



# Clinical Value of *EGFR* Copy Number Gain Determined by Amplicon-Based Targeted Next Generation Sequencing in Patients with *EGFR*-Mutated NSCLC

Jiacong Wei<sup>1,4</sup> · Pei Meng<sup>2,5</sup> · Miente Martijn Terpstra<sup>1</sup> · Anke van Rijk<sup>2</sup> · Menno Tamminga<sup>3</sup> · Frank Scherpen<sup>2</sup> · Arja ter Elst<sup>2</sup> · Mohamed Z. Alimohamed<sup>1,9</sup> · Lennart F. Johansson<sup>1</sup> · Jos Stigt<sup>6</sup> · Rolof P. G. Gijtenbeek<sup>7</sup> · John van Putten<sup>8</sup> · T. Jeroen N. Hiltermann<sup>3</sup> · Harry J. M. Groen<sup>3</sup> · Klaas Kok<sup>1</sup> · Anthonie J. van der Wekken<sup>3</sup> · Anke van den Berg<sup>2</sup>

Accepted: 3 February 2021 / Published online: 19 February 2021  
© The Author(s) 2021

## Abstract

**Background** The clinical relevance of epidermal growth factor receptor (*EGFR*) copy number gain in patients with *EGFR* mutated advanced non-small cell lung cancer on first-line tyrosine kinase inhibitor treatment has not been fully elucidated.

**Objective** We aimed to estimate *EGFR* copy number gain using amplicon-based next generation sequencing data and explored its prognostic value.

**Patients and Methods** Next generation sequencing data were obtained for 1566 patients with non-small cell lung cancer. *EGFR* copy number gain was defined based on an increase in *EGFR* read counts relative to internal reference amplicons and normal controls in combination with a modified *z*-score  $\geq 3.5$ . Clinical follow-up data were available for 60 patients treated with first-line *EGFR*-tyrosine kinase inhibitors.

**Results** Specificity and sensitivity of next generation sequencing-based *EGFR* copy number estimations were above 90%. *EGFR* copy number gain was observed in 27.9% of *EGFR* mutant cases and in 7.4% of *EGFR* wild-type cases. *EGFR* gain was not associated with progression-free survival but showed a significant effect on overall survival with an adjusted hazard ratio of 3.14 (95% confidence interval 1.46–6.78,  $p = 0.003$ ). Besides *EGFR* copy number gain, osimertinib in second or subsequent lines of treatment and the presence of T790M at relapse revealed significant effects in a multivariate analysis with adjusted hazard ratio of 0.43 (95% confidence interval 0.20–0.91,  $p = 0.028$ ) and 0.24 (95% confidence interval 0.1–0.59,  $p = 0.001$ ), respectively.

**Conclusions** Pre-treatment *EGFR* copy number gain determined by amplicon-based next generation sequencing data predicts worse overall survival in *EGFR*-mutated patients treated with first-line *EGFR*-tyrosine kinase inhibitors. T790M at relapse and subsequent treatment with osimertinib predict longer overall survival.

## 1 Introduction

In the past decade, targeted therapies have dramatically improved the clinical management of non-small cell lung cancer (NSCLC), especially of patients with lung adenocarcinoma [1]. Activating variants in epidermal growth factor receptor (*EGFR*) are observed in 10–35% of patients with

### Key Points

Amplicon-based targeted next generation sequencing data can be used to identify epidermal growth factor receptor (*EGFR*) amplifications.

High *EGFR* copy number is not associated with progression-free survival on first-line *EGFR*-tyrosine kinase inhibitors.

High *EGFR* copy number is associated with poor overall survival in T790M + patients treated with second-line osimertinib.

Jiacong Wei and Pei Meng contributed equally to the article.

✉ Anke van den Berg  
a.van.den.berg01@umcg.nl

Extended author information available on the last page of the article

lung adenocarcinoma, with deletions in exon 19 (E19DEL) and the L858R mutation in exon 21 being the most common [2]. Other tyrosine kinase inhibitor (TKI)-sensitive mutations are observed at amino acid positions 719, 768, and 861 [3]. Patients with these activating variants are treated with first-, second-, and third-generation TKIs including erlotinib, gefitinib, afatinib, dacomitinib, and osimertinib [4].

Patients with *EGFR*-sensitive mutations show variable overall survival and progression-free survival (OS and PFS) to EGFR-TKIs [5]. To explore the underlying cause, several studies analyzed the effect of differences in variant allele frequency (VAF) and the presence of *EGFR* amplifications, yet, the effects of VAF on survival to EGFR-TKIs were variable [6–8]. Amplifications were observed more frequently in tumor samples with *EGFR* mutation (range 8–81%) as compared to tumor samples without *EGFR* mutation (range 1–29%) and more frequently involved the mutant allele [9–16]. In an Asian cohort study and a Latino cohort study, patients with concurrent *EGFR* amplification and mutation had a better response to first/second-generation TKI as compared to patients without *EGFR* amplifications [11, 12]. Both studies used fluorescence in situ hybridization (FISH)-based assays to determine the presence of *EGFR* amplifications. However, the limited size of NSCLC biopsies are frequently not sufficient for multiple clinical tests.

To date, next generation sequencing (NGS) data are more commonly used for the detection of copy number variations, and validated protocols are available for whole genome sequencing data and hybridization-based targeted enrichment sequencing data sets [17]. The use of NGS data obtained by amplicon-based target enrichment is more challenging, and a consensus of “best practices” especially for aneuploidy tumor samples still needs to be reached. In a recent study, the ratio of the normalized read counts per amplicon and/or per gene compared to those in normal samples was used as an estimation of the copy number [18]. For targeted NGS data with a limited number of amplicons per gene, a modified approach has been proposed, using median ratio values and a modified  $z$ -score cut-off of 3.5 as recommended in previous studies [19–21]. The modified  $z$ -score is more robust as compared with the normal  $z$ -score because it relies on the median for the calculations and is therefore less susceptible to outliers. The sensitivity and specificity of calling NGS-based amplifications in formalin-fixed paraffin-embedded material were 100% and 99%, respectively. As a high degree of aneuploidy is frequently observed in lung cancer, we anticipated that a comparison to normal samples may result in an overestimation of gains [18, 22]. Moreover, the use of normal control samples as proposed in these studies may potentially be influenced by variability in inter-assay and experimental conditions over time. An alternative approach, using a selection of the amplicons that are not

involved in copy number aberrations as an internal reference is less biased by the aneuploidy state of the tumor sample.

Based on current literature there are no clear guidelines for defining *EGFR* amplification in lung cancer, and it remains unclear whether a gain of *EGFR* copies determined by amplicon-based targeted NGS is a marker for tumor response to first-line EGFR-TKI in *EGFR* mutated cases. In this study, we analyzed an amplicon-based diagnostic IonTorrent hotspot panel dataset of patients with advanced NSCLC. We used amplicon read depth relative to internal reference amplicons and relative to normal samples to identify copy number gains of *EGFR*. In addition, we evaluated whether *EGFR* copy number gain in patients with TKI-sensitive *EGFR* mutations is a prognostic marker for survival to EGFR-TKIs and which approach has a better performance to predict clinical outcome.

## 2 Materials and Methods

### 2.1 Patient/Sample Information

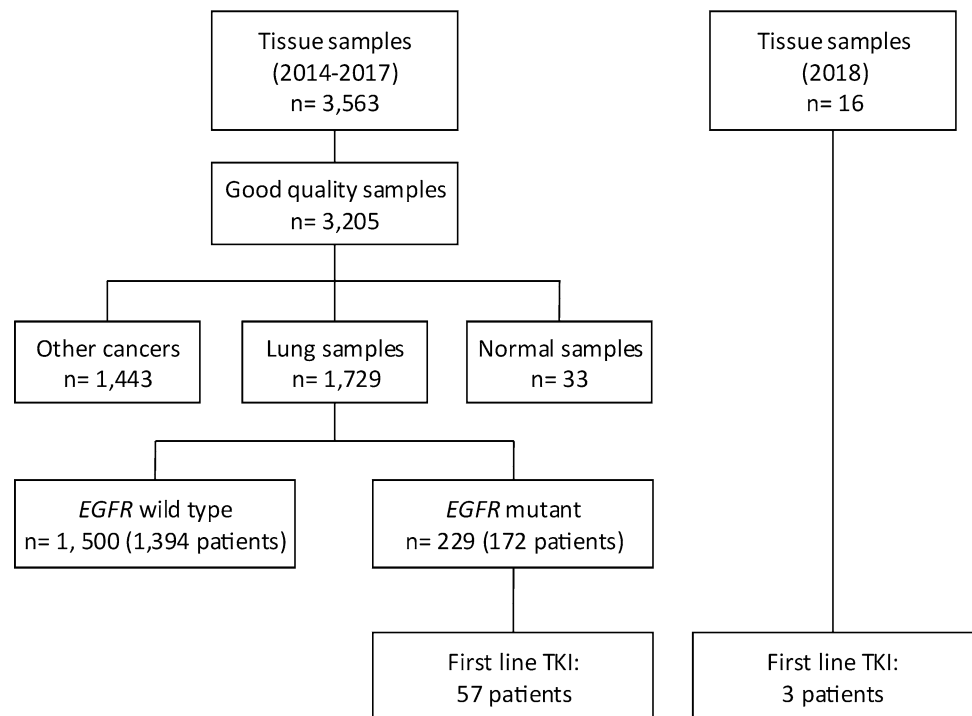
We retrieved data from 3563 diagnostic samples that were subjected to NGS analysis in the period 2014–17 (Fig. 1). Three hundred and fifty-eight samples were excluded based on low coverage (i.e., median read counts per amplicon < 50) resulting in 3205 data sets with sufficient coverage. For 57 *EGFR*-mutated and first-line EGFR-TKI-treated patients, clinical data could be retrieved. For 16 additional patients treated with first-line EGFR-TKI from January until September 2018, clinical data could be retrieved from three cases. Data of these three patients were used only for the survival analysis. Retrieved clinical data included age, sex, smoking, mutation type, TKI use and duration, second and subsequent lines of treatment, presence of T790M mutation in follow-up biopsies, PFS, and OS.

### 2.2 DNA Isolation, Library Preparation, and Sequencing

DNA was isolated from neoplastic cell-rich areas of four to eight 10- $\mu$ m formalin-fixed paraffin-embedded tissue sections to reach a tumor cell percentage > 20%, using Cobas kit (Roche Diagnostic Systems Inc., Branchburg, NJ, USA) according to the instructions of the manufacturer. The DNA concentration was measured by the Qubit™ dsDNA Broad range assay using Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). A minimum of 6 ng of DNA was used as input for the amplicon-based enrichment step. Estimated tumor cell percentages of biopsies from 58 of the 60 patients with clinical data ranged from 20 to 90% (median: 40%).

Two custom-designed AmpliSeq™ panels have been used in the period between 2014 and 2017. The first

**Fig. 1** Flow diagram of patient samples for copy number calling (2014–17) and of patients available for survival analysis (from 2014 to 2017 and 2018). A total of 1566 patients with non-small cell lung cancer were included for copy number gain analysis. Together with the patients from 2018, 60 patients were treated with first-line, epidermal growth factor receptor-tyrosine kinase inhibitors (EGFR-TKIs) and had clinical follow-up data. Clinical analysis was not performed for the other 1506 patients



custom-designed AmpliSeq panel (Design 1) was used from September 2014 to September 2016 and consisted of 30 amplicons covering mutational hotspots of 11 genes (Table S1 of the Electronic Supplementary Material [ESM]). The second design (Design 2) implemented in September 2016 included two separate amplicon pools (Pool 1 and Pool 2) with 44 and 40 different amplicons respectively for 36 genes (Table S1 of the ESM). Sample preparation was performed separately for each pool, whereas for sequencing both pools were combined. Barcoded AmpliSeq libraries were pooled and subjected to emulsion polymerase chain reaction using the OneTouch2 (Life Technologies, San Francisco, CA, USA). Resulting libraries were generated and processed for sequencing on the IonTorrent PGM sequencing system (Life Technologies).

### 2.3 Copy Number Analysis

Read counts for each amplicon were generated using a targeted NGS-based copy number variation detection (CoN-VaDING) pipeline [17]. Of note, as a pre-filtering step, we excluded reads that covered <80% of the amplicon as they cannot reliably be assigned in case of overlapping amplicons. For Design 2, the amplicons from the two polymerase chain reaction pools were analyzed separately and *AMELY*, a sex-differentiating gene located on the Y chromosome, was excluded.

Variability in coverage per amplicon was standardized using amplicon coverage divided by the total read counts of all other amplicons in the library pool. Reference amplicons

were selected per pool (one for design 1 and two for design 2) based on coefficient of variations (CVs) of standardized read counts per amplicon in samples that had a median read count > 50. The 25% amplicons with the lowest CV values across all samples were selected as internal reference amplicons. Starting from the CoN-VaDING-derived read counts, we calculated the coverage of each amplicon relative to the internal reference amplicons.

We next followed two approaches to estimate *EGFR* copy numbers, i.e., (1) within the tumor sample relative to a set of reference amplicons and (2) relative to a set of normal control samples. For comparison within the sample, we calculated the copy number ratio for *EGFR* per amplification pool using the formula: median *EGFR* amplicon read coverage/median reference amplicon read coverage. For design 1, this ratio indicates the relative *EGFR*-specific copy numbers using the within-tumor sample approach. For design 2, we averaged the ratio of the two pools as a measure for the relative *EGFR*-specific copy numbers within-tumor sample approach. Next, we calculated the modified *z*-score within the sample as a measure for the significance of the identified copy number changes [19, 20], by using the previously proposed formula for limited numbers of amplicons per gene:  $0.6745 \times (EGFR \text{ ratio}_{\text{tumor}} - \text{median ratio}_{\text{internal reference amplicons}}) / \text{median absolute deviation (MAD) of ratios}_{\text{internal reference amplicons}}$ . For comparison relative to normal control samples, we first calculated the *EGFR*-specific read count ratio within the normal control samples, following the same approach as described above for the tumor samples.

Next, the *EGFR*-specific ratio of the tumor samples was divided by the median of the *EGFR*-specific ratios of the normal control samples. The *z*-scores relative to normal control samples were calculated as  $0.6745 \times (EGFR \text{ ratio}_{\text{tumor}} - \text{median } EGFR \text{ ratio}_{\text{normal controls}}) / \text{MAD of } EGFR \text{ ratios}_{\text{normal controls}}$ . The optimal cut-off value of the calculated ratios was determined based on the results of the multiplex ligation-dependent probe amplification (MLPA) test (see below). The cut-off for the modified *z*-score was set as  $\geq 3.5$  according to a previous study [20].

## 2.4 MLPA

To determine the optimal cut-off value for the ratios, we performed a MLPA analysis [23] using the SALSA MLPA P105 Glioma-2 probe mix (MRC Holland, Amsterdam, the Netherlands), which is a validated assay in the molecular diagnostics group for glioblastoma. This assay determines *EGFR* copy number gains based on the signals of 11 *EGFR* probe pairs in an assay consisting of a total of 55 probe pairs. Multiplex ligation-dependent probe amplification was performed in accordance with the manufacturer's instruction on the same DNA sample as used for targeted NGS. For each run, we included three normal controls and one sample with a known *EGFR* amplification. DNA samples used for the analysis were retrieved from the molecular diagnostics archive, based on availability. We aimed to include a similar number of cases with and without amplification based on our NGS data. Copy number variation analysis was performed using Coffalyser net software (MRC Holland, Amsterdam, the Netherlands).

## 2.5 Statistical Analysis

Receiver operating characteristic curve analyses to determine the optimal ratio cut-off values based on the MLPA results were performed by both internal and normal comparison approaches using SPSS 23 (IBM SPSS Statistics, Armonk, New York, United States). Percentages of patients with and without *EGFR* gain were calculated using the optimal ratio cut-off and the *z*-score in the total patient group and in the *EGFR* mutant and wild-type subgroups. Progression-free survival was defined as the time between the start of the first TKI treatment and tumor progression or censored for end/loss of follow-up. Overall survival was defined as the time between the start of the first *EGFR*-TKI treatment and death or censored for end/loss of follow-up. Kaplan–Meier survival analysis and univariate and multivariate Cox regression analysis for both PFS and OS were performed using SPSS 23. Variates included in the univariate analysis were age, sex, smoking, *EGFR* variant types, variant allele frequency, first-line treatment drugs, and copy number gain as defined by both approaches. Covariates with  $p < 0.1$  for the

hazard ratio (HR) were included in the multivariate analysis. According to the results of these initial analyses, we subsequently tested the effect of osimertinib in second or subsequent lines of treatment and *EGFR* T790M mutation status after disease progression to first-line *EGFR*-TKI on OS using both univariate and multivariate Cox regression analyses. For these analyses, we changed second or subsequent lines of treatment to a binary variable based on being treated with osimertinib in second, third, or fourth line or no osimertinib. Kaplan–Meier plots to show the time to event distributions were generated by GraphPad Prism. Differences were considered to be statistically significant when the *p* value was 0.05 or less.

## 3 Results

### 3.1 Overview of Patient Samples

In total, we could include 2205 samples analyzed with design 1 and 1000 samples analyzed with design 2. These included 1729 NSCLC samples from 1566 patients, 1443 samples of other malignancies from 1334 patients, and 33 normal samples. We identified 172 patients with NSCLC, with a total of 229 tumor samples being analyzed by NGS (11%) with *EGFR* variants from diagnostic reports (Table S2 of the ESM). E19DELs (including some cases with an E19 insertion combined with a deletion) were the most common with a frequency of 4.2%, followed by the L858R variant with a frequency of 3.3%. Other variants included G719A/C/S with a frequency of 1.1%, S768I with a frequency of 0.4%, and E709A/K and L861Q with a frequency of 0.2% and 0.3%, respectively. Exon 20 INDEL variants (E20INDELs) were observed with a frequency of 1.1%. Variant allele frequencies of the activating *EGFR* mutation had a range of 9–95% (median: 51%), which is slightly higher than the median frequency as can be expected based on the median tumor cell content.

### 3.2 Estimating *EGFR* Copy Number Gains

For design 1, the eight reference amplicons had a CV range of 0.35–0.49. For design 2, the 11 reference amplicons in pool 1 had a CV range of 0.27–0.34 and the ten amplicons from pool 2 had a CV range of 0.26–0.35 (Table S1 of the ESM). Next *EGFR* ratios and *z*-scores were calculated according to internal reference amplicons and to normal control samples. Based on a *z*-score  $\geq 3.5$  and a ratio  $\geq 3.0$ , 49 samples were selected for an independent validation using MLPA. Of these, 22 were scored as amplification positive and 27 as non-amplified based on significantly elevated signals for at least ten out of the 11 *EGFR* probe pairs as calculated by the Coffalyser net

software (Table 1). Receiver operating characteristic curve analysis revealed an optimal cut-off value of 2.8 and 2.3 for the *EGFR* ratio as determined by the internal reference and normal comparison approach, respectively (Fig. S1 of the ESM). These ratios equal 5.6 and 4.6 copies in cases with a diploid genome and 100% tumor cells, respectively. *EGFR* copy number gain was defined by a ratio  $\geq 2.8$  or  $\geq 2.3$  for the internal and normal comparison approaches, respectively, and a modified  $z$ -score  $\geq 3.5$ . Using these criteria for the internal comparison approach, 20 out of 22 samples were correctly scored as *EGFR* amplified, resulting in a sensitivity value of 91%. Using the criteria for the normal comparison approach, 21 out of 22 samples were correctly scored positive indicating a sensitivity of 95%. For one sample, *EGFR* copy number gain was called only for the normal comparison and not for the internal reference amplicon approach.

*EGFR* copy number gains were identified in 151 (9.6%) patients with NSCLC by the internal reference amplicon approach and in 149 samples (9.5%) by the normal control approach. A total of 118 patients were positive for both

**Table 1** Validation of NGS-based *EGFR* copy number gain by MLPA

|       | Internal comparison |      | Normal comparison |      |
|-------|---------------------|------|-------------------|------|
|       | NGS+                | NGS- | NGS+              | NGS- |
| MLPA+ | 20                  | 2    | 21                | 1    |
| MLPA- | 2                   | 25   | 1                 | 26   |

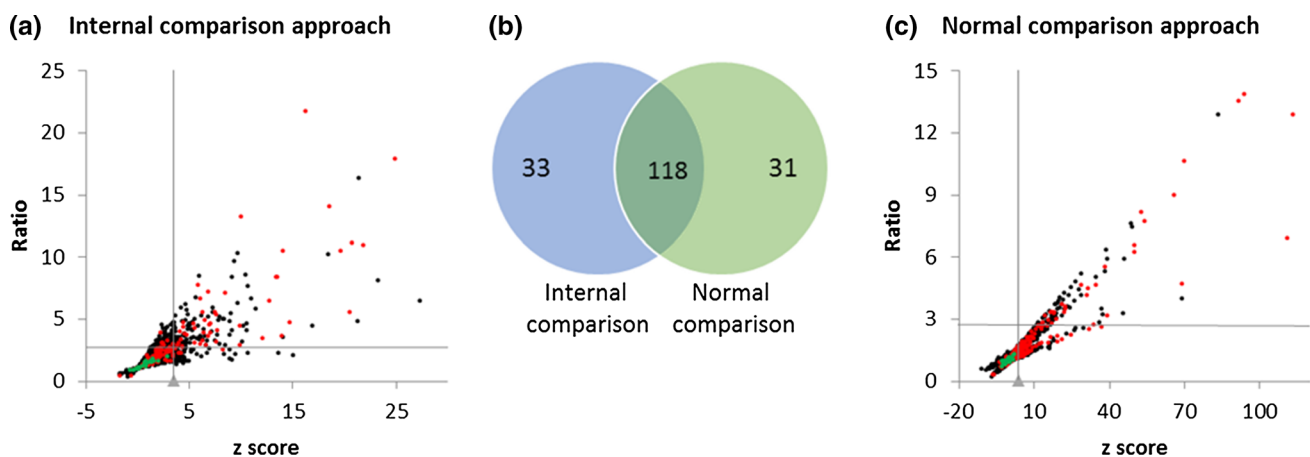
*EGFR* epidermal growth factor receptor, *MLPA* multiplex ligation-dependent probe amplification, *NGS* next generation sequencing

approaches (Fig. 2). The percentage of samples with a NGS-based amplification was 27.9% in the *EGFR* mutant group as compared with 7.4% in the *EGFR* wild-type group according to the within-tumor sample comparison approach and 25.6% vs 7.5% according to the normal sample comparison approach (Table 2).

### 3.3 Association of *EGFR* Copy Number Gain and Clinical Outcome

A total of 60 patients were treated with first-line *EGFR*-TKIs and had complete follow-up data (Table 3). The median PFS time was 8 months (95% confidence interval [CI] 5.9–10.1) and the median overall survival was 30 months (95% CI 23.6–36.4) in the total group.

Kaplan–Meier survival analysis showed no significant differences of PFS for patients with *EGFR* ratio  $\geq 2.8$ ,  $z$ -score  $\geq 3.5$ , or copy number gain as defined by the combination of both criteria for the internal comparison approach (Fig. 3a–c). In contrast, a significant shorter OS was seen for patients with an *EGFR* ratio  $\geq 2.8$ ,  $z$ -score  $\geq 3.5$ , or *EGFR* copy number gain ( $p$ -values were 0.011, 0.002, and 0.0003, respectively) (Fig. 3d–f). Patients with *EGFR* copy number gains had a median OS of 13 months, while the median OS of patients without gain was not reached. The univariate analysis showed no significant differences for PFS (Table S3 of the ESM), thus no multivariate analysis was performed for PFS. The HR for OS of *EGFR* gain by the internal comparison approach was 3.14 (95% CI 1.46–6.78,  $p = 0.003$ ) [Table S4 of the ESM].



**Fig. 2** Next generation sequencing analysis-based epidermal growth factor receptor (*EGFR*) high copy numbers using two different strategies. Ratios and  $z$ -scores of patients with non-small cell lung cancer as calculated by (a) the within-tumor sample and (c) the normal control sample approaches. For each patient, we only included the first biopsy analyzed by next generation sequencing. **b** Venn diagram showing the overlap between the two approaches. The lines in the

graphs of panels **a** and **c** represent the cut-off values used for calling high copy numbers. The two distinct subpopulations that can be seen in this graph represent samples analyzed by design 1 and design 2, the ratio values show two distinct subgroups owing to differences in normal controls used for calculation of the ratios. Green dots indicate normal samples, black dots indicate *EGFR* wild-type samples, and red dots indicate *EGFR* mutant samples

**Table 2** Next generation sequencing-based *EGFR* high copy number in patients with *EGFR* wild-type and mutant non-small cell lung cancer

| <i>EGFR</i>                         | High copy number |
|-------------------------------------|------------------|
| <i>Internal comparison approach</i> |                  |
| Mutant                              | 48/172 (27.9%)   |
| Wild type                           | 103/1394 (7.4%)  |
| <i>Normal comparison approach</i>   |                  |
| Mutant                              | 44/172 (25.6%)   |
| Wild type                           | 105/1394 (7.5%)  |

*EGFR* epidermal growth factor receptor

**Table 3** Characteristics of patients with *EGFR* mutation-positive adenocarcinoma treated with a first-line *EGFR* inhibitor

| Characteristics               | <i>n</i> = 60 (%) |
|-------------------------------|-------------------|
| <i>Age upon TKI treatment</i> |                   |
| Median                        | 66                |
| Range                         | 38–85             |
| <i>Sex</i>                    |                   |
| Male                          | 20 (33%)          |
| Female                        | 40 (67%)          |
| <i>Smoking status</i>         |                   |
| Never smoker                  | 24 (40%)          |
| Former and current smoker     | 36 (60%)          |
| <i>Mutation type</i>          |                   |
| E19DEL                        | 28 (47%)          |
| L858R                         | 17 (28%)          |
| Uncommon mutations*           | 15 (25%)          |
| <i>EGFR-TKI inhibitors</i>    |                   |
| Afatinib                      | 14 (23%)          |
| Erlotinib                     | 16 (27%)          |
| Gefitinib                     | 28 (47%)          |
| Osimertinib                   | 2 (3%)            |
| Median VAF (range)            | 51 (9–95)         |

\*Uncommon mutations: *EGFR* sensitive mutations except for E19DEL and L858R

*EGFR* epidermal growth factor receptor, *TKI* tyrosine kinase inhibitors, *VAF* variant allele frequency

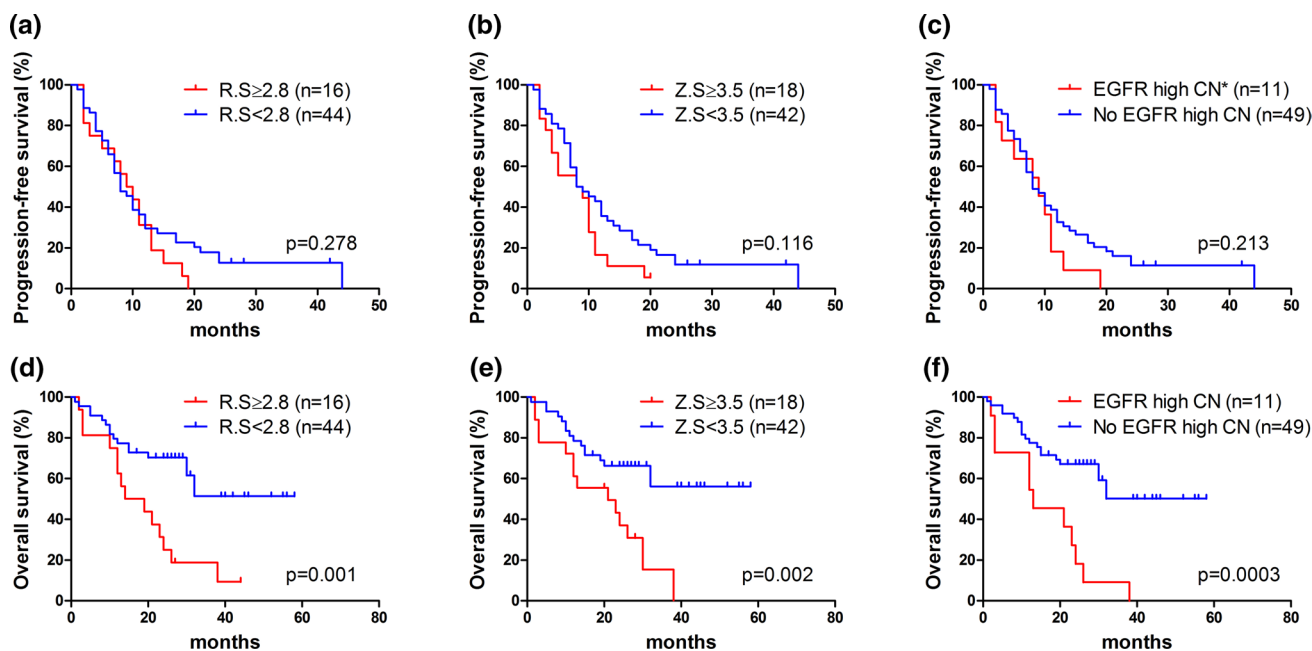
For the comparison relative to normal control samples, no significant differences were observed for PFS, similar to the within-tumor sample approach (Figs. S2a–c of the ESM). For OS, log-rank *p* values for the ratio, *z*-score, and *EGFR* copy number gain as defined by both criteria were 0.062, 0.096, and 0.062, respectively (Fig. S2d–f of the ESM). Median OS for patients with *EGFR* copy number gain as defined by the normal control approach was 23 months and 32 months for patients without gain. The HR for OS was 2.01 (95% CI 0.94–4.32, *p* = 0.074). Thus, the approach

using the internal reference amplicons showed a more significant effect on OS as compared with the relative to normal control comparison. Excluding the two patients who received first-line osimertinib treatment did not change the results on OS for both the internal and normal comparison approaches (Figs. S3 and S4 of the ESM).

We next checked second and/or subsequent lines of treatment and T790M status as potential variables for the effect on OS. An overview of the subsequent lines of treatment is shown in Table S5 of the ESM. Three patients were excluded based on receiving radiotherapy (*n* = 1), *EGFR* antibody treatment (*n* = 1), or with missing treatment information (*n* = 1). As all but one of the patients received osimertinib as a third-generation TKI, we further refer to this group as the osimertinib-treated patients. Median OS for patients who received osimertinib, first/second-generation *EGFR*-TKI, chemotherapy, or no further treatment was significantly different (*p* = 0.0003) with not reached, 30 months, 30 months, and 5 months respectively (Fig. 4a). A multivariate analysis including osimertinib and *EGFR* copy number gain for OS revealed significant and opposite HR for *EGFR* copy number gain (HR = 2.79, 95% CI 1.29–6.02, *p* = 0.009) and osimertinib treatment (HR = 0.43, 95% CI 0.20–0.91, *p* = 0.028) [Table S6 of the ESM]. Among the 29 patients who had osimertinib as second- or subsequent lines of treatment, five patients with *EGFR* copy number gain had a median OS of 13 months while the median OS for the 24 patients without *EGFR* copy number gain was not reached (Fig. 4b).

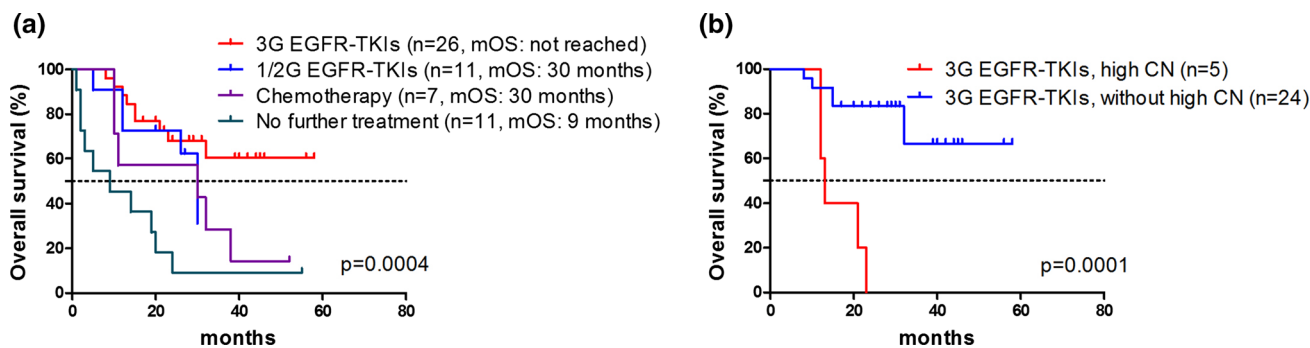
T790M mutation status in follow-up biopsies could be retrieved for 42 patients from the molecular diagnostics reports. No re-biopsy was available at progression or there was no progression during follow-up for the remaining 18 cases. T790M was identified in follow-up samples of 27 out of the 42 patients, including 16 patients with a baseline E19DEL, nine patients with a L858R, one patient with S768I, and one patient with a R776G. For the remaining 15 patients, we did not identify a T790M mutation in the biopsy retrieved at progression.

The 27 patients with a T790M mutation at progression had a longer OS compared with the 15 patients that did not develop a T790M (*p* = 0.0002) (Fig. 5a). Cox regression analysis of the 42 patients using *EGFR* gain and T790M mutation status as covariates showed an HR of 3.8 (95% CI 1.78–8.1, *p* = 0.014) and 0.24 (95% CI 0.1–0.59, *p* = 0.001), respectively, indicating opposite effects for T790M and *EGFR* gain on OS. The four patients with concurrent *EGFR* T790M and *EGFR* gains had a shorter OS than the 23 patients with T790M but without *EGFR* gain (log-rank *p* = 0.002) [Fig. S5a of the ESM]. Non-T790M mutation patients had a short OS irrespective of their *EGFR* copy number status (Fig. S5b of the ESM). As the number of patients with *EGFR* gain is rather low, we analyzed OS of *EGFR* T790M mutant and non-T790M patients based only



**Fig. 3** Kaplan–Meier plots of progression-free survival and overall survival for ratio, z-score, and the combined score (indicated as *EGFR* high copy number [CN] or no *EGFR* high CN) of epidermal growth factor receptor (*EGFR*) gain using the internal comparison approach of 60 *EGFR*-mutated patients who were treated with EGFR-TKIs. **a–c** Progression-free survival of *EGFR*-mutated patients with and without *EGFR* ratio  $\geq 2.8$ , z-score  $\geq 3.5$  as single parameters, and

for the combination. There is no significant difference in progression-free survival time between patients with and without *EGFR* gain. **d–f** Overall survival of *EGFR*-mutated patients with and without *EGFR* ratio  $\geq 2.8$ , z-score  $\geq 3.5$  as single parameters, and for the combination. Patients with a ratio  $\geq 2.8$  and/or a z-score  $\geq 3.5$  for *EGFR* had a worse overall survival



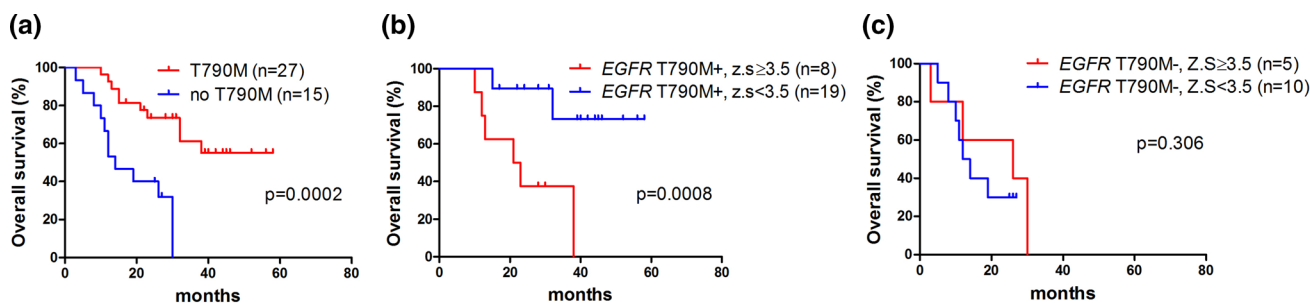
**Fig. 4** Kaplan–Meier plots of overall survival for patients stratified according to subsequent lines of treatment. **a** Overall survival of 55 patients stratified based on second-line treatment regimens. The patient for whom no second-line treatment information could be retrieved, the two patients received radiotherapy or epidermal growth factor receptor (*EGFR*) antibody treatment as second-line treatment, and the two patients who received osimertinib in the first line were

excluded. **b** Overall survival of 29 patients had third-generation *EGFR*-tyrosine kinase inhibitors as second-line or subsequent lines of treatment according to *EGFR* copy number. Three of the 29 patients received osimertinib as third- or fourth-line treatment. 1/2/3G first/second/third-generation, CN copy number, mOS median overall survival

on a z-score cut-off  $\geq 3.5$  (without using the ratio cut-off score) for calling *EGFR* copy number gain. These analyses revealed a longer OS for patients with T790M and a z-score for *EGFR* gain  $< 3.5$  ( $p=0.0008$ ), while there was no difference for OS for non-T790M patients ( $p=0.306$ ) (Fig. 5b,

c). Despite the limited number of patients, our data suggest that OS for patients with *EGFR* gain is different between T790M-positive and T790M-negative patients.

Progression-free survival was different ( $p=0.036$ ) among the mutation types (Fig. S6 of the ESM), with a median



**Fig. 5** Kaplan–Meier plots of overall survival for patients with or without an epidermal growth factor receptor (*EGFR*) T790M mutation in the follow-up biopsy. **a** Overall survival of all 42 patients. Patients who had a T790M mutation in follow-up biopsies had a longer overall survival compared with those without the T790M variant. **b** Overall survival of T790M-positive patients based on a

$z$ -score  $\geq 3.5$  or  $< 3.5$ . Patients with a  $z$ -score  $\geq 3.5$  had a shorter overall survival than patients with a  $z$ -score  $< 3.5$ . **c** Overall survival of non-T790M patients using a  $z$ -score cut-off of 3.5. There was no difference in overall survival for non-T790M patients stratified according to a  $z$ -score cut-off of 3.5

PFS of 10 months (95% CI 4.8–15.2) for E19DEL patients ( $n = 28$ ), 8 months (95% CI 5.6–10.4) for L858R patients ( $n = 17$ ), and 5 months (95% CI 1.2–8.8) for patients with uncommon activating *EGFR* mutations ( $n = 15$ ). No significant difference was observed for OS according to mutation type. Lower or higher *EGFR* VAFs according to median, upper, and lower quartiles of VAF were not associated with PFS or OS.

## 4 Discussion

In this study, we evaluated two approaches for calling *EGFR* copy number gain based on read counts of the amplicons obtained via a routinely applied, amplicon-based targeted NGS approach. The percentages of *EGFR* gain as determined by the within-tumor sample approach were 9.6%, 27.9%, and 7.4% in all patients, in patients with and without *EGFR* mutations, respectively. Moreover, a high *EGFR* copy number as estimated by the within-tumor sample approach in *EGFR*-mutated patients was associated with a significant shorter OS, but not with PFS, indicating that it is a worse prognostic factor. No association was observed for the relative to normal control samples. A longer OS was observed specifically for patients without *EGFR* gain who were treated with osimertinib as second- or subsequent lines of treatment and in patients who developed a T790M mutation at disease progression. The difference in OS between T790M- positive and T790M-negative groups should be interpreted carefully given the limited number of patients. As the presence of T790M is an indication for treatment with osimertinib, we cannot dissect the effect of both variables in a multivariate analysis.

The use of diagnostic data for copy number variation calling enables broad implementation, as targeted NGS data are available for most patients with advanced-stage

NSCLC. Several studies have analyzed the presence of *EGFR* amplifications in patients with *EGFR* mutant and/or *EGFR* wild-type NSCLC using FISH, chromogenic in situ hybridization, Southern blot, or MLPA assay [9–16]. The incidence of *EGFR* amplifications varied between 3.5% and 32% in these studies [9, 10, 13, 14, 16]. In *EGFR* wild-type cases, amplification percentages were lower, with a range of 1–29% as compared with cases with known *EGFR* driver variants (8–81%) [9–16]. Of note, these studies applied different techniques and used different cut-off values to define amplifications as there is no guideline for calling clinically relevant *EGFR* gains or amplifications. High percentages were reported in studies that were less strict for setting amplification cut-off values, e.g.,  $\geq 3$  or  $\geq 4$  *EGFR* copies/cell [12, 15]. One study with relatively strict criteria for amplification estimation showed *EGFR* amplifications in 8% of *EGFR* mutant patients. Their criterion was seven or more *EGFR* copies per cell as a cut-off value [10]. We used a ratio cut-off of 2.3 or 2.8 to define high *EGFR* copy number gain, indicating approximately five copies for diploid or near diploid tumor cells. As NSCLC is frequently highly aneuploid with three to four copies of each chromosome per cell [24, 25], our criterion of a ratio  $\geq 2.3$  or 2.8 will for most cases reflect around seven copies per tumor cell. We observed amplifications in 27.9% of the NSCLC cases with an *EGFR* mutation as compared with 7.4% in *EGFR* wild-type cases by an internal comparison approach. These frequencies are in line with previously published frequencies. All studies show a consistently higher percentage of *EGFR* copy number gain in *EGFR* mutation-positive NSCLC cases.

In our Dutch cohort, we showed that patients with both an *EGFR* mutation and copy number gain had a worse OS as compared with patients with *EGFR* mutation and without copy number gain while there was no significant effect on PFS. In an Asian cohort with gefitinib-treated *EGFR* mutant patients, *EGFR* was amplified in 48% of



the patients and the *EGFR* amplification group ( $n = 41$ ) had a better median PFS compared with the group without *EGFR* amplifications ( $n = 45$ ) [16 months vs 9.1 months] [12]. Overall survival was not reported in this study. In an erlotinib-treated Latino cohort, the median PFS of the *EGFR* amplification group ( $n = 22$ ) vs the no amplification group ( $n = 50$ ) was 28.5 months vs 11.0 months and the median OS was 37.8 months vs 27.1 months [11]. The patients in this study were included from 2013 to 2016 and included eight patients with, in addition to the activating *EGFR* mutation, a T790M mutation in the pre-treatment biopsy. Two different hypotheses can be proposed for the potential effect of *EGFR* amplifications on PFS/OS, i.e., either as favorable as the tumor cells are more dependent on *EGFR* and thus more sensitive to EGFR blockade, or alternatively as unfavorable as it might be harder to efficiently block all mutated *EGFR* receptors and tumor cells have an increased chance to develop the T790M mutation on one of the mutant *EGFR* copies. Our results that *EGFR* amplification is associated with poor OS fits with the second hypothesis in which the initial tumor response is not affected by *EGFR* copy numbers, but in the case of second-line treatment with a third-generation TKI, that targets *EGFR*-activating mutations and T790M is less effective, owing to amplification of *EGFR*, which is a known resistance mechanism toward osimertinib. In the diagnostic setting, we mostly see a lower T790M variant allele frequency as compared with the variant allele frequency of the driver mutations [26], which further supports that having a single copy of the resistant EGFR (because of T790M) is sufficient to induce a relapse. The T790M-positive relapse patients subsequently show more indolent progression [27] and still a good response towards third-generation TKIs, indicating that the tumor cells are dependent on EGFR signaling.

Potential differences with the above-mentioned studies might be related to differences in techniques as well as differences in ethnicity. We used amplicon-based targeted NGS data to detect *EGFR* copy number gain, while the previous studies used FISH. Our approach gives an estimation of the *EGFR* copy number level based on DNA extracted from a tumor cell-rich area of the tissue sample, aiming at a minimal tumor cell content of at least 20%, while FISH results are based on counting copies in a limited number of tumor cells [11, 12, 28]. As marked differences have been shown for minor allele frequencies for the intronic CA repeat between different ethnicities [29], and *EGFR* polymorphisms have been linked to survival on EGFR-TKIs [30], this might at least in part explain differences in survival between our study and previous studies with respect to the clinical effect of *EGFR* amplifications.

To explore the potential underlying reason why patients with *EGFR* gain had a worse OS in our cohort, we

determined the potential influence of second- or subsequent lines of treatment and T790M as a resistance mechanism. This revealed a longer median OS for patients who received osimertinib as second or subsequent lines of treatment and/or developed a T790M as compared to patients who did not develop a T790M at disease progression. Patients with a T790M as a resistance mechanism were mostly successfully treated with a third-generation TKI at progressive disease. Combined analysis of these variables indicated that the favorable OS was restricted to the group of patients without *EGFR* copy number gains who developed a T790M mutation and were treated with osimertinib. In our study, lack of *EGFR* copy number gains at baseline is a good prognostic factor for OS, which fits with the observation that *EGFR* amplification has been reported as a resistance mechanism to osimertinib or other third-generation EGFR-TKIs [31–33].

In addition to studies focusing on *EGFR* gain in relation to survival, some other studies determined the relation between the VAF of the *EGFR* mutation and survival. A shorter PFS was shown for cases with lower E19DEL and L858R VAF in the study of Li et al. [8], whereas this was observed only for L858R in a study of Hung et al. [6]. Ono et al. [7] observed a shorter PFS for patients with a low L858R VAF but did not observe a difference in PFS for E19DEL VAF. In our study, VAF of both E19DEL (14–95%) and L858R (21–86%) were above the cut-off values used in the above-described papers (i.e., 4.9–9.9%) in all patients, while estimated tumor cell percentages were not higher in our study. Based on results as reported in the literature and our own data showing no effect of VAF on survival, we conclude that VAF is not a reliable predictor of survival in EGFR-TKI-treated patients.

Our targeted NGS panels included only a limited number of amplicons at chromosome 7. Therefore, we were limited in differentiating a high copy number because of polysomy or focal amplification. Moreover, the small NGS panel (Table S1 of the ESM) precluded a more in-depth analysis of other concurrent genomic alterations that may affect survival [34, 35]. Pseudo-amplification due to decreased copy numbers of reference amplicons by internal comparison approach can occur, but the impact will be low, as we selected the most stable amplicons and used the median value of these amplicons for the calculations. Read count normalization using the within-tumor sample approach had a better correlation with survival than the approach using normal control samples in this study. The differences between the two methods might be related to the high degree of aneuploidy in the tumor samples, with up to three or four copies per chromosome and related to the batch-wise analysis of normal control samples, whereas tumor samples were retrieved over a period of 3 years. The number of *EGFR* mutation-positive patients

with available clinical data for first-line TKI treatment was limited and further studies are needed.

## 5 Conclusions

Amplicon-based targeted NGS data can be used to estimate the presence of *EGFR* gain. The *EGFR* copy number status as defined by the internal amplicon comparison can be used as a prognostic marker for OS in *EGFR* mutation-positive patients treated with EGFR-TKIs. The observed clinical value of *EGFR* gain was observed in cases that were treated with osimertinib in second or subsequent lines of treatment based on the presence of T790M upon resistance to first-line TKIs. Our findings warrant routine testing of *EGFR* copy number gains in the clinical setting, especially because it might be associated with a lower tumor response to second-line EGFR-TKIs with osimertinib. Based on our results, we hypothesize that alternative combination therapies may improve outcome for patients with *EGFR* gain and mutation.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11523-021-00798-2>.

**Acknowledgments** We acknowledge the UMCG molecular diagnostic team in the Pathology Department for technical assistance with the experimental work. We thank the UG Center for Information Technology and their sponsors BBMRI-NL & TarGet for storage and computer infrastructure. We thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at <http://exac.broadinstitute.org/about>.

## Declarations

**Funding** This work was funded by a KWF grant (RUG2015-8044) and the University Medical Centre Groningen.

**Conflicts of Interest/Competing Interests** Harry J.M. Groen reports grants from Boehringer-Ingelheim, Takeda, BMS, Novartis, and Merck outside the submitted work. Anthonie J. van der Wekken has received research grants from AstraZeneca, Pfizer, Boehringer-Ingelheim, Roche, and Takeda. Jeroen Hiltermann reports grants from AstraZeneca, Pfizer, Boehringer-Ingelheim, Roche, BMS, and MSD, outside the submitted work. Jiacong Wei, Pei Meng, Miente Martijn Terpstra, Anke van Rijk, Menno Tamminga, Frank Scherpen, Arja ter Elst, Mohamed Z. Alimohamed, Lennart F. Johansson, Jos Stigt, Rolof P.G. Gijtenbeek, John van Putten, Klaas Kok, and Anke van den Berg have no conflicts of interest that are directly relevant to the content of this article.

**Ethics Approval** The study protocol is consistent with the Research Code of the University Medical Centre Groningen (<https://www.rug.nl/umcg/research/documents/research-code-info-umcg-nl.pdf>) and national ethical and professional guidelines (“Code of conduct; Dutch Federation of Biomedical Scientific Societies”, <http://www.feder.a.org>).

**Consent to Participate** Not applicable.

**Consent for Publication** Not applicable.

**Availability of Data and Material** Not applicable.

**Code Availability** Not applicable.

**Authors' Contributions** All authors contributed to the study conception and design, material preparation, data collection, and analysis. The first draft of the manuscript was written by Jiacong Wei and Pei Meng. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## References

- Hirsch FR, Suda K, Wiens J, Bunn PA Jr. New and emerging targeted treatments in advanced non-small-cell lung cancer. *Lancet*. 2016;388(10048):1012–24.
- Ladanyi M, Pao W. Lung adenocarcinoma: guiding EGFR-targeted therapy and beyond. *Mod Pathol*. 2008;21:S16.
- Tu HY, Ke EE, Yang JJ, Sun YL, Yan HH, Zheng MY, et al. A comprehensive review of uncommon EGFR mutations in patients with non-small cell lung cancer. *Lung Cancer*. 2017;14:96–102.
- Ettinger DS, Wood DE, Aggarwal C, Aisner DL, Akerley W, Bauman JR, et al. NCCN guidelines insights: non-small cell lung cancer, version 1.2020. *J Natl Compr Canc Netw*. 2019;7(12):1464–72.
- Zhong J, Li L, Wang Z, Bai H, Gai F, Duan J, et al. Potential resistance mechanisms revealed by targeted sequencing from lung adenocarcinoma patients with primary resistance to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs). *J Thorac Oncol*. 2017;12(12):1766–78.
- Hung MS, Lung JH, Lin YC, Fang YH, Hsieh MJ, Tsai YH. The content of mutant EGFR DNA correlates with response to EGFR-TKIs in lung adenocarcinoma patients with common EGFR mutations. *Medicine*. 2016;95(26):e3991.
- Ono A, Kenmotsu H, Watanabe M, Serizawa M, Mori K, Imai H, et al. Mutant allele frequency predicts the efficacy of EGFR-TKIs in lung adenocarcinoma harboring the L858R mutation. *Ann Oncol*. 2014;5(10):1948–53.
- Li X, Cai W, Yang G, Su C, Ren S, Zhao C, et al. Comprehensive analysis of EGFR-mutant abundance and its effect on efficacy of EGFR TKIs in advanced NSCLC with EGFR mutations. *J Thorac Oncol*. 2017;12(9):1388–97.
- Li AR, Chitale D, Riely GJ, Pao W, Miller VA, Zakowski MF, et al. EGFR mutations in lung adenocarcinomas: clinical testing experience and relationship to EGFR gene copy

- number and immunohistochemical expression. *J Mol Diagn.* 2008;10(3):242–8.
10. Yokoyama T, Kondo M, Goto Y, Fukui T, Yoshioka H, Yokoi K, et al. EGFR point mutation in non-small cell lung cancer is occasionally accompanied by a second mutation or amplification. *Cancer Sci.* 2006;97(8):753–9.
  11. Ruiz-Patiño A, Castro CD, Ricaurte LM, Cardona AF, Rojas L, Zatarain-Barrón ZL, et al. EGFR amplification and sensitizing mutations correlate with survival in lung adenocarcinoma patients treated with erlotinib (MutP-CLICaP). *Targ Oncol.* 2018;13(5):621–9.
  12. Shan L, Wang Z, Guo L, Sun H, Qiu T, Ling Y, et al. Concurrence of EGFR amplification and sensitizing mutations indicate a better survival benefit from EGFR-TKI therapy in lung adenocarcinoma patients. *Lung Cancer.* 2015;89(3):337–42.
  13. Morinaga R, Okamoto I, Fujita Y, Arai T, Sekijima M, Nishio K, et al. Association of epidermal growth factor receptor (EGFR) gene mutations with EGFR amplification in advanced non-small cell lung cancer. *Cancer Sci.* 2008;99(12):2455–60.
  14. Lewandowska MA, Czuba K, Klonowska K, Jozwicki W, Kowalewski J, Kozłowski P. The use of a two-tiered testing strategy for the simultaneous detection of small EGFR mutations and EGFR amplification in lung cancer. *PLoS One.* 2015;10(2):e0117983.
  15. Varella-Garcia M, Mitsudomi T, Yatabe Y, Kosaka T, Nakajima E, Xavier AC, et al. EGFR and HER2 genomic gain in recurrent non-small cell lung cancer after surgery: impact on outcome to treatment with gefitinib and association with EGFR and KRAS mutations in a Japanese cohort. *J Thorac Oncol.* 2009;4(3):318–25.
  16. Miller VA, Riely GJ, Zakowski MF, Li AR, Patel JD, Heelan RT, et al. Molecular characteristics of bronchioloalveolar carcinoma and adenocarcinoma, bronchioloalveolar carcinoma subtype, predict response to erlotinib. *J Clin Oncol.* 2008;26(9):1472–8.
  17. Johansson LF, van Dijk F, de Boer EN, van Dijk-Bos KK, Jongbloed JDH, van der Hout AH, et al. CoNVaDING: single exon variation detection in targeted NGS data. *Hum Mutat.* 2016;37(5):457–64.
  18. Eijkelenboom A, Tops BBJ, van den Berg A, van den Brule AJC, Dinjens WNM, Dubbink HJ, et al. Recommendations for the clinical interpretation and reporting of copy number gains using gene panel NGS analysis in routine diagnostics. *Virchows Arch.* 2019;474(6):673–80.
  19. Crosby T. How to detect and handle outliers. *Technometrics.* 1994;36(3):315–6.
  20. Hoogstraat M, Hinrichs JWJ, Besselink NJM, Radersma-van Loon JH, de Voijs CMA, Peeters T, et al. Simultaneous detection of clinically relevant mutations and amplifications for routine cancer pathology. *J Mol Diagn.* 2015;7(1):10–8.
  21. Iglewicz B, Hoaglin D. Volume 16: how to detect and handle outliers. The ASQC basic references in quality control: statistical techniques 16. Milwaukee: ASQ Quality Press; 1993.
  22. Grasso C, Butler T, Rhodes K, Quist M, Neff TL, Moore S, et al. Assessing copy number alterations in targeted, amplicon-based next-generation sequencing data. *J Mol Diagn.* 2015;7(1):53–63.
  23. Jeuken J, Cornelissen S, Boots-Sprenger S, Gijzen S, Wesseling P. Multiplex ligation-dependent probe amplification: a diagnostic tool for simultaneous identification of different genetic markers in glial tumors. *J Mol Diagn.* 2006;8(4):433–43.
  24. Testa JR, Siegfried JM. Chromosome abnormalities in human non-small cell lung cancer. *Cancer Res.* 1992;52(9 Suppl.):2702s–6s.
  25. Choma D, Daurès JJP, Quantin X, Pujol JL. Aneuploidy and prognosis of non-small-cell lung cancer: a meta-analysis of published data. *Br J Cancer.* 2001;85(1):14–22.
  26. Kuo CS, Huang CH, Liu CY, Pavlidis S, Ko HW, Chung FT, et al. Prior EGFR-TKI treatment in EGFR-mutated NSCLC affects the allelic frequency fraction of acquired T790M and the subsequent efficacy of osimertinib. *Targ Oncol.* 2019;14(4):433–40.
  27. Oxnard GR, Arcila ME, Sima CS, Riely GJ, Chmielecki J, Kris MG, et al. Acquired resistance to EGFR tyrosine kinase inhibitors in EGFR-mutant lung cancer: distinct natural history of patients with tumors harboring the T790M mutation. *Clin Cancer Res.* 2011;7(6):1616–22.
  28. Cappuzzo F, Finocchiaro G, Grossi F, Bidoli P, Favaretto A, Marchetti A, et al. Phase II study of afatinib, an irreversible ErbB family blocker. EGFR FISH-positive non-small-cell lung cancer. *J Thorac Oncol.* 2015;10(4):665–72.
  29. Liu W, Innocenti F, Chen P, Das S, Cook EH Jr, Ratain MJ. Interethnic difference in the allelic distribution of human epidermal growth factor receptor intron 1 polymorphism. *Clin Cancer Res.* 2003;9(3):1009–12.
  30. Winther-Larsen A, Fynboe Ebert EB, Meldgaard P, Sorensen BS. EGFR Gene Polymorphism predicts improved outcome in patients with EGFR mutation-positive non-small cell lung cancer treated with erlotinib. *Clin Lung Cancer.* 2019;20(3):161–166.
  31. 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(9):725–37.
  32. Knebel FH, Bettoni F, Shimada AK, Cruz M, Alessi JV, Negrão MV, et al. Sequential liquid biopsies reveal dynamic alterations of EGFR driver mutations and indicate EGFR amplification as a new mechanism of resistance to osimertinib in NSCLC. *Lung Cancer.* 2017;108:238–41.
  33. Zhang YC, Chen ZH, Zhang XC, Xu CR, Yan HH, Xie Z, et al. Analysis of resistance mechanisms to abivertinib, a third-generation EGFR tyrosine kinase inhibitor, in patients with EGFR T790M-positive non-small cell lung cancer from a phase I trial. *EBioMedicine.* 2019;43:180–7.
  34. Li S, Li L, Zhu Y, Huang C, Qin Y, Liu H, et al. Coexistence of EGFR with KRAS, or BRAF, or PIK3CA somatic mutations in lung cancer: a comprehensive mutation profiling from 5125 Chinese cohorts. *Br J Cancer.* 2014;110(11):2812–20.
  35. De Marchi F, Haley L, Fryer H, Ibrahim J, Beierl K, Zheng G, et al. Clinical validation of coexisting activating mutations within EGFR, mitogen-activated protein kinase, and phosphatidylinositol 3-kinase pathways in lung cancers. *Arch Pathol Lab Med.* 2019;143(2):174–82.

## Authors and Affiliations

Jiacong Wei<sup>1,4</sup> · Pei Meng<sup>2,5</sup> · Miente Martijn Terpstra<sup>1</sup> · Anke van Rijk<sup>2</sup> · Menno Tamminga<sup>3</sup> · Frank Scherpen<sup>2</sup> · Arja ter Elst<sup>2</sup> · Mohamed Z. Alimohamed<sup>1,9</sup> · Lennart F. Johansson<sup>1</sup> · Jos Stigt<sup>6</sup> · Rolof P. G. Gijtenbeek<sup>7</sup> · John van Putten<sup>8</sup> · T. Jeroen N. Hiltermann<sup>3</sup> · Harry J. M. Groen<sup>3</sup> · Klaas Kok<sup>1</sup> · Anthonie J. van der Wekken<sup>3</sup> · Anke van den Berg<sup>2</sup>

<sup>1</sup> Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

<sup>2</sup> Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, HPC: EA10, Room F0-15, Hanzeplein 1, 9713 GZ Groningen, The Netherlands

<sup>3</sup> Department of Pulmonary Diseases, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

<sup>4</sup> Department of Pathology, Cancer Hospital Chinese Academy of Medical Sciences, Beijing, China

<sup>5</sup> Department of Pathology, Collaborative and Creative Centre, Shantou University Medical College, Shantou, Guangdong, China

<sup>6</sup> Department of Pulmonary Diseases, Isala Clinic, Zwolle, The Netherlands

<sup>7</sup> Department of Pulmonary Diseases, Medical Center Leeuwarden, Leeuwarden, The Netherlands

<sup>8</sup> Department of Pulmonary Diseases, Martini Hospital, Groningen, The Netherlands

<sup>9</sup> Department of Haematology and Blood Transfusion, Muhimbili University of Health and Allied Sciences, Dar-es-Salaam, Tanzania