The frequency of EGFR and KRAS mutations in non-small cell lung cancer (NSCLC): routine screening data for central Europe from a cohort study

Christian Boch,1 Jens Kollmeier,1 Andreas Roth,2 Susann Stephan-Falkenau,2 Daniel Misch,1 Wolfram Grüning,3 Torsten Thomas Bauer,1 Thomas Mairinger2

ABSTRACT

Objectives: Owing to novel therapy strategies in epidermal growth factor receptor (EGFR)-mutated patients, molecular analysis of the EGFR and KRAS genome has become crucial for routine diagnostics. Till date these data have been derived mostly from clinical trials, and thus collected in pre-selected populations. We therefore screened ‘allcomers’ with a newly diagnosed non-small cell lung carcinoma (NSCLC) for the frequencies of these mutations.

Design: A cohort study.

Setting: Lung cancer centre in a tertiary care hospital.

Participants: Within 15 months, a total of 552 cases with NSCLC were eligible for analysis.

Primary and secondary outcome measures: Frequency of scrutinising exons 18, 19 and 21 for EGFR and KRAS mutations.

Results: Of the 552 patients, 27 (4.9%) showed a mutation of EGFR. 19 of these patients (70%) had deletion E746-A750 in codon 19 or deletion L858R in codon 21. Adenocarcinoma (ACA) was the most frequent histology among patients with EGFR mutations (ACA, 22/254 (8.7%) vs non-ACA, 5/298 (1.7%); p<0.001). Regarding only ACA, the percentage of EGFR mutations was higher in men (16/116 (14%) women vs 6/138 (4.3%) men; p=0.008). Tumours with an activating EGFR mutation were more likely to be from non-smokers (18/27; 67%) rather than smoker (9/27; 33%).

KRAS mutation was present in 85 (15%) of all cases. In 73 patients (86%), the mutation was found in exon 12 and in 12 cases (14%) in exon 13. Similarly, ACA had a higher frequency of KRAS mutations than non-ACA (67/254 (26%) vs 18/298 (6.0%); p<0.001).

Conclusions: We found a lower frequency for EGFR and KRAS mutations in an unselected Caucasian patient cohort as previously published. Taking our results into account, clinical trials may overestimate the mutation frequency for EGFR and KRAS in NSCLC due to important selection biases.

INTRODUCTION

Lung cancer is still one of the most common cancers in Germany and worldwide, and the most common cause of death due to malignant tumours. Despite intensive research, the prognosis of patients with non-small cell lung carcinoma (NSCLC) continues to be restricted, which applies especially to patients with distant metastases. However, in patients with NSCLC and an activating mutation of the epidermal growth factor receptor (EGFR), therapy has undergone a significant change. The application of first-generation tyrosine-kinase inhibitors (TKI) Gefitinib an Erlotinib was able to achieve not only a better tolerability of the therapy and an improvement in progression-free survival in observational studies, but also compared with platinum-based chemotherapy. However, the effect on overall survival (OS) is still controversial; one study could not detect a benefit, whereas Zhou and co-workers reported a significant increase in OS.
In the last few years, intensive research has been conducted to analyse the effect of TKIs in respect of the complex relationship of biomarkers which are associated with the EGFR pathway like \(v\)-Ki-ras2 Kirsten rat sarcoma viral oncogene homologue (KRAS). In addition to the superiority in EGFR-mutated patients, TKIs seem to have an effect on the wild-type NSCLC as well, while some studies have disputed this for KRAS mutated patients.\(^2\)\(^8\) However, the significance of KRAS mutations for the diagnosis or treatment of NSCLC remains under debate. Since double mutations (EGFR and KRAS) have been described in only very few cases, KRAS may help exclude EGFR mutation.\(^9\)\(^10\)

Therefore, the question arises as to how many patients with NSCLC can benefit from the new treatment option and whether screening should be implemented in the routine diagnostics. In addition, numerous studies came from Southeast Asia, where patients with NSCLC have higher EGFR mutation rates.\(^11\)\(^12\) Since the data available today are hampered either by biases owing to study exclusion criteria or different geographical peculiarities, we established routine screening for consecutive patients to receive reliable data for the routine setting in Central Europe.

### PATIENTS AND METHODS

#### Study population

From October 2009 until December 2010 we have screened prospectively all newly diagnosed patients with NSCLC (\(n=753\)) for the ability to be analysed for EGFR and KRAS mutations. All subsequent biopsies have been obtained by surgery, routine bronchoscopy or CT-guided biopsy. Staging of the tumour war performed according to the Union for International Cancer Control (UICC, Seventh Edition).\(^13\) The complete staging was available for 753/753 patients (97.6%). The histological diagnosis was performed according to the WHO criteria.\(^14\) A total of 552/753 (75%) cases with NSCLC were eligible for analysis. In 183 patients (25%) none of the two analyses could be performed. In the majority of cases, this was explained by an insufficient amount of tissue sample. In addition, some patients refused their consent to carry out the mutation analysis. Informed consent was obtained from all patients prior to biopsy.

### DNA extraction

After preparation of a 20 µm thick section from the formalin-fixed and paraffin-embedded tissue sample, the area of tumour cells was marked by comparison with the corresponding HE-stained slide. This area with a size of 2–10 mm\(^2\) was scraped with a pipette tip and transferred into 180 µl AL-buffer. The AL-buffer is included in the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). Later the DNA was extracted and dissolved in 150 µl AE-buffer according to the manufacturer’s instructions using the QIAamp DNA Mini Kit.

#### Table 1

<table>
<thead>
<tr>
<th>Gender</th>
<th>Histology</th>
<th>Total</th>
<th>Female</th>
<th>Male</th>
<th>Adeno</th>
<th>Squamos</th>
<th>Large cell</th>
<th>Adenosquamous</th>
<th>NOS</th>
<th>EGFR positive</th>
<th>EGFR negative</th>
<th>KRAS positive</th>
<th>KRAS negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>552</td>
<td>467</td>
<td>85</td>
<td>254</td>
<td>79</td>
<td>21</td>
<td>5</td>
<td>19</td>
<td>(15.4%)</td>
<td>(73.6%)</td>
<td>(13.9%)</td>
<td>(97.5%)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>200</td>
<td>158</td>
<td>42</td>
<td>158</td>
<td>79</td>
<td>21</td>
<td>5</td>
<td>19</td>
<td>(4.9%)</td>
<td>(97.2%)</td>
<td>(21.0%)</td>
<td>(79.0%)</td>
</tr>
<tr>
<td></td>
<td>Adeno</td>
<td>352</td>
<td>309</td>
<td>43</td>
<td>309</td>
<td>79</td>
<td>21</td>
<td>5</td>
<td>19</td>
<td>(63.8%)</td>
<td>(91.1%)</td>
<td>(21.0%)</td>
<td>(79.0%)</td>
</tr>
<tr>
<td></td>
<td>Squamos</td>
<td>254</td>
<td>232</td>
<td>22</td>
<td>232</td>
<td>79</td>
<td>21</td>
<td>5</td>
<td>19</td>
<td>(46.0%)</td>
<td>(91.3%)</td>
<td>(42.2%)</td>
<td>(57.8%)</td>
</tr>
<tr>
<td></td>
<td>Large cell</td>
<td>186</td>
<td>181</td>
<td>5</td>
<td>181</td>
<td>79</td>
<td>21</td>
<td>5</td>
<td>19</td>
<td>(33.7%)</td>
<td>(97.3%)</td>
<td>(2.7%)</td>
<td>(97.3%)</td>
</tr>
<tr>
<td></td>
<td>Adenosquamous</td>
<td>27</td>
<td>22</td>
<td>5</td>
<td>22</td>
<td>79</td>
<td>21</td>
<td>5</td>
<td>19</td>
<td>(4.9%)</td>
<td>(91.1%)</td>
<td>(13.9%)</td>
<td>(86.1%)</td>
</tr>
<tr>
<td></td>
<td>NOS</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>(1.3%)</td>
<td>(20%)</td>
<td>(9.5%)</td>
<td>(100%)</td>
</tr>
</tbody>
</table>

No double mutations were found in this cohort. Bold is used to highlight the most important findings with regard to the values published in the literature. EGFR, epidermal growth factor receptor; NOS, non other specified.

Determination of EGFR mutations

For analysing the EGFR mutation status, the extracted DNA was subjected to conventional PCR for amplification of relevant portions of exons 18, 19 and 21 of the EGFR gene. Further, the PCR products were analysed for the presence of a mutation by using the Pyrosequencing PyroMark Instrument (Qiagen, Hilden, Germany). The primers used for the PCR are listed in table 1. The method of pyrosequencing has been performed according to the reagent manufacturer (Qiagen, Hilden, Germany) and the following primer names and sequences (3’–5’)
have been used for exons 18, 19 and 21:

- 18_F (forward) CAGCATGGTAGGGCTGAG
- 18_R (reverse) GCCTGTGCCAGGGACCTTAC
- 19_F (forward) GTGGCACCATCTCACAATTGC
- 19_R (reverse) CACACAGCAAAGCAGAAACTCAC
- 21_F (forward) CCTCACAGCAGGGTCTTCTCTG
- 21_R (reverse) CTGGCTGACCTAAAGCCACC

Samples with a mutation in the pyrogram were additionally examined by dideoxy sequencing according to Sanger et al for confirmation (figure 1).

Determination of KRAS mutations

The DNA was analysed according to manufacturer’s instructions using the K-RAS codon 12/13 kit from Molbiol TIB (Berlin, Germany) and the LightCycler 2.0 instrument (Roche Diagnostics, Mannheim, Germany).

This is a real-time PCR with FRET probes, which facilitate the differentiation of mutant and wild-type DNA in codons 12 and 13 on exon 1 of KRAS gene after successful amplification and analyses of the melting curve of the PCR products.

Statistical analyses

Quantitative variables were summarised as means±SD. For comparisons Student t test was used and in cases with more than two groups, a one-way analysis of variance (ANOVA) with Bonferroni correction was calculated. Categorical variables are reported as frequencies and for comparisons of categorical variables between groups χ² tests were used. The level of significance for all tests was set to 5% and all tests were two-sided. Analyses were carried out with the Statistical Package for Social Sciences (SPSS, V.13.0) on a Windows operating system.

RESULTS

Patients

The mean age of patients with wild-type EGFR (n=440; 66.7±9.5 years) was comparable with the mean age of patients with EGFR mutation (n=27; 70.3±11.4 years) and that of patients with a KRAS mutation in the overall comparison (n=85; 65.3±9.8 years; p=0.059). Table 1 summarises frequencies according to mutation, gender and histology.

A breakdown of patients according to the genealogy was not performed because all patients with a migration background except two came from other European countries. Patients from Asia were not diagnosed or treated in our clinic during the corresponding period.

EGFR mutation

Overall 27 (4.9%) of all 552 patients showed an activating EGFR mutation (3/27 mutations in exon 18 (11%), 15/27 mutations in exon 19 (56%) and 8/27 mutations (30%) in exon 21; 1/27 tumour (3%) has shown a double mutation (exons 19 and 21)). The most common mutation in exon 19 was delE746-A750 with 10/15 (67%) cases. In exon 21, all detected eight mutations were located on L848R (100%).

Considering the frequency of mutations according to histology, adenocarcinomas (ACAs) show mutations more often than any other histology (22/254 (8.7%) vs 5/298 (1.7%); p<0.001).

When all histologies were analysed together, women also showed a significantly higher frequency of EGFR mutation compared with men (17/200 (8.5%) vs 10/352 (2.8%); p=0.003; table 2). In contrast, among men with NSCLC no significant difference in the rate of mutation could be detected with regard to histology (ACA, 6/138 (4.3%) vs non-ACA, 4/214 (1.9%); p=0.172; table 2).

Smoking status was closely associated with EGFR mutation rate. Among patients with EGFR mutations (n=27) the rate of never-smokers was twice as high compared with smokers (18/27 (67%) vs 9/27 (33%); table 3).

KRAS mutation

A mutation in the KRAS gene was seen in a total of 85/552 (15%) of all analysed tumours. While 73/85 (86%) of these mutations were found in exon 12, in exon 13...
only 12/85 (14%) mutations were localised. With 29/85 (34%) cases the 12 Cys mutation was the most common. Similar to the EGFR the subgroups of ACAs had significantly (p<0.001) higher rate of KRAS mutations compared with other histologies (67/254 (26%) vs 18/298 (6%); table 1). Within the subgroup of ACAs there was an even gender distribution regarding the frequency of KRAS mutations (women, 31/116 (27%) vs men, 36/138 (26%); p=0.909; table 2 upper and lower panels).

However, regardless of histology the rate of KRAS mutations was significantly (p=0.007) more common in women than in men (42/200 (21%) vs 43/352 (12%); table 1). No patient had a double mutation (KRAS and EGFR).

Unlike in patients with an activating EGFR mutation the proportion of smokers is higher in patients with a KRAS mutation (75/85 (88%) vs 10/85 (22%); table 3).

**Intention-to-treat analyses**

Table 4 concludes the distribution according to histology and gender for all patients not included in this analysis (n=180). The percentage of tumours that had to be left unclassified (NOS; non other specified) was substantially higher (38/180, 21%) compared with the patient group included (21/552; 3.8%; p<0.001). This was most likely owing to an insufficient probe size, which hampered both plain histology and mutation analysis. The percentage of patients with ACAs excluded was slightly lower (39% vs 46%), however, this difference was not statistically significant (p=0.112; tables 1 and 4).

When we analysed the distribution according to the clinical stage, we found a bias with regard to higher clinical stages in the non-tested group (table 5). However, when stages IIIB and IV were compared with the potentially operable stages (IA–IIIA) we found this difference not to be statistically significant (not tested: 102/180; 56.7% vs tested: 277/552; 50.2%, p=0.077).

**DISCUSSION**

The present study delivers sound data on the frequency of EGFR and KRAS mutations in a large and unselected cohort of patients with newly diagnosed non-small cell lung cancer (NSCLC) over a well-defined study period. The EGFR mutation rate reported here (4.9%) differs largely from frequencies reported in previous studies (8–60%). This suggests that previous study designs had import selection biases concerning this mutation analyses. In the most extensive study on the prevalence of activating EGFR mutations in NSCLC at 126 Spanish hospitals during the years 2005–2008, a total of 2105

---

**Table 2** Proportion of EGFR and KRAS mutations found according to women (upper table) and men (lower table)

<table>
<thead>
<tr>
<th>Total number</th>
<th>EGFR positive (%)</th>
<th>KRAS positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total women</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>116</td>
<td>17</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>47</td>
<td>0</td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>Sarcomatoid carcinoma</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Adenosquamous carcinoma</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NOS (non other specified)</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Total men</td>
<td>352</td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>138</td>
<td>10</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>139</td>
<td>2</td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>51</td>
<td>1</td>
</tr>
<tr>
<td>Sarcomatoides Karzinom</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Adenosquamous carcinoma</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>NOS (non other specified)</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 3** Smoker status and gender distribution of all 112 patients with either EGFR or KRAS

<table>
<thead>
<tr>
<th></th>
<th>Smokers</th>
<th>Never-smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number</td>
<td>84</td>
<td>28</td>
</tr>
<tr>
<td>EGFR positive</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>EGFR positive female</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>EGFR positive male</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>KRAS positive</td>
<td>75</td>
<td>10</td>
</tr>
<tr>
<td>KRAS positive female</td>
<td>38</td>
<td>4</td>
</tr>
<tr>
<td>KRAS positive male</td>
<td>37</td>
<td>6</td>
</tr>
</tbody>
</table>

Bold is used to highlight the most important findings with regard to the values published in the literature.
patients were analysed and a mutation rate of 16.6% was reported.\textsuperscript{17} One possible explanation for the higher mutation rate in this study lies in the unrepresentative distribution of the histological subgroups. While the expected proportion of ACAs in routine studies is between 32% and 54% depending on the literature, this study reported up to 78% ACAs. Since ACAs were also in our study more likely to exhibit an activating EGFR mutation, the frequency was overestimated in the Spanish study.\textsuperscript{17} The disproportion of ACAs is also found in other studies reporting EGFR mutation frequencies between 9% and 31\%.\textsuperscript{18,19} In our analysis, however, the proportion of ACAs was well within the expected range (46%) with no selection bias induced by patient selection.

Another factor with great influence on the EGFR mutation rate is the geographical origin of the analysed patients. For quite some years it has been known that East Asian patients with NSCLC have a higher incidence of smokers.\textsuperscript{20} 21\textsuperscript{21} Studies with a high proportion of KRAS mutations compared with other ethnic groups at the same time.\textsuperscript{12,20,21} Studies with a high proportion of patients from Far Eastern countries or studies which have recruited their patients entirely in the East Asian region have consequently shown a corresponding effect on the frequency of the detected EGFR and KRAS mutations.\textsuperscript{6,22,23} None of our patients with a migration background came from an East Asian country and thus our figures are not prone to this bias.

A study from Basel seems to be the most comparable in terms of patient’s cohort. Pathologists examined 307 NSCLC during a 4.5-year period for EGFR mutation and showed a mutation rate of 8.1%. However it remains unclear in this study, which preliminary selection bias may have been introduced by the clinicians, because mutation analyses were not requested for all samples. Moreover, it has not been reported how many patients with NSCLC were newly diagnosed in the participating hospitals.\textsuperscript{16}

As demonstrated in our study, patients with NSCLC and EGFR mutation are non-smokers in the majority of cases, which has been confirmed by numerous studies.\textsuperscript{5,24,25} In Europe and in Germany as well it is estimated that 85% of lung cancers-related deaths are associated with cigarette smoking.\textsuperscript{26} Such a high proportion of smokers among patients with lung cancer make a high EGFR mutation rate questionable. However, the question arises why this mutation occurs much more frequently in the Asian populations.

Independent of the rate of mutation in total, the frequency of EGFR mutations in exons 19 and 21 of our samples is comparable with those of numerous studies.\textsuperscript{22,27–29} In addition, the occurrence of the most common mutation in our evaluation delE746-A750 in exon 19 (10/27 (37%)) corresponds to the literature.\textsuperscript{30} Another interesting aspect is that the restriction on the search for the two most common mutations in exons 19 (delE746-A750) and 21 (L858R), as discussed in the literature, would have detected only 18/27 (69.2%) of all EGFR mutations in our study.

Regarding the KRAS mutations our analysis confirms the known fact that KRAS mutations occur in smokers with NSCLC significantly more often than in non-smokers.\textsuperscript{31,32} This is similar to the distribution of mutations of KRAS at codon 12 (72/85 (85%)) and 13 (13/85 (15.3%), which appears to be comparable with the frequency of the other studies.\textsuperscript{5,33} Therefore, our results regarding the frequency of KRAS mutation are in keeping with the published reports. The missing gender bias and the higher frequency among smokers repeated in our sample support the validity of our data. Till date, the clinical applicability of the KRAS mutation is limited, because

\begin{table}
\centering
\begin{tabular}{|l|c|c|c|}
\hline
Histology of the patients not included in the study, because of insufficient probe size or denial of informed consent & Per cent & Total & Men & Women \\
\hline
Total & 180 & 111 & 69 \\
Adenocarcinoma & 70 & 38.9 & 38 & 32 \\
Squamous cell carcinoma & 47 & 26.1 & 34 & 13 \\
Large cell carcinoma & 17 & 9.4 & 10 & 7 \\
Sarcomatoides Karzinom & 8 & 4.4 & 7 & 1 \\
Adenosquamous & 0 & 0.0 & 0 & 0 \\
NOS (non other specified) & 38 & 21.1 & 22 & 16 \\
\hline
\end{tabular}
\caption{Histology of the patients not included in the study, because of insufficient probe size or denial of informed consent.}
\end{table}

\begin{table}
\centering
\begin{tabular}{|l|c|c|c|c|c|}
\hline
Lung cancer stages of the investigated cohort and the excluded patients (according to UICC 7th edition) & Intention to treat population (n=732) & Per cent & Study Population (n=522) & Per cent & Patients excluded from analyses (n=180) & Per cent \\
\hline
Clinical Stage (n, %) & & & & & & \\
Staging not completed & 13 & 1.8 & 9 & 1.6 & 4 & 2.2 \\
IA & 76 & 10.4 & 56 & 10.1 & 20 & 11.1 \\
IB & 50 & 6.8 & 36 & 6.5 & 14 & 7.8 \\
IIA & 44 & 6.0 & 35 & 6.3 & 9 & 5.0 \\
IIB & 54 & 7.4 & 44 & 8.0 & 10 & 5.6 \\
IIIA & 116 & 15.8 & 95 & 17.2 & 21 & 11.7 \\
IIIB & 82 & 11.2 & 70 & 12.7 & 12 & 6.7 \\
IV & 297 & 40.6 & 207 & 37.5 & 90 & 50.0 \\
\hline
UICC, Union for International Cancer Control.
\end{tabular}
\caption{Lung cancer stages of the investigated cohort and the excluded patients (according to UICC 7th edition).}
\end{table}
there are no clinical associations strong enough for inclusion in clinical or therapeutic algorithms. However, KRAS mutation exclude or are only very rarely associated with other mutations (eg, EGFR). KRAS may therefore help to build up a testing algorithm in order to get a reasonable genetic array for this type of cancer.

In summary, for the first time we offer unbiased mutation rates for EGFR and KRAS in an unselected cohort of Caucasian patients. Previously described mutation rates for EGFR need to be corrected to approximately 5%, thus lowering the proportion of patients with NSCLC entitled to personalised medicine in Central Europe. However, in keeping with the literature non-smoking women with ACA do have the highest frequency of tumours with an activating EGFR mutation.

Contributors CB, JK and TM worked on the idea of the study and were responsible for the design and the setup. AR, SSF, DM and WS organised patient specimens and analyses and thus produced the data and were responsible for its analysis in tables and graphs. CB, JK, DM and TTB performed data analyses and wrote the initial manuscript. All authors contributed to the final manuscript and the interpretation of the data. All authors read and approved the final manuscript.

Funding Oskar-Helene Heim Foundation, Berlin, Germany.

Competing interests None.

Ethics approval Routine informed consent was obtained for the diagnostic procedure.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES