Value of CT features for predicting EGFR mutations and ALK positivity in patients with lung adenocarcinoma

Xiaoyu Han1,2,4, Jun Fan3,4, Yumin Li1,2,4, Yukun Cao1,2, Jin Gu1,2, Xi Jia1,2, Yuhui Wang1,2* & Heshui Shi1,2*

The aim of this study was to identify the relationships of epidermal growth factor receptor (EGFR) mutations and anaplastic large-cell lymphoma kinase (ALK) status with CT characteristics in adenocarcinoma using the largest patient cohort to date. In this study, preoperative chest CT findings prior to treatment were retrospectively evaluated in 827 surgically resected lung adenocarcinomas. All patients were tested for EGFR mutations and ALK status. EGFR mutations were found in 489 (59.1%) patients, and ALK positivity was found in 57 (7.0%). By logistic regression, the most significant independent prognostic factors of EGFR effective mutations were female sex, nonsmoker status, GGO air bronchograms and pleural retraction. For EGFR mutation prediction, receiver operating characteristic (ROC) curves yielded areas under the curve (AUCs) of 0.682 and 0.758 for clinical only or combined CT features, respectively, with a significant difference ($p < 0.001$). Furthermore, the exon 21 mutation rate in GGO was significantly higher than the exon 19 mutation rate ($p = 0.029$). The most significant independent prognostic factors of ALK positivity were age, solid-predominant-subtype tumours, mucinous lung adenocarcinoma, solid tumours and no air bronchograms on CT. ROC curve analysis showed that for predicting ALK positivity, the use of clinical variables combined with CT features (AUC = 0.739) was superior to the use of clinical variables alone (AUC = 0.657), with a significant difference ($p = 0.0082$). The use of CT features for patients may allow analyses of tumours and more accurately predict patient populations who will benefit from therapies targeting treatment.

Abbreviations

CT  Computed tomography
EGFR  Epidermal growth factor receptor
ALK  Anaplastic large-cell lymphoma kinase
NSCLC Non-small cell lung carcinoma
TKIs  Tyrosine kinase inhibitors
GGOs Ground-glass opacities
pGGO Pure GGO
mGGO Mixed GGO
PFS Progression-free survival
TNM Tumour node metastasis
ROC Receiver operating characteristic
AUC Area under the curve
IASLC International Association for the Study of Lung Cancer
ARMS Amplified drug resistance mutation system
IHC Immunohistochemistry

1Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1277 Jiefang Rd, Wuhan, Hubei Province 430022, The People's Republic of China. 2Hubei Province Key Laboratory of Molecular Imaging, Wuhan 430022, China. 3Department of Pathology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1277 Jiefang Rd, Wuhan, Hubei Province 430022, The People's Republic of China. 4These authors contributed equally: Xiaoyu Han, Jun Fan and Yumin Li. *email: wangyuhui1983@163.com; heshuishi@hust.edu.cn
Lung cancer is one of the leading causes of cancer-related death worldwide. In China, approximately 733,300 patients were diagnosed with lung cancer in 2015, with 610,200 deaths, and the number of lung cancer-related deaths is expected to exceed one million by 2025.3,4 In recent decades, the emergence of new therapies targeting signalling pathways activated by genetic alterations has revolutionized a new treatment approach for non-small-cell lung cancer (NSCLC), especially adenocarcinoma. The two most common druggable targets are epidermal growth factor receptor (EGFR) mutations and anaplastic lymphoma kinase (ALK) rearrangement in lung adenocarcinomas. EGFR mutations are associated with tumour sensitivity to EGFR-tyrosine kinase inhibitors (TKIs)3.

For tumours with mutations in exons 18, 19, 21, and 20 of the EGFR gene, 80% gefitinib is effective 4,5, but it is useless for tumours with EGFR wild-type mutations6. Crizotinib was the first drug approved for adenocarcinomas harbouring ALK rearrangement7. Randomized clinical trials have demonstrated longer progression-free survival (PFS) following treatment with TKIs than chemotherapy in advanced NSCLC harbouring EGFR mutations or ALK rearrangements7,8. Therefore, it is critical to determine EGFR mutation and ALK statuses prior to the use of TKIs in patients with adenocarcinoma. Two types of methods are currently available: “screening” assays and “specific” methods9,10. However, both methods for detecting EGFR mutations or ALK status are costly and not feasible for every patient with lung cancer. Accordingly, clinical factors are needed to enrich the analysis of gene status in patients with nonresectable lung adenocarcinomas by computed tomography (CT) because this imaging is readily available (Table 1).

Although several studies have investigated relationships between CT imaging features and EGFR mutation11–15 and ALK status16–19, such associations in lung adenocarcinoma are conflicting due to the small sample sizes. Glynn et al. and Sugano et al.13,14 found no significant association between CT features and EGFR mutations, whereas Liu et al.15 reported that a ground glass opacity (GGO) appearance and 15 other CT features were significantly associated with EGFR mutations. Furthermore, the difference in CT features between EGFR exon 21- and 19-mutated adenocarcinomas remains unclear, since it has been proven that TKIs show different targeted effects in cases of EGFR exon 21 and -19 mutations.20

Consequently, our study reviewed data for 827 patients with adenocarcinoma to assess the association between CT and clinical characteristics and EGFR and ALK mutations. We also explored the difference in clinical and CT features between EGFR exon 21- and 19-mutated adenocarcinomas.

Results
Clinical characteristics. A total of 827 eligible patients (average age, 59 ± 9 years; 418 males) were included in the study, and their clinical and pathological characteristics are summarized in Table 2. EGFR mutations were found in 489 (59.1%) patients, and exon 21, 19, 20, and 18 mutation rates were 49.5%, 44.0%, 3.1% and 3.4%, respectively. ALK positivity was found in 57 (7.0%) patients. Six patients had concomitant EGFR mutations and ALK positivity (0.7%). EGFR mutations were more common in females than in males (p < 0.001) and in those who had never smoked (p < 0.001). No significant association was found between age or TNM stage and EGFR mutation (p = 0.320 and p = 0.831, respectively). Pathologically, EGFR mutation was associated with a high frequency of the lepidic predominant subtypes (p < 0.001) and a low frequency of lymph node metastasis proven by surgery (p = 0.006). Pleural invasion did not differ between patients with wild-type and mutant disease (p = 0.268). ALK positivity was found more frequently in younger patients (p < 0.001) (Fig. 1a), and the optimal cut-off value for age was 56 years old (Fig. 1c). A high frequency of solid growth or mucus patterns in ALK-positive tumours was observed in the present study. There was no significant difference in sex, smoking history, TNM stage, lymph node metastasis or pleural invasion between patients with wild-type and mutant EGFR (p = 0.268). The concordance between the two observers was also good, with k coefficients ranging between 0.613 and 0.984 (Table 3).

Correlation of EGFR mutations and ALK positivity with CT features. In regard to radiologic features, tumours with EGFR mutations tended to be smaller than EGFR-wild-type tumours (p < 0.001) (Fig. 1b), and the optimal cut-off value of the tumour maximum diameter was 33 mm (Fig. 1d). EGFR-mutated tumours also tended to have peripheral lesions (p = 0.035) and GGOs (p < 0.001), especially mixed GGOs (mGGOs) (p = 0.045), well-defined margins (p = 0.022), air bronchograms (p < 0.001), pleural retraction (p < 0.001) and no lymphadenopathy (p = 0.001) (Fig. 2). No other CT signs were associated with EGFR mutation status (Table 4). ALK positivity was associated with solid tumours (p = 0.009) (Fig. 3) and the absence of air bronchograms (p < 0.001) and with metastasis (p < 0.001) (Fig. 3). No other CT signs were associated with ALK rearrangement status (Table 4).

Differences in CT features between EGFR exon 21 and -19 mutations. The exon 21 mutation rate in GGOs was significantly higher than the exon 19 mutation rate (34% vs 24%, p = 0.029). However, no differences in sex, smoking history, predominant subtype, type, tumour maximum diameter, air bronchogram, margin, or lymphadenopathy were found between patients with EGFR exon 19 and exon 21 mutations (Table 5).
Multivariable and ROC curve analyses of prognostic factors for EGFR mutations. In the model including both clinical variables and CT features, regression showed that the most significant independent prognostic factors of EGFR + were female sex (OR = 1.713, 95% CI: 1.117, 2.628), non-smoking status (OR = 0.557, 95% CI: 0.357,0.868), GGO (OR = 3.035, 95% CI: 1.841, 5.004), air bronchograms (OR = 1.912, 95% CI: 1.336, 2.737) and pleural retraction (OR = 2.183, 95% CI:1.557, 3.061) (Table 6). ROC curve analysis yielded area under the curve (AUC) values of 0.682 and 0.758 for clinical only or combined CT features, respectively, for the prediction of EGFR mutation, and a significant difference was found between them ($p$ < 0.0001) (Fig. 4a).

Multivariable and ROC curve analyses of prognostic factors for ALK mutations. In logistic regression analysis, age (OR = 0.940, 95% CI: 0.910, 0.972), tumours with the solid-predominant subtype or mucinous lung adenocarcinoma (OR = 7.994, 95% CI: 4.183, 15.278), solid tumours (OR = 0.292, 95% CI: 0.997, 0.878) and lesions without air bronchograms (OR=0.307, 95% CI: 0.123, 0.767) were significantly associated with ALK positivity (Table 7). ROC curve analysis showed that the use of clinical variables combined with CT features (AUC=0.739) was superior to the use of clinical variables alone (AUC=0.657) for the prediction of ALK positivity, and a significant difference was found between them ($p=0.0082$) (Fig. 4b).

Correlation analysis of histopathology subtype with lesion texture on CT. Tumours displaying GGOs on CT correlated positively with the lepidic-predominant subtype ($\beta=0.325$, $p<0.001$). A positive correlation was also found between lesions with a solid appearance on CT and the solid-predominant subtype or mucinous adenocarcinoma ($\beta=0.363$, $p<0.001$).

---

| Variable                  | Definition                                                                 |
|---------------------------|-----------------------------------------------------------------------------|
| Type                      | Central, tumour located in the segmental or more proximal bronchi; peripheral, tumour located in the subsegmental bronchi or more distal airway |
| Location                  | The distribution of each lesion in the lung was recorded, including left upper lobe, left lower lobe, right upper lobe, right middle lobe and right lower lobe |
| Tumour size               | Longest diameter of the tumour in MPR images                                |
| Texture                   | Predominantly solid. Tumour solid component/ground glass component $>0.5$; Predominantly ground glass opacity, tumour solid diameter/ground glass diameter $<0.5$ |
| Ground glass opacity      | Ground glass dense nodules with internal vessels and bronchi visible        |
| Mix ground glass opacity  | Composition of both ground glass opacity and solid                          |
| Pure ground glass opacity | Composition of ground glass opacity only                                     |
| Shape                     | Indicated as lobulated, others (round, or oval)                             |
| Lobulated                 | The surface of the tumour showed as multiple arc-shaped projections         |
| Spiculate                 | Evaluated in the lung window, and indicated as different degrees of spinous or burr-like protrusions at the tumour margin |
| Margin definition         | Evaluated in the lung window, and indicated as well-defined, or poor-defined |
| Air bronchogram           | Tube like or branched air structure within the tumour                        |
| Bubble-like lucency       | The 1–3 mm of air density area within the mass                              |
| Margins                   | Evaluated in the lung window, and indicated as smooth, or spiculated        |
| Heterogeneity             | Evaluated in the soft tissue window, and heterogeneity indicated as the difference of CT values in tumour was greater than 20HU |
| Pleural attachment        | Retraction of the pleura towards the tumour                                 |
| Thickening of the adjacent pleura | Thickening of the adjacent fissural or peripheral pleura pleura            |
| Pleural effusion          | Presence or absence of pleural effusion                                     |
| Cavitation                | Presence or absence of cavitation                                           |
| Intramodular calcifications | Presence or absence of calcifications                                     |
| Necrosis                  | Low-density area in the tumour, without enhancement in enhance CT            |
| Peripheral emphysema      | Presence or absence of peripheral emphysema                                 |
| Peripheral fibrosis       | Pulmonary fibrosis around the tumour                                       |
| Vascular convergence      | Convergence of vessels to the tumour, applied to the peripheral tumours    |
| Enhancement               | “mild” = 0–20 HU; “moderate” = 20–40 HU, ”marked”>40 HU                     |
| Lymphadenopathy           | Presence or absence of lymphadenopathy thoracic lymph nodes (hilari or mediastinal) with short-axis diameter greater than 1 cm |
| Distance metastases       | Including metastasis in bone, brain, liver metastasis, etc. Detected at the same time as the diagnosis of primary tumour or occurred within 6 months following up after surgery |

Table 1. CT features for lung adenocarcinoma.
Discussion

The EGFR mutation rate has been reported to be 27–56% in Asian patients\textsuperscript{13,21,22}. Patients had a similar EGFR mutation rate in our study (59.1%), mainly composed of exons 19 and 21 (93.5%). Previous studies have reported that females and nonsmokers have an increased risk of EGFR mutations\textsuperscript{17,23}, which was confirmed in our series. A recent study found that EGFR mutations more commonly occurred in early-stage NSCLC patients than in advanced-stage patients\textsuperscript{24}. The present study showed that the mutation rate of EGFR (55.6%) in lung adenocarcinoma of TNM stage I-II was similar to that in stage III-IV (56.5%). Such a discrepancy in results may be due to individual differences in the samples included between studies. Moreover, in our study, EGFR mutations were associated with a small tumour size and lymph node metastasis proven by surgery, suggesting a low TNM stage in those cases. In addition, our study demonstrated that EGFR mutations were significantly more common in lepidic predominant adenocarcinomas\textsuperscript{13,14}. This result is supported by gene expression profiling microarray studies in which high EGFR mutation frequencies were observed in terminal respiratory unit adenocarcinoma\textsuperscript{13,25}.

Several studies have explored the correlation between EGFR gene mutations and GGOs on CT\textsuperscript{11–15}. For example, Glynn et al. and Sugano et al.\textsuperscript{11,12} found no significant association between GGOs and EGFR mutations, whereas Liu et al.\textsuperscript{13} reported that a GGO appearance and 15 other CT features were significantly associated with EGFR mutations. In the present study, three main results regarding correlations between GGO and EGFR mutations were identified. First, EGFR mutations were more frequently associated with GGOs by CT, consistent with most previous studies\textsuperscript{10,13,17}. This finding may be due to the inverse relationship between the replication of EGFR and the percentage of GGOs on CT\textsuperscript{26,27}. Moreover, this result was supported by the pathological-imaging correlation in this study: tumours displaying GGOs on CT correlated positively with the lepidic-predominant subtype. Second, our series showed that mixed GGO (mGGO) lesions are more susceptible to EGFR mutations than pure GGO (pGGO) or solid lesions. Hsu et al.\textsuperscript{15} stated that EGFR mutations are more common in invasive solid patterns and significantly less common in pGGO patterns in stage I lung adenocarcinoma. Possible explanations for

| Variable | EGFR + N (%) | EGFR- N (%) | Total N (%) | P | ALK + N (%) | ALK − N (%) | Total N (%) | P |
|----------|--------------|-------------|-------------|---|-------------|-------------|-------------|---|
| Age (years) | 59 ± 8 | 59 ± 9 | 0.320 | 54 ± 9 | 59 ± 9 | 59 ± 9 | <0.001* |
| Sex | | | | | | | | 0.131 |
| Male | 195 (39) | 223 (66) | 418 (51) | 23 (40) | 395 (51) | 418 (51) |
| Female | 294 (61) | 115 (34) | 409 (49) | 34 (60) | 375 (49) | 409 (49) |
| Smoking history | 109 (23) | 165 (49) | 274 (33) | <0.001* | 19 (33) | 255 (33) | 274 (33) | 1.000 |
| TNM stage\textsuperscript{a} | | | | | | | | 0.213 |
| I–II | 272 (56) | 191 (57) | 463 (56) | 27 (47) | 436 (57) | 463 (56) |
| III–IV | 217 (44) | 147 (43) | 364 (44) | 30 (53) | 334 (43) | 364 (44) |
| Histological subtype | | | | | | | | |
| Lepidic predominant\textsuperscript{b} | 48 (10) | 11 (3) | 59 (7) | <0.001* | 4 (7) | 55 (7) | 59 (7) | 1.000 |
| Others subtypes\textsuperscript{c} | 441 (90) | 327 (40) | 768 (93) | 53 (93) | 715 (93) | 768 (93) |
| Solid or Mucinous | 43 (9) | 106 (13) | 149 (18) | <0.001* | 27 (47) | 122 (16) | 149 (18) | <0.001* |
| Others subtypes\textsuperscript{d} | 446 (91) | 232 (69) | 678 (82) | 30 (53) | 648 (84) | 678 (82) |
| Acinar predominant | 261 (65) | 148 (44) | 409 (49) | 16 (28) | 390 (51) | 409 (49) |
| Papillary predominant | 128 (26) | 63 (18) | 324 (39) | 6 (11) | 318 (41) | 324 (39) |
| Solid predominant | 36 (7) | 83 (25) | 119 (14) | 17 (30) | 102 (13) | 119 (14) |
| Mucinous predominant | 7 (1) | 23 (7) | 30 (4) | 10 (18) | 20 (3) | 30 (4) |
| Micropapillary | 8 (2) | 7 (2) | 15 (2) | 3 (5) | 12 (2) | 15 (2) |
| Sieve predominant | 1 (0.2) | 3 (1) | 4 (0.5) | 1 (2) | 3 (0.4) | 4 (0.5) |
| Lymph node metastasis | 134 (27) | 101 (30) | 235 (28) | 0.006* | 18 (32) | 217 (28) | 235 (28) | 0.426 |
| Pleural invasion | 142 (29) | 86 (25) | 228 (28) | 0.268 | 18 (32) | 210 (28) | 228 (28) | 0.539 |

Table 2. Association between clinical characteristics and EGFR and ALK status in adenocarcinoma. * P values < 0.05 were based on comparisons between the two groups; & According to the IASLC 8th TNM Lung Cancer Staging System; $ Lepidic predominant includes: adenocarcinoma in situ, minimally invasive adenocarcinoma, and lepidic predominant invasive adenocarcinoma; @ Other subtypes include: acinar, papillary, micropapillary, and solid predominant adenocarcinoma, as well as variants of invasive mucinous adenocarcinoma; % other subtypes include: lepidic predominant acinar, papillary, micropapillary, as well as variants of invasive adenocarcinoma; EGFR, epidermal growth factor receptor; ALK, anaplastic large-cell lymphoma kinase EGFR +, EGFR mutation; EGFR-, EGFR wild type mutation; ALK +, ALK positive; ALK-, ALK negative.
The above findings may be that EGFR mutations promote the conversion of pGGO to mGGO and indicate more aggressive behaviour. Third, GGOs were significantly more highly related to the exon 21 mutation rate than to the exon 19 mutation rate, in line with Lee et al.\textsuperscript{14}. The results indicated that GGO is not only a factor for EGFR mutation but also an incidence factor for EGFR exon 21 mutation. Because TKIs show different targeted effects in cases of EGFR exon 21 and -19 mutations\textsuperscript{20}, analysis of clinical or imaging variables between these cases can provide a comprehensive baseline for selecting targeted treatment.

In the present study, patients with EGFR mutations frequently presented with smaller tumours (tumour maximum diameter ≤ 33 mm) and peripheral lesions, although Liu et al.\textsuperscript{13} found that EGFR mutations are more common in peripheral lung adenocarcinomas with diameters < 3 cm. This difference may be attributed to the different methods used between the studies, as the cut-off was self-defined in the previous study\textsuperscript{13}. In our study, ROC curve analysis was used to determine the cut-off. In addition, our study found associations between EGFR mutations, air bronchogram and pleural retraction, which agrees with Zhou et al.\textsuperscript{17} and Stefania et al.\textsuperscript{16} but contradicts Mizue et al.\textsuperscript{28}. These controversial results may be the result of different ethnicities, grouping methods and sample sizes. Air bronchograms reflect tumour invasion or expansion without destroying the intratumoural bronchus, suggesting reduced aggressiveness. Pleural retraction is a common sign of visceral pleural invasion, which is one of the most important prognostic factors after the surgical resection of NSCLC\textsuperscript{29}. Nevertheless, in the present study, pleural invasion (proven by surgery) did not differ between wild-type and mutant EGFR. The reason for this difference may be that pleural retraction does not always mean pleural invasion pathologically. Besides, our study showed that EGFR mutation was associated with less lymphadenopathy and less lymph node metastasis pathologically, suggesting a previous study\textsuperscript{13}, suggesting a lower invasiveness of tumours with EGFR mutations.

ALK rearrangement has been identified in 0.4 to 13.5% of unselected NSCLC patients\textsuperscript{30,31}, consistent with the present study (7.0%). Previous studies have reported that patients with ALK positivity tend to be younger and are more often never smokers than patients with ALK-negative tumours\textsuperscript{32}. We found that a younger age (≤ 56 years old) was associated with ALK positivity. However, no significant difference in smoking status was
Table 3. Analysis of interreader agreement percent of concordance and kappa of agreement.

| CT features                      | N (% of concordance) | Kappa (95%CI)        | Kappa interpretation |
|----------------------------------|----------------------|----------------------|----------------------|
| Shape                            | 767/827              | 0.726 (0.655–0.788)  | Substantial          |
| Location                         | 803/827              | 0.828 (0.724–0.906)  | Almost perfect       |
| Texture                          | 788/827              | 0.847 (0.794–0.893)  | Almost perfect       |
| Bubblelike lucency                | 787/827              | 0.797 (0.727–0.852)  | Almost perfect       |
| Margins                          | 777/827              | 0.851 (0.807–0.888)  | Almost perfect       |
| Vascular convergence              | 758/827              | 0.767 (0.707–0.818)  | Almost perfect       |
| Air bronchogram                  | 762/827              | 0.827 (0.781–0.867)  | Almost perfect       |
| Cavitation                       | 809/827              | 0.859 (0.781–0.909)  | Almost perfect       |
| Pleural retraction                | 821/827              | 0.984 (0.969–0.995)  | Almost perfect       |
| Pleural effusion                 | 797/827              | 0.926 (0.893–0.950)  | Almost perfect       |
| Spiculate                        | 782/827              | 0.613 (0.511–0.711)  | Substantial          |
| Calcifications                   | 785/827              | 0.626 (0.510–0.714)  | Substantial          |
| Enhancement degree               | 215/360              | 0.748 (0.677–0.816)  | Substantial          |
| Peripheral emphysema             | 801/827              | 0.894 (0.853–0.928)  | Almost perfect       |
| Peripheral fibrosis              | 779/827              | 0.860 (0.891–0.899)  | Almost perfect       |
| Lymphadenopathy                  | 794/827              | 0.884 (0.842–0.920)  | Almost perfect       |
| Thickening pleura                | 789/827              | 0.749 (0.673–0.818)  | Substantial          |
| Heterogeneity                    | 773/827              | 0.873 (0.828–0.902)  | Almost perfect       |
| Necrosis                         | 747/827              | 0.694 (0.623–0.751)  | Substantial          |
| Metastasis                       | 793/827              | 0.908 (0.874–0.939)  | Almost perfect       |

Figure 2. A 53-year-old female with primary lung adenocarcinoma with EGFR mutation and ALK negative in the right upper lobe (a, b) that showed a mixed ground grass opacity (GGO) appearance with air bronchograms and pleural retraction on CT imaging. Haematoxylin–eosin staining (c) shows the lepidic predominant histological type of adenocarcinoma. The results of the amplification-refractory mutation system (ARMS) method (d) identified an L858R mutation in exon 21 of the EGFR gene in the tumour.
found. A recent report by Li et al.\(^{33}\) conducted on a relatively large sample demonstrated that ALK rearrangements are more commonly observed in the solid predominant subtype of adenocarcinoma. A high frequency of solid growth or mucus patterns in ALK-positive tumours was observed in the present study, consistent with the above report. Previous studies have demonstrated that ALK positivity is more common in stage IV disease (ranging from 9.7–28.0%). In contrast, ALK positivity showed no significant difference between the stage I–II and III–IV groups in our study. We indeed found that ALK positivity was associated with metastasis detected by CT, suggesting a higher TNM staging.

Table 4. Association between CT features and EGFR and ALK status in adenocarcinoma. * P values < 0.05 were based on comparisons between the two groups; \(^a\) The maximum diameter of the lesion (in centimetres) evaluated on multiplanar reconstructions (MPRs) with a soft tissue window; \(^\uparrow\) Comparison between solid and GGO. \(^\downarrow\) Comparison between solid and pGGO. \(^\#\) Comparison between solid and mGGO. \(^!\) Comparison between pGGO and mGGO. EGFR epidermal growth factor receptor; ALK anaplastic large-cell lymphoma kinase; EGFR+ EGFR mutation; EGFR−, EGFR wild-type mutation; ALK+ ALK positive; ALK− ALK negative. GGO ground-glass opacity; pGGO pure ground-glass opacity; mGGO mixed ground-glass opacity.
Figure 3. A 61-year-old female with primary lung adenocarcinomas in the right lower lobe (a, b) who presented as a lobulated solid tumour with concomitant EGFR mutation and ALK rearrangement positive. Th12 vertebral bone metastasis can be seen in the bone window (c). Haematoxylin–eosin staining (d) shows a solid predominant histological type of adenocarcinoma. The results of immunohistochemistry (IHC) the tumour was positive (e); the ARMS method (f) identified EGFR mutations in exons 19 and 20 in the tumour.

Table 5. Differences in clinical and CT features between exon 21 and 19 EGFR-mutated adenocarcinoma. * P values < 0.05 were based on comparisons between the two groups; GGO, ground glass opacity. $ Lepidic predominant includes adenocarcinoma in situ, minimally invasive adenocarcinoma, and lepidic predominant invasive adenocarcinoma; @ other subtypes include acinar, papillary, micropapillary, and solid predominant adenocarcinoma, as well as variants of invasive mucinous adenocarcinoma.
### Table 6. Logistic regression analyses of various predictive factors for EGFR mutation in adenocarcinoma. OR, odds ratio; CI, confidence interval. Lepidic predominant includes adenocarcinoma in situ, minimally invasive adenocarcinoma, and lepidic predominant invasive adenocarcinoma.

| Variables | B       | Wald X² | P value | OR       | 95% CI     |
|-----------|---------|---------|---------|----------|------------|
| Gender    | 0.538   | 6.076   | 0.014   | 1.713    | 1.117–2.628|
| Smoking status | −0.586   | 6.690   | 0.010   | 0.557    | 0.357–0.868|
| Lepidic predominant | 0.493   | 1.495   | 0.221   | 1.637    | 0.743–3.609|
| Lymphatic metastasis | −0.189   | 0.993   | 0.319   | 1.208    | 0.833–1.751|
| Maximum diameter | 0.000   | 0.003   | 0.957   | 1.000    | 0.989–1.012|
| Type      | −0.188  | 0.271   | 0.603   | 0.829    | 0.408–1.683|
| Texture   | 1.110   | 18.935  | 0.000   | 3.035    | 1.841–5.004|
| Air bronchograms | 0.648   | 12.566  | 0.000   | 1.912    | 1.336–2.737|
| Margin    | 0.062   | 0.104   | 0.747   | 1.063    | 0.732–1.545|
| Pleural retraction | 0.781   | 20.502  | 0.000   | 2.183    | 1.557–3.061|
| Lymphadenopathy | −0.307  | 2.005   | 0.157   | 0.736    | 0.481–1.125|

### Table 7. Logistic regression analyses of various predictive factors for ALK-positive status in adenocarcinoma. OR, odds ratio; CI, confidence interval.

| Variables | B         | Wald X² | P value | OR       | 95% CI     |
|-----------|-----------|---------|---------|----------|------------|
| Age       | −0.062    | 13.391  | 0.000   | 0.940    | 0.910–0.972|
| Solid predominant or Mucinous | 2.079   | 39.570  | 0.000   | 7.994    | 4.183–15.278|
| Texture   | −1.230    | 4.806   | 0.028   | 0.292    | 0.097–0.878|
| Air bronchograms | −1.180   | 6.389   | 0.011   | 0.307    | 0.123–0.767|

**Figure 4.** (a) Receiver operating characteristic curves for the prediction of EGFR mutation status using clinical variables alone and combined with CT features in patients with lung adenocarcinoma. (b) Receiver operating characteristic curves for the prediction of ALK-positive status using clinical variables alone and combined with CT features in patients with lung adenocarcinoma.
A limited number of reports have focused on the association between ALK positivity and CT features because of the low rate of ALK positivity in lung adenocarcinoma. According to Zhou et al. and Chang et al., a solid pattern is the main characteristic of ALK-positive tumours. Zhou et al. also found that a pure GGO appearance was significantly less common in ALK-positive cases than in EGFR mutation cases. Similarly, in our series, ALK positivity was associated with solid nodules or masses without GGOs on CT. This imaging finding was consistent with the correlation between solid lesions on CT and solid-predominant subtypes or mucinous adenocarcinoma in the present study, whereas ALK positivity commonly occurred in tumours of solid-predominant subtypes or mucus patterns. Therefore, the correlation between CT signs and gene mutations may be due to different histological growth characteristics.

Previous studies have reported that ALK-positive lung adenocarcinoma frequently occurs with extensive lymph node metastasis, pleural retraction, pleural effusion, and distant metastasis. Among them, distant metastasis in ALK-positive patients was supported by our study, although we did not find significant correlations between ALK positivity and other CT features, which may be caused by the small sample size in the current studies. In addition, the present study found that ALK-positive tumours lack air bronchograms. As mentioned above, the sign of air bronchogram represents reduced tumour invasiveness. When taken together, we may reasonably consider that ALK-positive lung adenocarcinoma is associated with high invasion. The lack of GGO manifestations in our study also supports this biological behaviour.

Our study has some limitations. First, this study was conducted at a single institution, and the patients in our study were all Chinese and thus had a genetic alteration pattern distinct from that of other races, which may impede the application of our results to other ethnicities. To improve the generalization ability and optimization of the model, multidisciplinary and prospective research is needed. Second, due to the limitation of the retrospective analysis method, our study is just a preliminary research study on CT features for predicting gene mutations. However, this report serves as a basis for comprehensive, thorough investigations analysing these patient populations. Third, although we strictly used double-blind methods to record CT signs, EGFR and ALK status, selection bias was inevitable. Finally, the present study analysed only adenocarcinoma and did not include other histologic subtypes, which could explain the results. However, this is understandable, as the majority of EGFR mutations and ALK positivity are found in adenocarcinomas, with an extremely low mutation rate in squamous cell carcinoma (< 5%). Finally, CT findings of distance metastases were not pathologically confirmed. Thus, we did not include distance metastases in the multivariate logistics analysis to predict ALK positivity.

In conclusion, combining clinical variables and CT features was more effective in predicting EGFR and ALK than using clinical variables alone. In addition, GGO is not only a factor for EGFR mutation but also an incidence factor for EGFR exon 21 mutation. Therefore, the use of CT features for patients can allow analyses of tumours and more accurately predict patient populations who will benefit from EGFR-TKIs or ALK crizotinib treatment.

Materials and methods

Patients and inclusion criteria. A total of 1,459 patients evaluated by the multidisciplinary thoracic oncology group between January 2010 and February 2017 at the Union Hospital of Tongji Medical College were retrospectively screened. Initially, 1,186 patients were included according to the inclusion criteria: (1) lung adenocarcinoma confirmed by surgical resection; (2) available pathology reports (including predominant histologic subtypes, lymph node metastasis and pleural invasion, etc.); (3) available results for both EGFR mutations and ALK status; and (4) available clinical data. In total, 359 patients were excluded because of the following three exclusion criteria: (1) thin-section CT was not available (n = 251); (2) heavy CT image artefacts (n = 65); and (3) received preoperative treatment with chemotherapy or radiation therapy before surgery (n = 43). Ultimately, 827 patients were included. The patient's clinical characteristics, including age, sex, smoking history, histopathology, nodal involvement, and tumour stage, among others, were recorded in accordance with Lu et al., non-smoking was defined as lifetime exposure to fewer than 100 cigarettes, and the remaining patients were categorized as ever-smokers. TNM staging was based on the IASLC 8th TNM Lung Cancer Staging System. This retrospective study was approved by the Institutional Review Board of Union Hospital of Tongji Medical College. All subjects enrolled signed a written consent form after being informed of the details of the research. This study was conducted in compliance with the Declaration of Helsinki.

EGFR mutation analysis. EGFR mutations were analysed according to the principle of the amplified drug resistance mutation system (ARMS). Primary tumours or lymph nodes were simply excised, aspirated, or biopsied, followed by 10% neutral buffered formalin fixation and paraffin embedding. DNA was extracted from formalin-fixed paraffin-embedded (FFPE) tissue sections, and the Qiagen FFPE Tissue Kit (Netherlands Roots NV) was used according to the manufacturer's instructions. PCR was carried out using the Mx3000PTM system (Stratagene, La Jolla, USA) with an EGFR 29 Mutations Detection Kit (Amoy Diagnostics, Xiamen, People's Republic of China), and the results were interpreted according to the manufacturer's instructions. Molecular analysis of EGFR mutations was defined as mutations in EGFR exons 18, 19, 21, or 20; other types of EGFR mutations were defined as wild-type EGFR.

VENTANA ALK immunohistochemical (IHC) assay. VENTANA is a fully automated IHC detection method based on the monoclonal antibody D5F3. The VENTANA IHC assay has been approved by the US FDA and the China FDA for the identification of patients with NSCLC who are eligible for treatment with ALK TKIs. FFPE tissue sections with a thickness of 4 μm were cut according to the manufacturer's instructions and a scoring algorithm. The result was dichotomous, whereby the presence of any percentage of positive tumour cells with strong granular cytoplasmic staining was regarded as ALK positivity and all other observations were regarded as ALK negativity.
CT image acquisition. CT was performed at our institution using a multislice spiral CT system (SOMATOM Definition AS+, Siemens Healthineers, Germany)\(^9\). The scan ranged from the level of the chest inlet to the inferior level of the costophrenic angle. The CT parameters were as follows: detector collimation width, 64 × 0.6 mm and 128 × 0.6 mm; tube voltage, 120 kV. The tube current was regulated by an automatic exposure control system (CARE Dose 4D). Images were reconstructed with a slice thickness of 1.5 mm and an interval of 1.5 mm. The reconstructed image is transmitted to the workstation and picture archiving and communication systems (PACS) for multiplanar reconstruction (MPR) postprocessing. Nonionic iodine contrast agents (60–80 ml iohexol 350 mg/ml, Bellu Pharmaceutical Co., Ltd; Beijing, China) at a dose of 3 ml/s were intravenously injected into 360 patients.

Two radiologists with different degrees of experience in interpreting chest CT images independently performed all qualitative image analyses\(^9\). One of them was a senior radiologist with 10 years of chest imaging (H.S.); the other is a fellow with 4 years of experience in CT image interpretation (J.G.). Both analysed the Digital Imaging and Communications in Medicine (DICOM) images from the CT studies without access to clinical and histologic findings but were aware of the presence and sites of tumours. They assessed CT features using both axial CT images and MPR images. After separate evaluations were performed, differences were resolved by consensus. For each CT scan, the data shown in Table 1 were recorded.

Statistical analysis. The analyses were performed using SPSS Statistics (SPSS, version 21, IBM, Chicago, IL, USA) and MedCalc 16.2.0 (MedCalc Software, Mariakerke, Belgium)\(^38\). Distribution normality was assessed using the Kolmogorov–Smirnov test. Normally and nonnormally distributed data and categorical variables are expressed as the mean ± standard deviation, median (interquartile range) and frequency (percentage), respectively. The independent-sample Student's t test was applied to compare two groups of normally distributed variables, and one-way ANOVA and the chi-square test were used to compare categorical variables. Multivariate linear regression analyses (binary logistic regression) were performed to identify independent factors predictive of EGFR or ALK mutation status. The final model was selected by using the enter elimination method, with a cut-off P value of 0.05. A P value < 0.05 (two-tailed) was considered to be statistically significant. Receiver operating characteristic (ROC) curves were constructed for the ability of combined independent factors to predict EGFR mutations or ALK positivity. Comparison of the ROC curve for clinical characteristics alone and clinical characteristics combined with CT signs was performed by the nonparametric approach of DeLong et al. Patient age and the tumour maximum diameter were applied to examine the diagnostic performance of ALK positivity and EGFR mutation by ROC curve analysis. The sensitivity, specificity and optimal cut-off value were calculated. The repeatability test of the maximum tumour diameter was evaluated by intraclass correlation coefficient (ICC) analysis and the 95% CI. For other CT signs, interobserver agreement was assessed by the k coefficient. Pearson's correlation was used to analyse the relationship between histopathology subtype and lesion texture. A P value < 0.05 (two-tailed) was considered to be statistically significant.

Ethics declarations

This study was approved by the ethics committee of Tongji Medical College of Huazhong University of Science and Technology. All subjects provided written informed consent.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 8 October 2020; Accepted: 2 February 2021
Published online: 11 March 2021

References

1. Yang, L., Parkin, D. M., Ferlay, J., Li, L. & Chen, Y. Estimates of Cancer Incidence in China for 2000 and Projections for 2005. Cancer Epidemiol. Biomark. Prev. 14, 243–250 (2005).
2. Chen, W. et al. Cancer statistics in China, 2015. CA Cancer J Clin 66, 115–132 (2016).
3. Jorge, S. E. D. C., Kobayashi, S. S. & Costa, D. B. Epidermal growth factor receptor (EGFR) mutations in lung cancer: preclinical and clinical data. Braz. J. Med. Biol. Res. 47, 929–939 (2014).
4. Hellmann, M. D. et al. Clinical and in vivo Evidence that EGFR S768I Mutant Lung Adenocarcinomas Are Sensitive to Erlotinib. J. Thor. Oncol. 9, e73–e74 (2014).
5. Arcila, M. E. et al. EGFR Exon 20 Insertion Mutations in Lung Adenocarcinomas: Prevalence, Molecular Heterogeneity, and Clinicopathologic Characteristics. Mol. Cancer Ther. 12, 220–229 (2013).
6. Besse, B., Ropert, S. & Soria, J. C. Targeted therapies in lung cancer. Ann. Oncol. 18(9), 135–214 (2007).
7. Solomon, B. J. et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. N. Engl. J. Med. 371, 2167–2177 (2014).
8. Mok, T. S. et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N. Engl. J. Med. 361, 947–957 (2009).
9. Gillian, E. et al. EGFR mutation testing in lung cancer: a review of available methods and their use for analysis of tumour tissue and cytology samples. J. Clin. Pathol. 66, 79–89 (2013).
10. Khoo, C., Rogers, T. M., Fellowes, A., Bell, A. & Fox, S. Molecular methods for somatic mutation testing in lung adenocarcinoma: EGFR and beyond. Transl. Lung Cancer Res. 4, 126 (2015).
11. Glynn, C., Zakowski, M. F. & Ginsberg, M. S. Are there imaging characteristics associated with epidermal growth factor receptor and KRAS mutations in patients with adenocarcinoma of the lung with bronchioloalveolar features?. J. Thor. Oncol. 5, 344–348 (2010).
12. Masayuki, S. et al. Correlation between computed tomography findings and epidermal growth factor receptor and KRAS gene mutations in patients with pulmonary adenocarcinoma. Oncol. Rep. 26, 1205–1211 (2011).
13. Liu, Y. et al. CT features associated with epidermal growth factor receptor mutation status in patients with lung adenocarcinoma. Radiology 280, 1514–1555 (2016).
14. Hyun-Ju, L. et al. Epidermal growth factor receptor mutation in lung adenocarcinomas: relationship with CT characteristics and histologic subtypes. Radiology 268, 254–264 (2013).
15. Hsu, K. H. et al. Epidermal growth factor receptor mutation status in stage I lung adenocarcinoma with different image patterns. J. Thor. Oncol. 6, 1066–1072 (2011).
16. Rizzo, S. et al. CT Radiogenomic characterization of EGFR, K-RAS, and ALK mutations in non-small cell lung cancer. Eur. Radiol. 26, 32–42 (2016).
17. Zhou, J. Y. et al. Comparative analysis of clinicoradiologic characteristics of lung adenocarcinomas with ALK rearrangements or EGFR mutations. Eur. Radiol. 25, 1257–1266 (2015).
18. Halpenny, D. F. et al. Are there imaging characteristics associated with lung adenocarcinomas harboring ALK rearrangements?. Lung Cancer 86, 190–194 (2014).
19. Chang-Min, C. et al. Advanced adenocarcinoma of the lung: comparison of CT characteristics of patients with anaplastic lymphoma kinase gene rearrangement and those with epidermal growth factor receptor mutation. Radiology 275, 272–279 (2015).
20. Remon, J., Morán, T., Reguart, N., Majem, M. & Llanes, P. Beyond EGFR-TKI in EGFR-mutant non-small cell lung cancer patients: Main challenges still to be overcome. Cancer Treat Rev. 40, 723–729 (2014).
21. Guillermo Paez, J. et al. EGFR Mutations in Lung Cancer Correlation with Clinical Response to Gefitinib Therapy. Science 304, 1497–1500 (2004).
22. Maemondo, M. et al. Gefitinib or chemotherapy for non–small-cell lung cancer with mutated EGFR. N. Engl. J. Med. 362, 2380–2388 (2010).
23. Locatelli-Sanchez, M. et al. Routine EGFR molecular analysis in non–small-cell lung cancer patients is feasible: exons 18–21 sequencing results of 753 patients and subsequent clinical outcomes. Lung 191, 491–499 (2013).
24. Tu, W. et al. Radiomics signature: a potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology. Lung Cancer 132, 28–35 (2019).
25. Takeuchi, T. et al. Expression profile–defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors. J. Clin. Oncol. 24, 1679–1688 (2006).
26. Park, E. A. et al. EGFR gene copy number in adenocarcinoma of the lung by FISH analysis: investigation of significantly related factors on CT, FDG–PET, and histopathology. Lung Cancer 64, 179–186 (2009).
27. Lee, Y. et al. Imaging characteristics of stage I non–small cell lung cancer on CT and FDG–PET: relationship with epidermal growth factor receptor protein expression status and survival. Korean J Radiol 14, 375–383 (2013).
28. Hasegawa, M. et al. CT features of epidermal growth factor receptor–mutated adenocarcinoma of the lung: comparison with nonmutated adenocarcinoma. J Thorac Oncol 11, 819–826 (2016).
29. Manac’h, D., Riquet, M., Medioni, J., Pimpec-Barthes, F. L. & Danel, C. Visceral pleura invasion by non-small cell lung cancer: An underrated bad prognostic factor. Ann. Thorac. Surg. 71, 1088–1093 (2001).
30. Benjamin, S., Marilella, V. G. & D Ross C., ALK gene rearrangements: a new therapeutic target in a molecularly defined subset of non–small cell lung cancer. J. Thor. Oncol. 4, 1450–1454 (2009).
31. Rikova, K. et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell 131, 1190–1203 (2007).
32. Shaw, A. T. et al. Clinical features and outcome of patients with non–small-cell lung cancer Who Harbor EML4–ALK. J. Clin. Oncol. 27, 4247–4253 (2009).
33. Li, P. et al. Comparison of clinicopathological features and prognosis between ALK rearrangements and EGFR mutations in surgically resected early-stage lung adenocarcinoma. J Cancer 10, 61–71 (2019).
34. Iv, Z. et al. Value of (18)F–FDG PET/CT for predicting EGFR mutations and positive ALK expression in patients with non–small cell lung cancer: a retrospective analysis of 849 Chinese patients. Eur J Nucl Med Mol Imaging 45, 735–750 (2018).
35. Pao, W. & Girard, N. New driver mutations in non–small-cell lung cancer. Lancet Oncol. 12, 175–180 (2011).
36. Goldstraw, P. et al. The IASLC lung cancer staging project: proposals for? Revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer. J. Thor. Oncol. 11, 39–51 (2016).
37. Antoniello, A., Cafarotti, S., Indimi, A., Galli, A. & Dutly, A. E. EGFR-targeted therapy for non-small cell lung cancer: focus on EGFR oncogenic mutation. Int. J. Med. Sci. 10, 320–330 (2013).
38. Han, X. et al. CT features associated with EGFR mutations and ALK positivity in patients with multiple primary lung adenocarcinomas. Cancer Imaging 20, 51 (2020).

Acknowledgements
We would like to thank all colleagues for helping us during the current study. We are also very grateful to all selfless volunteers who participated in the study. This work was supported by the National Natural Science Foundation of China (grant numbers: 82071921).

Author contributions
Conception and design: H.S., Y.W., X.H., J.F., Y.L.; Collection and assembly of data: X.H., J.F., Y.L., J.G., X.J.; Data analysis and interpretation: X.H., J.F., Y.L., Y.C.; Manuscript writing: All authors; Final approval of manuscript: All authors; Accountable for all aspects of the work: All authors.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to Y.W. or H.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
