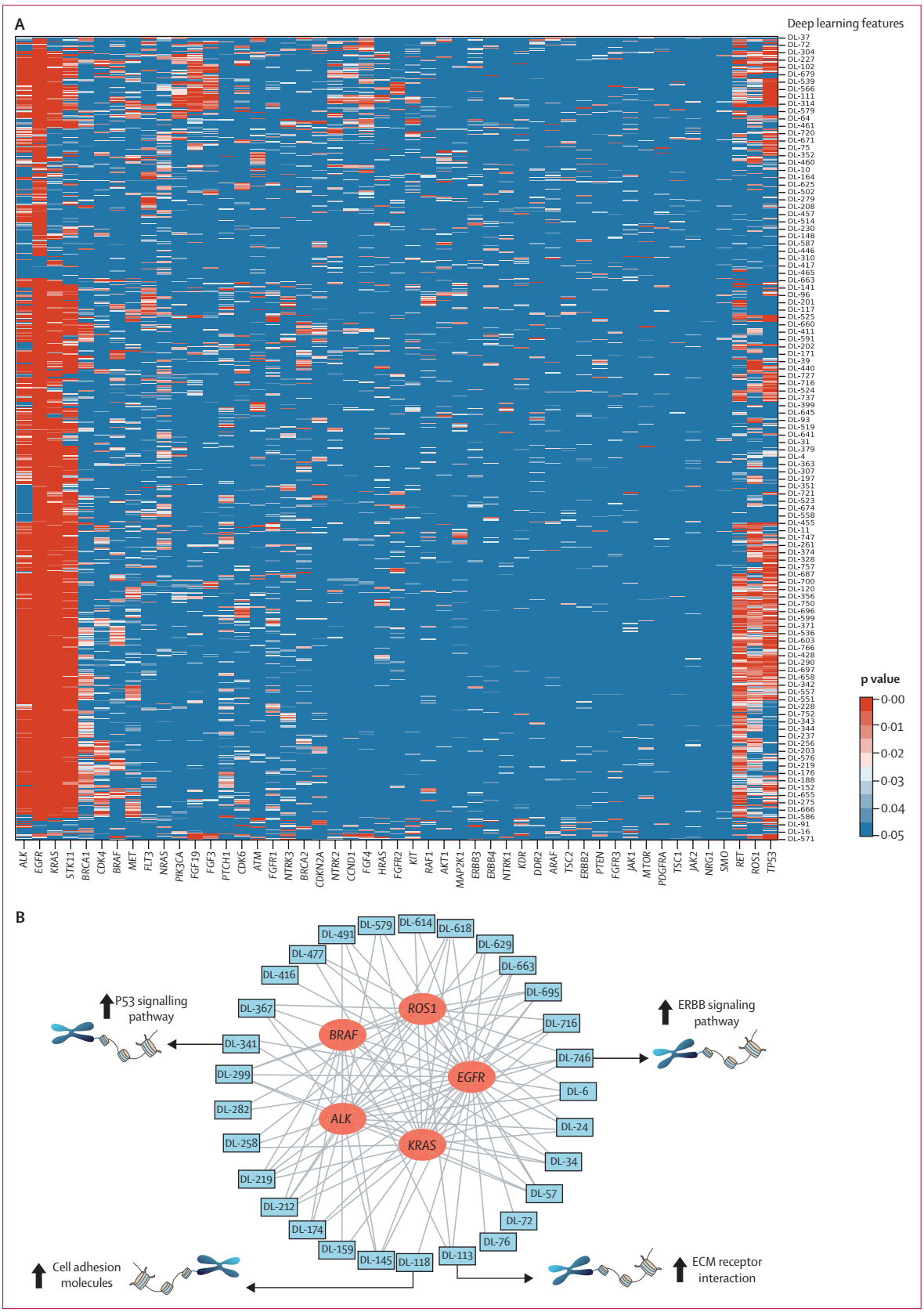


***Figure 4:*** **Association between genetic activities and deep learning features extracted from the whole lung**

The figure presents the association between each deep learning feature and the mutation status of 56 genes (A) and the association between the 29 prognostic deep learning features and the five genes linked to the prognosis with epidermal growth factor receptor-tyrosine kinase inhibitors (B). Lines between circles and blue boxes indicate significant association ($p<0·05$).

These pathways were associated with a resistance mechanism to targeted therapy, suggesting that the 341 and 716 deep learning features might convey information about the risk of drug resistance. Similarly, the 118 deep learning feature was associated with the cell adhesion molecular pathway, which is a sign of cancer metastasis. The 113 deep learning feature was associated with the ECM receptor interaction pathway, which is part of the epithelial mesenchymal transition mechanism, previously linked to cancer progression.

Except for mining genotype and prognostic information, whole-lung analysis has good potential for analysing other clinical outcomes in lung cancer. For instance, we used FAIS to predict lung cancer metastasis and achieved AUC values of 0·826 in the internal validation set and 0·787 in the testing set in primary cohort A (appendix p 16), which outperformed the tumour-based deep learning model (0·765, 95% CI 0·748–0·781, in the internal validation set and 0·746, 0·727–0·765 in the testing set).

## Discussion

*EGFR* genotype and individualised prognosis to EGFR-TKI treatment are crucial for targeted therapy planning in patients with lung cancer. Our findings suggest that routine CT imaging combined with a fully automated whole-lung analysis artificial intelligence system can non-invasively predict *EGFR* genotype and identify patients with an *EGFR* mutation and with a poor prognosis to EGFR-TKI treatment. The performance improvement of FAIS-C model over FAIS suggests that CT imaging provides information that complements clinical factors.

Although gene sequencing is the gold standard for detecting genotypes, it faces the difficulties brought by tumour heterogeneity and invasive biopsy. Most importantly, gene sequencing cannot identify patients with poor prognosis to targeted therapy. In these situations, FAIS could be a good supplement to biopsy sequencing because *EGFR* genotype predicted by FAIS was significantly associated with patients' prognosis with EGFR-TKI treatment. Patients confirmed to have an *EGFR* mutation by both gene sequencing and FAIS showed good prognosis to EGFR-targeted therapy (figure 3A). However, those with a confirmed *EGFR* mutation by gene sequencing, but who were predicted to have wild-type *EGFR* by FAIS showed a poor prognosis (figure 3A). Moreover, FAIS was able to predict personalised PFS for patients receiving EGFR-TKIs, providing a method to stratify *EGFR*-mutant genotype according to patients' therapeutic response, which is a great supplement to gene sequencing.

By contrast with previous artificial intelligence-based studies, FAIS predicts *EGFR* genotype and personalised prognosis simultaneously, and it showed better performance than the commonly used tumour-based methods[7,13] in a large-scale cohort (18 232 patients, over 20-times larger than those of previous studies). Unlike the commonly used tumour-based deep learning model that only extracts tumour information, FAIS mines both tumour information and global lung information, thus achieving better performance. For instance, FAIS mines pleural characteristics for inferencing *EGFR* genotype in some patients (figure 2E–H). These findings are consistent with previous studies,[16,17] where the rate of pleural retraction and invasion differed between patients with an *EGFR* mutation and those with wild-type *EGFR*. Notably, the *EGFR* mutation rate differs between ethnicities (about 15% in a White population and 50% in an Asian population).[9,28,29] To assess the generalisability of the proposed model, we trained the model using data from an Asian population and tested it using data from a White population; FAIS performed well in these different populations. Moreover, this study included all types of lung cancer instead of only adenocarcinoma cases,[7,13] eliminating the need to identify adenocarcinoma when using this system. FAIS is fully automated and does not require any time-consuming CT imaging annotation, which is more convenient for use in clinical practice. Most importantly, we found that genotype and prognostic information can be obtained from the whole lung instead of only from tumour tissues. The effectiveness of FAIS in predicting lung cancer metastasis further showed that many clinical outcomes of lung cancer are probably associated with the macro-level changes in the whole lung instead of only tumour tissues. Consequently, mining whole-lung information should have great potential in analysing lung cancer.

Figure 4A showed that much genotype information can be captured by analysing the whole lung in CT imaging. Although FAIS was trained to predict *EGFR* genotype, the strong associations between the deep learning features and many other genes showed the potentiality of FAIS in predicting multiple gene markers in lung cancer. Further gene pathway analysis revealed that image features mined from the whole lung were associated with multiple important gene pathways linked to drug resistance or cancer progression mechanisms, which explains why the CT features mined from the whole lung conveyed prognostic information that complemented gene sequencing. The associations between CT-derived whole-lung features and gene pathways could help clinicians to better understand the inference process and biological mechanisms of FAIS.

Our study has several limitations. First, in addition to the *EGFR* genotype, other important genes are also relevant to targeted therapy, such as *ALK* and *KRAS*. A method of concurrently predicting several target genes would be valuable. Second, although whole-lung analysis performed better than the tumour-based method, a combination of these methods might achieve optimum performance and requires future research.

In conclusion, FAIS provides a non-invasive method to predict *EGFR* genotype and targeted therapy response through a whole-lung analysis method using CT images, which shows that genotype information can be obtained from the whole lung instead of only tumor tissues.

### Contributors

SW and HY designed this study and wrote the paper. SW, ZW, YZ, and LW built the deep learning models. KW, ZaL, and WW processed and analysed the data. YG, EL, XL, JC, HD, MX, YW, XM, DL, BC, PT, ZQ, JX, JR, FX, ZhL, QW, XX, and JS collected the clinical dataset and performed data preprocessing. JT and WL conceived the project and edited the paper. All authors reviewed and approved the final manuscript for submission. We ensured that all authors had access to all the raw datasets. SW and HY have verified the data and are independent of any company or investor. JT had full access to all the data in the study and had final responsibility for the decision to submit for publication.

### Declaration of interests

We declare no competing interests.

### Data sharing

Data in the TCIA cohort is publicly available at https://wiki.cancerimagingarchive.net/display/Public/NSCLC+Radiogenomics. Other datasets used in this study belong to the involved hospitals. For researchers who want to access these datasets for academic research purposes, please direct your request to the corresponding author or the first author (shuo_wang@buaa.edu.cn). The code used in this study including the FAIS, the tumour-based deep learning models, the DenseNet121-FPN lung segmentation model, and the trained weights of these models is available at https://github.com/wangshuocas and can be accessed by contacting the first author.

### Acknowledgments

This work was supported by the Ministry of Science and Technology of China (2017YFA0205200), National Natural Science Foundation of China (62027901, 81227901, 81930053, 82001913, 91859203, 82072598, 82027805, 81972916, 62176166, 61976223, and 82001986), National Key Development Plan for Precision Medicine Research (2017YFC0910004), Key R&D Program of Guangdong Province (2021B0101420006), National Science Fund for Distinguished Young Scholars (81925023), the National Key R&D Program of China (2021YFF1201003), High-level Hospital Construction Project (DFJHBF202105), Key R&D project of Sichuan Provincial Department of science and Technology (2021YFS0072), Clinical Research Foundation of Shanghai Pulmonary Hospital (FK1944), China Postdoctoral Science Foundation (2019TQ0019, 2020M670101), Interdisciplinary Innovation Project of 135 Project of West China Hospital of Sichuan University (ZYJC21028), Central Guide Place-Free Exploration Project, Sichuan Provincial Department of Science and Technology (2020ZYD005), project of High-Level Talents Team Introduction in Zhuhai City (Zhuhai HLHPTP201703), Applied Basic Research Projects of Yunnan Province, Outstanding Youth Foundation (202101AW070001), and Yunnan Digitalization, Development and Application of Biotic Resource (202002AA100007). We acknowledge the instrumental and technical support of the Multimodal Biomedical Imaging Experimental Platform, Institute of Automation, Chinese Academy of Sciences.

### References

1 Akamatsu H, Toi Y, Hayashi H, et al. Efficacy of osimertinib plus bevacizumab vs osimertinib in patients with EGFR T790M-mutated non-small cell lung cancer previously treated with epidermal growth factor receptor-tyrosine kinase inhibitor: West Japan Oncology Group 8715L phase 2 randomized clinical trial. *JAMA Oncol* 2021; **7:** 386–94.

2 Soria J-C, Ohe Y, Vansteenkiste J, et al. Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. *N Engl J Med* 2018; **378:** 113–25.

3 Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin* 2019; **69:** 127–57.

4 Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017; **14:** 749–62.

5 Bai H, Wang Z, Chen K, et al. Influence of chemotherapy on EGFR mutation status among patients with non-small-cell lung cancer. *J Clin Oncol* 2012; **30:** 3077–83.

6 Mu W, Jiang L, Zhang J, et al. Non-invasive decision support for NSCLC treatment using PET/CT radiomics. *Nat Commun* 2020; **11:** 5228.

7 Rios Velazquez E, Parmar C, Liu Y, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res* 2017; **77:** 3922–30.

8 Girard N, Sima CS, Jackman DM, et al. Nomogram to predict the presence of EGFR activating mutation in lung adenocarcinoma. *Eur Respir J* 2012; **39:** 366–72.

9 Recondo G, Facchinetti F, Olaussen KA, Besse B, Friboulet L. Making the first move in EGFR-driven or ALK-driven NSCLC: first-generation or next-generation TKI? *Nat Rev Clin Oncol* 2018; **15:** 694–708.

10 Sun R, Limkin EJ, Vakalopoulou M, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* 2018; **19:** 1180–91.

11 Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 2019; **25:** 954–61.

12 Rossi G, Barabino E, Fedeli A, et al. Radiomic detection of EGFR mutations in NSCLC. *Cancer Res* 2021; **81:** 724–31.

13 Wang S, Shi J, Ye Z, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J* 2019; **53:** 1800986.

14 Mu W, Jiang L, Zhang J, et al. Non-invasive decision support for NSCLC treatment using PET/CT radiomics. *Nat Commun* 2020; **11:** 5228.

15 Dong Y, Hou L, Yang W, et al. Multi-channel multi-task deep learning for predicting *EGFR* and *KRAS* mutations of non-small cell lung cancer on CT images. *Quant Imaging Med Surg* 2021; **11:** 2354–75.

16 Hasegawa M, Sakai F, Ishikawa R, Kimura F, Ishida H, Kobayashi K. CT features of epidermal growth factor receptor–mutated adenocarcinoma of the lung: comparison with nonmutated adenocarcinoma. *J Thorac Oncol* 2016; **11:** 819–26.

17 Liu Y, Kim J, Qu F, et al. CT features associated with epidermal growth factor receptor mutation status in patients with lung adenocarcinoma. *Radiology* 2016; **280:** 271–80.

18 Rizzo S, Petrella F, Buscarino V, et al. CT radiogenomic characterization of EGFR, K-RAS, and ALK mutations in non-small cell lung cancer. *Eur Radiol* 2016; **26:** 32–42.

19 Liu Y, Kim J, Balagurunathan Y, et al. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas. *Clin Lung Cancer* 2016; **17:** 441–48.e6.

20 Meyer M, Ronald J, Vernuccio F, et al. Reproducibility of CT radiomic features within the same patient: influence of radiation dose and CT reconstruction settings. *Radiology* 2019; **293:** 583–91.

21 Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. IEEE Conference on Computer Vision and Pattern Recognition; July 21–26, 2017.

22 Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. IEEE Conference on Computer Vision and Pattern Recognition; July 21–26, 2017.

23 Wang S, Zha Y, Li W, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J* 2020; **56:** 2000775.

24 Yuan Z, Yan Y, Sonka M, Yang T. Large-scale robust deep AUC maximization: a new surrogate loss and empirical studies on medical image classification. IEEE/CVF International Conference on Computer Vision; Oct 11–17, 2021 (abstr 2911).

25 Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: visual explanations from deep networks via gradient-based localization. *Proc IEEE Int Conf Comput Vis* 2017; **2017:** 618–26.

26 Takamochi K, Oh S, Matsunaga T, Suzuki K. Prognostic impacts of EGFR mutation status and subtype in patients with surgically resected lung adenocarcinoma. *J Thorac Cardiovasc Surg* 2017; **154:** 1768–1774.e1.

27 Sutiman N, Tan SW, Tan EH, et al. EGFR mutation subtypes influence survival outcomes following first-line gefitinib therapy in advanced Asian NSCLC patients. *J Thorac Oncol* 2017; **12:** 529–38.

28 Wu Y-L, Cheng Y, Zhou J, et al. Tepotinib plus gefitinib in patients with EGFR-mutant non-small-cell lung cancer with MET overexpression or MET amplification and acquired resistance to previous EGFR inhibitor (INSIGHT study): an open-label, phase 1b/2, multicentre, randomised trial. *Lancet Respir Med* 2020; **8:** 1132–43.

29 Leonetti A, Sharma S, Minari R, Perego P, Giovannetti E, Tiseo M. Resistance mechanisms to osimertinib in EGFR-mutated non-small cell lung cancer. *Br J Cancer* 2019; **121:** 725–37.