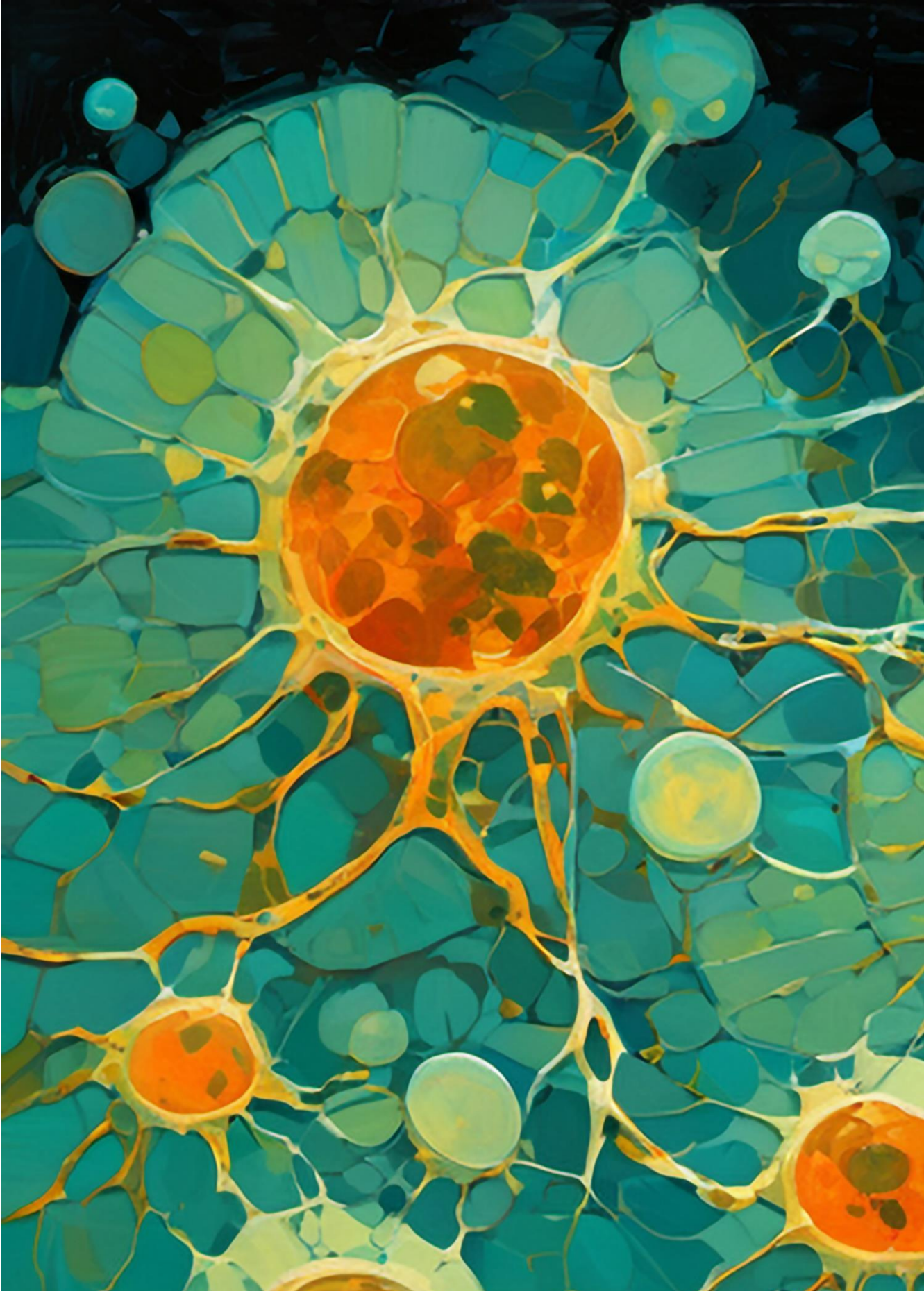
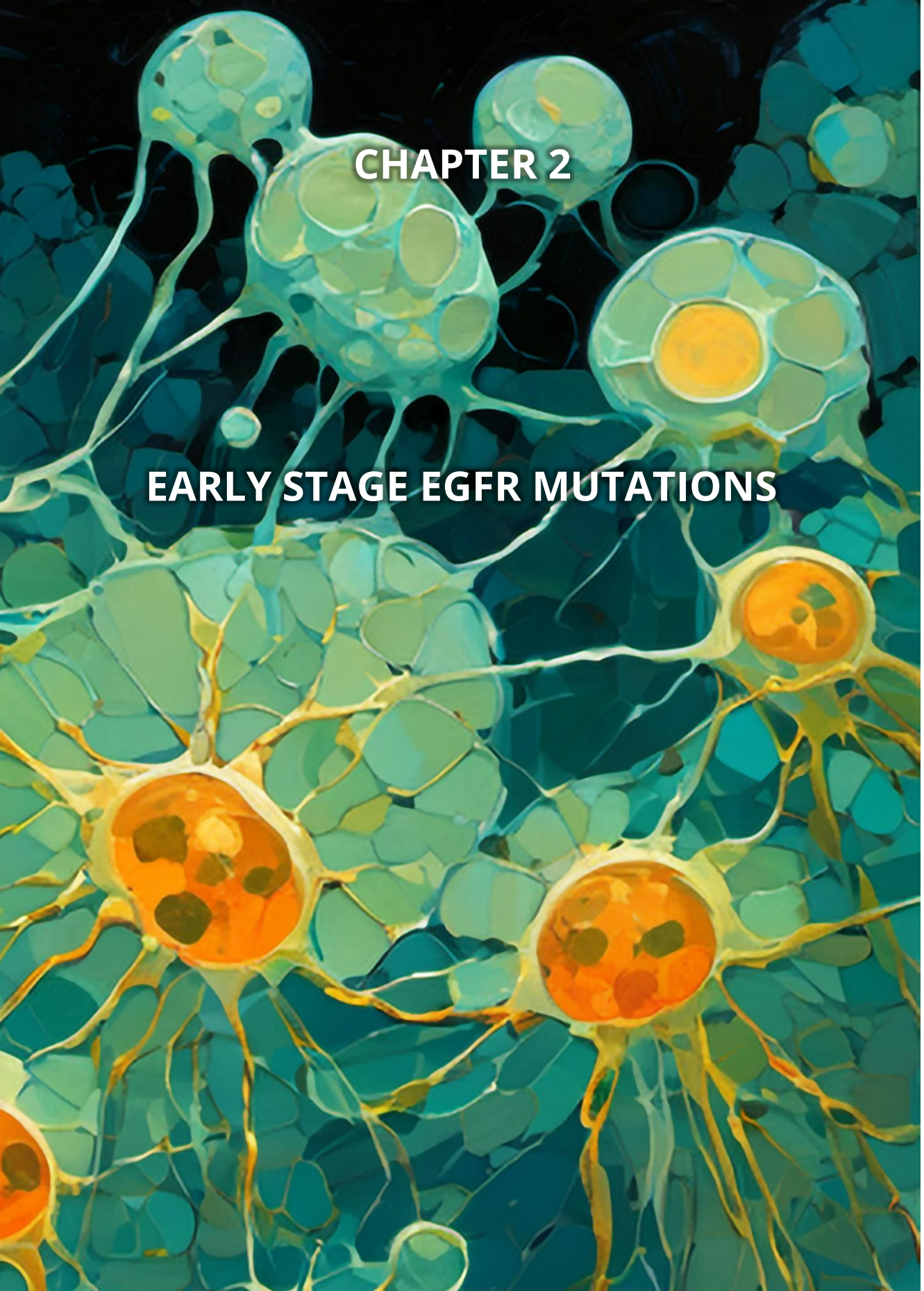
Diagnostic challenges of today's lung cancer pathology: personalizing therapy by immunohistochemical and molecular biomarkers

Hondelink, L.M.





CHAPTER 2

EARLY STAGE EGFR MUTATIONS

Chapter 2: Early stage EGFR mutations

2.1 Title page

Prevalence, clinical and molecular characteristics of early-stage *EGFR*-mutated lung cancer in a real-life West-European cohort: implications for adjuvant therapy.

Short running title: *EGFR* in early-stage lung cancer

Published: European Journal of Cancer, 2023. DOI: 10.1016/j.ejca.2022.12.010

2.1.1 Authors

Authors: Liesbeth M. Hondelink[1], Sophie M. Ernst[2], Peggy Atmodimedjo[3], Danielle Cohen[1], Janina L. Wolf[3], Anne-Marie C. Dingemans[2], Hendrikus J. Dubbink[3], Jan H. von der Thüsen[3]

2.1.2 Affiliations

1. Department of Pathology, Leiden University Medical Center, the Netherlands
2. Department of Respiratory Medicine, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, the Netherlands
3. Department of Pathology and Clinical Bioinformatics, Erasmus Medical Center, the Netherlands
4.

2.1.4 Acknowledgement

The authors thank the molecular diagnostics team at the Department of Pathology and Clinical Bioinformatics, Erasmus Medical Center.

2.1.5 Funding

This study received funding from AstraZeneca for the generation of data.

2.2 Abstract

2.2.1 Objectives

The landmark ADAURA study recently demonstrated a significant disease-free survival benefit of adjuvant osimertinib in patients with resected EGFR-mutated lung adenocarcinoma. However, data on prevalence rates and stage distribution of EGFR mutations in NSCLC in Western populations are limited since upfront EGFR testing in early-stage lung adenocarcinoma is not common practice. Here we present a unique, real-world, unselected cohort of lung adenocarcinoma to aid in providing a rationale for routine testing of early-stage lung cancers for EGFR mutations in the West-European population.

2.2.2 Material and Methods

We performed routine unbiased testing of all cases, regardless of TNM stage, with targeted next generation sequencing (NGS) on 486 lung adenocarcinoma cases between 01-01-2014 and 01-02-2020. Clinical and pathological data, including co-mutations and morphology, were collected. *EGFR*-mutated cases were compared to *KRAS*-mutated cases to investigate *EGFR*-specific characteristics.

2.2.3 Results

53 of 486 lung adenocarcinomas (11%) harbored an *EGFR* mutation. In early-stages (stage 0-IIIA) the prevalence was 13%, versus 9% in stage IIIB-IV. 9 out of 130 (7%) stage IB-IIIA patients fit the ADAURA criteria. Early-stage cases harbored more *L858R* mutations ($p = 0.02$), fewer exon 20 insertions ($p = 0.048$), fewer *TP53* co-mutations ($p = 0.007$), and were more frequently never smokers ($p = 0.04$) compared to late-stage cases with *EGFR* mutations. The *KRAS*-mutated cases were distributed more evenly across TNM stages compared to the *EGFR*-mutated cases.

2.2.4 Conclusion

As (neo-)adjuvant targeted therapy regimes enter the field of lung cancer treatment, molecular analysis of early-stage NSCLC becomes relevant. Testing for *EGFR* mutations in early-stage lung adenocarcinoma holds a substantial yield in our population, as our number needed to test ratio for adjuvant osimertinib was 14.4. The observed differences between early- and late-stage

disease warrants further analysis to work towards better prognostic stratification and more personalized treatment.

2.3 Introduction

Almost 30% of patients with non-small cell lung cancer (NSCLC) present with resectable early-stage disease. [1] Unfortunately, recurrence rates after resection are high: up to 50% of patients present with lung cancer recurrence within 5 years, which underscores the need for effective (neo)adjuvant treatment strategies. [2] Currently, in most patients with completely resected stage II-IIIA disease adjuvant platinum-based chemotherapy is recommended. However, the 5-year survival benefit of adjuvant chemotherapy remains limited. [3] Therefore, certain therapies that have proven to be effective in the advanced setting, such as immunotherapy and tyrosine kinase inhibitors (TKI), are now also of interest for the adjuvant setting. For instance, the landmark ADAURA trial has recently led to the approval of osimertinib, a third generation TKI, as adjuvant treatment after complete resection in patients with stage IB-IIIA NSCLC harboring EGFR exon 19 deletions or L858R substitution mutations. [4]

Pathogenic mutations in the *EGFR* gene are one of the most common oncogene driver mutations in metastatic NSCLC. The incidence of *EGFR* mutations in advanced non-squamous NSCLC varies greatly, from around 10% in West-European populations, to as high as 64% in the East Asian population. [5-11] The introduction of TKIs that inhibit the downstream pathways of *EGFR*, have greatly improved the outcome of patients with metastatic *EGFR*-mutated NSCLC. [12, 13] Osimertinib increased the median progression-free survival to 18.9 months [12] and the overall survival to 38.6 months. [14] Recently, the ADAURA investigators also demonstrated a substantial clinical benefit of adjuvant osimertinib in patients with resected *EGFR*-mutated lung adenocarcinoma. The study was discontinued early due to a significant efficacy benefit shown at interim analysis: patients with stage IB-IIIA disease receiving adjuvant osimertinib had a 24-month disease-free survival of 89%, versus only 52% in the placebo group ($p<0.001$), with a hazard ratio of 0.20 for disease recurrence and death. [15] However, currently the secondary endpoint of overall survival remains immature, and is hampered by the early unblinding of the study.

Until now, molecular screening for *EGFR* has only been routinely performed as part of standard care in stage IIIB and IV disease to select patients for treatment with osimertinib or other *EGFR* TKIs. [5, 6, 16] The expansion of routine molecular analysis to all early-stage lung adenocarcinomas to select patients for adjuvant treatment warrants a well-founded approach. To construct such an approach, several questions still need to be answered. There is a considerable amount of literature available on the prevalence of *EGFR* mutations in late-stage NSCLC and in the East Asian population. [17] However, as upfront *EGFR* testing in early-stage disease is not common practice, most reports on early-stage *EGFR*-mutated lung adenocarcinoma are from preselected cohorts, often enriched for *EGFR* mutations. [18] Therefore, it is still unclear how prevalent *EGFR* mutations are in early-stage *EGFR*-mutated lung adenocarcinomas in the Western population, and how to identify the patients who are at higher risk of recurrence and would therefore potentially have greater benefit of adjuvant treatment. These lacunae are essential to fill, as they could have implications for justified patient selection for adjuvant TKI treatment.

In the Erasmus Medical Center in Rotterdam, the Netherlands, all lung adenocarcinomas are subject to targeted next generation sequencing (NGS) testing regardless of TNM stage, so-called 'reflex-testing'. This provides a unique opportunity to investigate the real-world prevalence of *EGFR* mutations in early-stage NSCLC in a West-European patient population. Here we present our prospective unselected cohort of consecutive lung adenocarcinomas that were diagnosed in our center over the course of 6 years, using patients with *KRAS*-mutated NSCLC as a comparator for *EGFR*-mutated NSCLC. Additionally, we investigated the clinicopathological features, such as co-mutations and morphology, that are potentially associated with a higher risk for disease recurrence in early-stage *EGFR*-mutated NSCLC.

2.4 Materials and methods

2.4.1 Case collection and study setup

All in-house lung adenocarcinoma core needle biopsies, cytology specimens or resection samples of the Erasmus Medical Center Rotterdam (EMC) that were submitted to the pathology department for routine diagnostic purposes between 01-01-2014 and 01-02-2020 were evaluated for inclusion. Cases had

to have been analyzed with targeted DNA NGS with a customized oncogene-panel and have complete TNM staging for inclusion. In the case of multiple primary tumors per patient, each primary adenocarcinoma was eligible for inclusion if NGS had been performed. Both cytology and histology specimens were included, consisting of metastatic as well as primary tumor specimens. Only primary diagnostic specimens were allowed; liquid biopsy specimens and sequential biopsies after start of systemic treatment were excluded. Cases with insufficient tissue for DNA NGS or without complete TNM staging were excluded, which for example occurred if the patient opted to be referred to another medical center for staging, or if the patient was terminally ill with a concurrent disease.

To investigate whether possible differences between early- and late-stage cases are *EGFR*-specific, we compared the *EGFR* cases to the *KRAS*-mutated cases of our cohort.

2.4.2 DNA isolation

Formalin-fixed paraffin-embedded (FFPE) tissue, including cytology cell blocks, was used for DNA isolation. The DNA was isolated as previously described. [19] The acquired DNA was stored at -20°C until analysis.

2.4.3 DNA NGS

For targeted DNA NGS, an IonTorrent custom targeted NGS panel was used, including the following genes: *CDKN2A* (coverage 98%), *PTEN* (coverage 94%), *TP53* (coverage 100%) and mutation hotspots in *AKT1* (exon 3), *ALK* (20, 22-25), *APC* (14), *ARAF* (7), *BRAF* (11, 15), *CTNNB1* (3, 7, 8), *EGFR* (18-21), *HER2* (19-21), *EZH2* (16), *FBWX7* (9, 10), *FGFR1* (4, 7, 12), *FGFR2* (7, 9, 12), *FGFR3* (7, 9), *FOXL2* (1), *GNA11* (4, 5), *GNAQ* (4, 5), *GNAS* (8, 9), *HRAS* (2-4), *IDH1* (4), *IDH2* (4), *KIT* (8, 9, 11, 13, 14, 17), *KRAS* (2-4), *MAP2K1* (2, 3), *MET* (2, 14, 19), *MYD88* (5), *NOTCH1* (26, 27), *NRAS* (2-4), *PDGFRA* (12, 14, 18), *PIK3CA* (10, 21), *POLD1* (12), *POLE* (9, 13), *RAF1* (7), *RET* (11, 16), *RNF43* (3, 4, 9), *ROS1* (38, 41), *SMAD4* (3, 9, 12), *STK11* (4, 5, 8) and *TERT* promotor, as previously described. [20] Copy number calling was performed with SNPitty. [21, 22]

Genomic alterations were classified according to the ACMG/AMP consensus paper in 5 classes of ascending likelihood of pathogenicity. [23] For *EGFR* mutations, both class 4 or 5 pathogenic mutations and variants of unknown

significance (VUS) were included. We considered non-*EGFR* and non-*KRAS* mutations as co-mutations, including other driver mutations. Only class 4 and 5 pathogenic mutations were included, VUS were not considered co-mutations. Pathogenicity was assessed with reference databases, including Alamut, ClinVar, IARC, CKB and cBioportal. *KRAS* mutations were classified in G12C, G12D, G12V, Q61H and other mutations.

Additionally, we assessed the immunohistochemical expression pattern of *p53* in the *EGFR*-mutated cases if available.

2.4.4 Clinical parameters

For all cases, clinical data regarding age at diagnosis, TNM stage (7th edition) and sex were collected. For patients with *EGFR*-mutated adenocarcinoma, we collected additional data on the smoking history, recurrence-free survival (RFS) for early-stage cases, previous cytotoxic therapy for another malignancy, follow up time, symptoms at the time of diagnosis and prior lung cancer screening or monitoring. Stage 0-IIIA were considered early-stage disease, and stage IIIB and IV were considered late-stage disease. RFS was defined as time from date of diagnosis until disease recurrence.

Patients were categorized as 'current smokers' if they smoked in the month before diagnosis. Patients were considered to be 'former smokers' if they quit smoking at least one month before diagnosis. Patients were considered to be 'never smokers' if they had accumulated less than one pack year and had not smoked in the month before diagnosis.

2.4.5 Morphology

Growth patterns were assessed by one or multiple experienced thoracic pathologists, using a continuous score for each of the following categories: percentage lepidic, percentage acinar-papillary, percentage micropapillary-solid. The continuous scores for each category were used to assess the 'most prevalent growth pattern' and the 'worst growth pattern'. The 'most prevalent growth pattern' was the pattern which was most prevalent. If two patterns were equally prevalent, the worst growth pattern was used as the most prevalent growth pattern.

Literature has previously suggested that the type of growth pattern has potential prognostic value, with micropapillary-solid having the worst

prognosis, followed by acinar-papillary, and a lepidic growth pattern having the most favorable prognosis. [24] We therefore also scored the cases according to the pattern with the assumed worst prognosis, i.e. the 'worst growth pattern', to evaluate whether the presence of a less favorable growth pattern indeed has prognostic value. Growth pattern assessment was only performed for cases in which tissue from the primary tumor was available. Cytology specimens and metastasis biopsies were not scored for growth pattern. Examples of these scoring systems are outlined in Supplementary Table 1.

2.4.6 Statistics

We used IBM SPSS Statistics software, version 25 for statistical analysis. Statistical significance was set at $p < 0.05$. Categorical data were compared using the Chi Square test or Fisher Exact test, as appropriate. For t-distributed stochastic neighbor embedding (t-SNE) data visualization, we adapted the dataset. We normalized all continuous and ordinal data, such as age and TNM stage to values between 0 and 1. We used one-hot-encoding for non-ordinal categorical data, including *EGFR* mutations and co-mutations. We performed Mean Imputation for missing values in normally distributed continuous data and binary data. We performed Median Imputation for missing non-normally distributed continuous data and categorical data. [25] T-SNE was created with Python 3.7, using scikit-learn and perplexity values of 4 and 12 to plot these t-SNE figures. [26] The stage labels were excluded from the t-SNE data.

2.4.7 Ethics

This study was approved by the local medical ethical committee, registration number: MEC-2020-0732. Informed consent was not necessary and patient data were anonymized before processing.

2.5 Results

2.5.1 Case characteristics

We included 486 new lung adenocarcinoma cases, 53 (11%) harbored an *EGFR* mutation and 129 (27%) harbored a *KRAS* mutation. Cases were spread unevenly across TNM stages, with fewer patients in stage 0 (in situ carcinoma) and II and more patients in stage I and IV (Table 1).

CASE CHARACTERISTICS	ALL CASES (N = 486)	*EGFR*-MUTATED (N = 53)	*KRAS*-MUTATED (N = 129)
STAGE 0	11 (2%)	3 (6%)	3 (2%)
STAGE IA	114 (23%)	21 (40%)	31 (24%)
STAGE IB	38 (8%)	4 (8%)	13 (10%)
STAGE IIA	16 (3%)	2 (4%)	3 (2%)
STAGE IIB	17 (3%)	0	5 (4%)
STAGE IIIA	59 (12%)	3 (6%)	13 (10%)
STAGE IIIB	25 (5%)	1 (2%)	7 (5%)
STAGE IV	206 (42%)	19 (36%)	54 (42%)
EARLY-STAGE (0-IIIA)	255 (52%)	33 (62%)	68 (53%)
LATE-STAGE (IIIB-IV)	231 (48%)	20 (38%)	61 (47%)

***Table 1**: Case overview per TNM stage (TNM 7th edition).*

2.5.2 Prevalence of EGFR mutations per TNM stage

EGFR mutations were more prevalent in early-stage adenocarcinoma (13% of stage 0-IIIA patients harbored an *EGFR* mutation), compared to late-stage (9% of stage IIIB-IV patients harbored an *EGFR* mutation). The percentage of patients harboring *EGFR* mutations was especially high in stage 0 (27%) and 1A (18%), compared to the other stages (p = 0.03) (Figure 1). Of the 33 patients with early-stage *EGFR*-mutated NSCLC, 9 (27%) fit the ADAURA criteria (*L858R* mutation or exon 19 deletion, stage IB-IIIA). Since we included 130 stage IB-IIIA in our EMC cohort, the number of stage IB-IIIA cases needed to test in order to identify one patient eligible for adjuvant osimertinib following the ADAURA regimen, is 14.4.

2.5.3 Characteristics of early versus late-stage EGFR-mutated adenocarcinoma

We compared clinical, molecular and morphological parameters between the early-stage and the late-stage *EGFR* cases (Table 2), as well as between *EGFR* and *KRAS* cases (Figure 2). *EGFR*-mutated, early-stage cases harbored significantly more *EGFR L858R* mutations (45% vs 15%, p = 0.02), and were more likely to have a predominantly lepidic growth pattern (65% versus 0%, p = 0.003) than the late-stage *EGFR*-mutated cases. Late-stage cases more often harbored *EGFR* exon 20 insertions (25% versus 6%, p = 0.048) and were enriched for *TP53* co-mutations (65% versus 27%, p = 0.007). Within the *TP53* mutated cases, late-stage harbored more disruptive *TP53* mutations than early-stage cases (40% versus 0%, p < 0.001). The *KRAS* early- and late-stage cohorts

differed with regard to *TP53* mutation prevalence (31% versus 52%, respectively, p = 0.02), with late-stage cases again harboring more disruptive *TP53* mutations, though not significantly (15% versus 9%, p = 0.4).

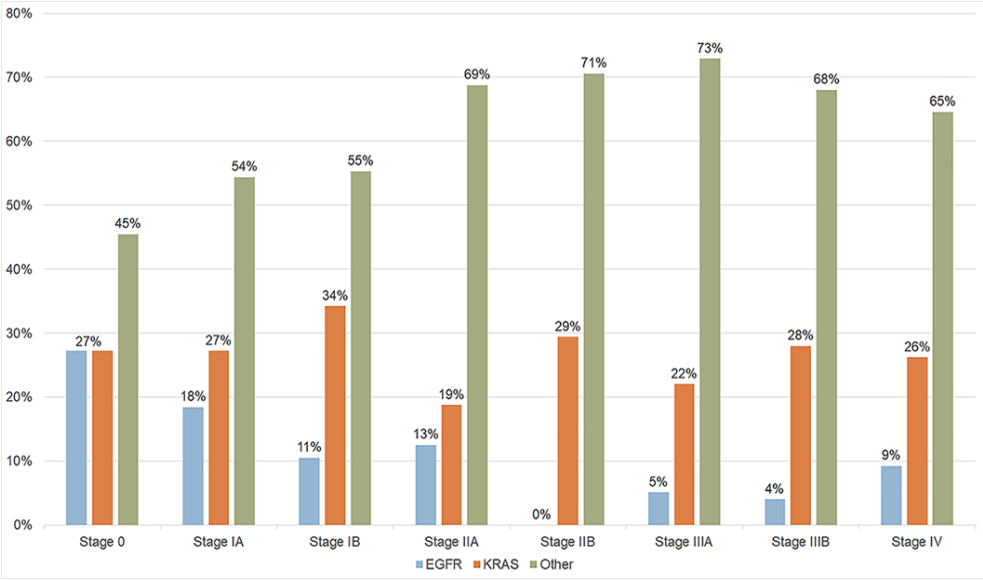


Figure 1: *Mutation prevalence across stages. Prevalence of EGFR-mutated cases, KRAS-mutated cases and Other cases per TNM stage (TNM 7th edition). Blue: EGFR; Orange: KRAS; Green: other cases. KRAS is evenly distributed across stages, whereas EGFR prevalence differs across stages.*

Additionally, early- and late-stage *EGFR*-mutated cases differed significantly with regard to smoking history (p = 0.04). We did not identify differences in age, sex, and worst growth pattern between early- and late-stage disease. In 8 of the *TP53* mutated cases *p53* immunohistochemistry was performed: 7 showed strong nuclear expression for *p53*, whereas one had absent nuclear expression.

Prior to diagnosis, 9 patients (27% of all early-stage *EGFR*-mutated cases) were monitored with computed tomography (CT) scans for a 'ground glass' lesion or pulmonary node, for an average time period of 3.1 years (range 1-7 years). Of these cases, 4 harbored a *L858R* mutation, 4 an exon 19 deletion, and one an exon 20 insertion. 4 cases harbored a non-disruptive *TP53* mutation. 7 had a predominantly lepidic growth pattern, and the remaining 2 cases had acinar growth patterns. Two other patients were not monitored, but the tumor had in

retrospect been visible on previous imaging, 15 and 17 years prior to the diagnosis, respectively.

From the 486 included cases, 129 were *KRAS*-mutated, including 68 early-stage and 61 late-stage cases. The characteristics for the *KRAS* cohort are outlined in Supplementary Table 2. The *EGFR*-mutated and *KRAS*-mutated cohorts differ with regard to smoking history and pre-diagnosis follow up, with more current smokers in the *KRAS* cohort (42% versus 11%, p<0.001), more never-smokers in the *EGFR* cohort (28% versus 2%, p<0.001) and more often pre-diagnosis follow-up in the *EGFR* cohort (17% versus 5%, p = 0.03). In contrast to the *EGFR*-mutated cases, the *KRAS*-mutated cases were distributed more evenly across TNM stages (Figure 1). Also, *EGFR* early- and late-stage cases differed significantly with regard to mutation type, predominant growth pattern, and co-mutation prevalence, whereas this was not the case for the *KRAS* cohort.

FEATURE N (%)	EARLY-STAGE *EGFR* (N = 33)	LATE-STAGE *EGFR* (N = 20)	P-VALUE	EARLY-STAGE *KRAS* (N = 68)	LATE-STAGE *KRAS* (N = 61)	P-VALUE
EGFR L858R	15 (45%)	3 (15%)	0.02[a]	N/A	N/A	N/A
EGFR EXON 20 INS	2 (6%)	5 (25%)	0.048[a]	N/A	N/A	N/A
EGFR EXON 19 DEL	13 (39%)	9 (45%)	0.7	N/A	N/A	N/A
OTHER *EGFR*	3* (9%)	3• (15%)	0.5[a]			
TP53	9 (27%)	13 (65%)	0.007[a]	21 (31%)	32 (52%)	0.02[a]
TP53 DISRUPTIVE	0	8 (40%)	<0.001[a]	6 (9%)	9 (15%)	0.4[a]
MOST PREVALENT GROWTH PATTERN			0.003[b]			0.6[b]
LEPIDIC	20 (65%)	0 (0%)		22 (38%)	6 (33%)	
ACINAR OR PAPILLARY	9 (29%)	3 (15%)		31 (53%)	9 (50%)	
SOLID OR MICROPAPILLARY	2 (6%)	3 (15%)		5 (9%)	3 (17%)	
NOT SCORED	2 (6%)	14 (70%)				
SMOKING STATUS			<0.001[b]			0.4[b]

NEVER SMOKER	21 (64%)	1 (5%)		2 (3%)	0	
FORMER SMOKER	10 (30%)	15 (75%)		34 (56%)	27 (48%)	
CURRENT SMOKER	2 (6%)	4 (20%)		25 (41%)	29 (52%)	
UNKNOWN	0	0		7 (10%)	5 (8%)	
PRIOR TO DIAGNOSIS			0.02[a]			0.7[a]
PRIOR FOLLOW-UP	9 (27%)	0		5 (7%)	1 (2%)	
NO PRIOR FOLLOW-UP	24 (73%)	19 (95%)		52 (76%)	26 (43%)	
UNKNOWN	0	1 (5%)		11 (16%)	34 (56%)	

***Table 2**: Significant differences between early-stage EGFR-mutated lung adenocarcinomas (n = 33) and late-stage EGFR-mutated lung adenocarcinomas (n = 20). Co-mutations were assessed only in cases with complete coverage of the panel, as described in the Methods. Predominant growth pattern was not available for cytology and metastasis specimens. P-values were calculated with (a) Fisher's Exact test or (b) Chi-squared test. For categories 'Smoking status' and 'Prior to diagnosis', missing data was omitted from percentage calculations and statistic testing. * 'Other' EGFR mutations included p.G779F, p.G719A and p.L861R. • 'Other' EGFR mutations included p.G719A, concomitant p.G719S and p.S768I, and p.V774L.*

2.5.4 Recurrence free survival (RFS)

Within the early-stage *EGFR* cases (n = 33), 3 patients (9%) had presented with disease recurrence after 7, 48 and 60 months respectively, 12 patients (36%) were recurrence-free for at least 2 years after resection, and 18 (55%) patients had a follow-up duration of less than 2 years. Type of *EGFR* mutation, presence of *TP53* mutations and clinical characteristics for the recurrence-free, recurrence and late-stage cases are summarized in Supplementary Figure 1. This illustrates that most late-stage cases harbor similar clinicopathological features (*EGFR* exon 20 insertions, presence of (*TP53*) co-mutations, growth pattern, previous or current tobacco smoke exposure), which can also partly be identified in the early-stage cases with recurrence although in a limited number of cases, and in some recurrence-free cases. With regard to the growth patterns, the recurrence-free cases were predominantly characterized by a lepidic growth pattern (67%), followed by an acinar growth pattern (10%).

Growth patterns differed in the 3 cases with recurrence: one case had a predominantly solid, one predominantly acinar and one predominantly lepidic growth pattern. The patient with the solid growth pattern had a RFS of 7 months, versus 48 months in the patient with predominantly acinar growth pattern and 60 months in the patient with the lepidic growth pattern.

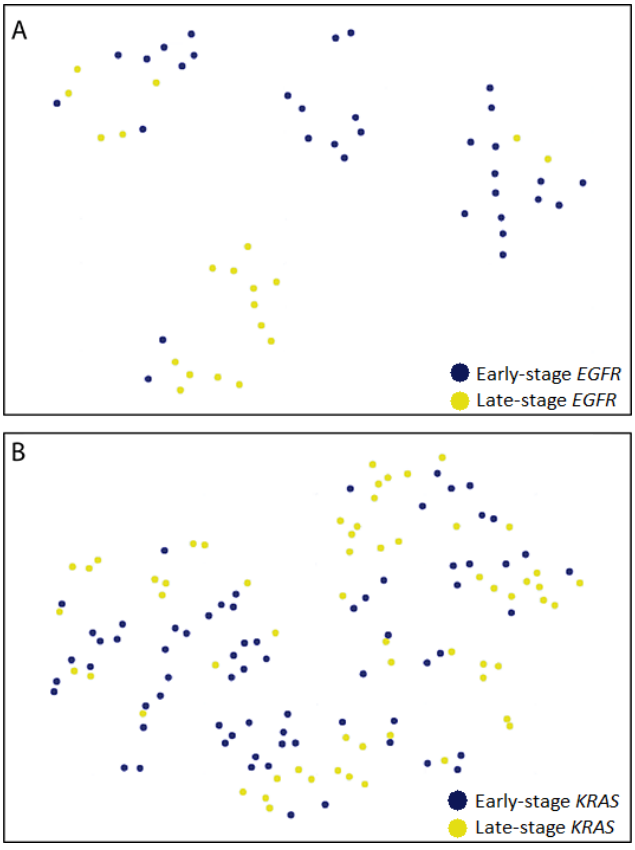


Figure 2: *Unsupervised clustering of EGFR- and KRAS-mutated cases. Unsupervised clustering, using t-distributed stochastic neighbor embedding (t-SNE). A: t-SNE of EGFR-mutated lung adenocarcinoma features, perplexity value 4. B: t-SNE of KRAS-mutated lung adenocarcinoma features, perplexity value 12. Blue dots: early-stage (0-IIIA, TNM 7th edition); yellow dots: late-stage (IIIB-IV). Features used for this t-SNE include: smoking history, symptoms, prior follow-up, T-stage, sex, age, growth pattern, EGFR mutations, KRAS mutations and co-mutations.*

To illustrate these different growth patterns, figure 3A depicts the aforementioned case with a solid growth pattern and disease recurrence after 7 months. This 64-year-old woman was referred to the pulmonologist with an

asymptomatic pulmonary nodule, discovered via a coincidental finding. She was a former smoker and had accumulated 22 pack years. A lung biopsy was taken (Figure 3A), and the patient was diagnosed with a lung adenocarcinoma with 100% solid growth pattern. Staging showed that the tumor is stage cT2aN0M0, and the patient is eligible for surgical resection. In the resection specimen, the tumor had infiltrated the visceral pleura (pT2aN0M0PL1) and harbored an *EGFR* L858R mutation. After 7 months, she was diagnosed with bone metastases, and treated with *EGFR* TKIs.

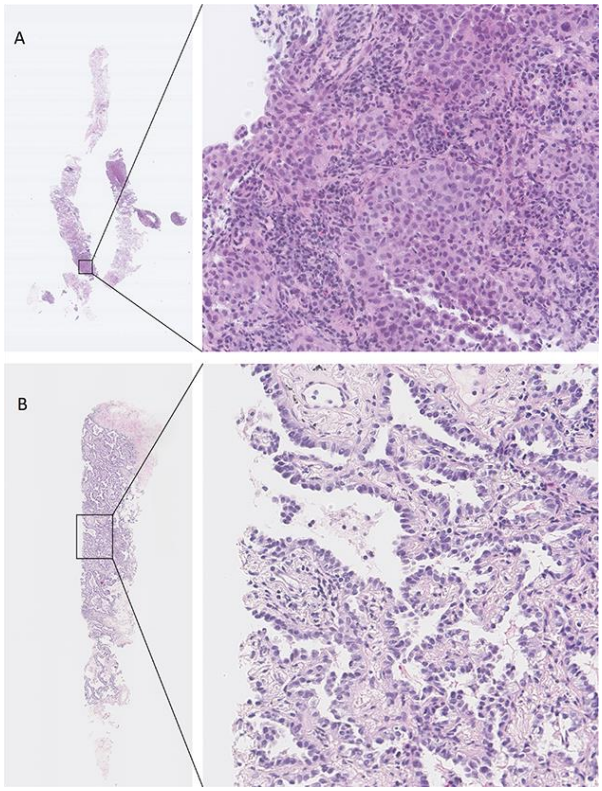


Figure 3: *Case descriptions. A: Case 1 biopsy. First image: 4x, close-up: 40x. B: Case 2 biopsy. First image: 4x, close-up: 40x.*

In contrast, figure 3B illustrates a case with a lepidic growth pattern in which no disease recurrence occurred. This 65-year-old woman was referred to the pulmonologist with a pulmonary lesion on CT-scan, discovered via a coincidental finding. She had smoked in the past, but had accumulated less than 10 pack years. On CT, a 'ground glass' lesion was identified, not suspicious for invasive malignancy. She was followed every 6 months with a CT-scan. After

2 years, the lesion had grown a few millimeters, and now had a small solid component. A lung biopsy (Figure 3B) revealed a 100% lepidic lung adenocarcinoma (IASLC grade 1). The patient was diagnosed with a cT1aN0M0 lung adenocarcinoma. Next-generation sequencing revealed an exon 19 deletion in *EGFR*, and no co-mutations. After surgical resection of the tumor, the patient is now recurrence free for 6 years.

2.6 Discussion

In this study, we investigated the prevalence of *EGFR* mutations across TNM stages in an unselected West-European cohort of 486 lung adenocarcinomas in which NGS reflex testing was performed. We found that *EGFR* mutations are unevenly spread over TNM stages, with a prevalence of 13% in early-stage, and 9% in late-stage. The latter is in line with previously reported prevalence rates of *EGFR* mutations in metastatic NSCLC in the Netherlands. [9, 11] 9 out of 130 (7%) stage IB-IIIA cases met the ADAURA inclusion criteria (*L858R* or exon 19 deletion), [15] which indicates that the number of stage IB-IIIA tumors needed to test in order to identify one patient eligible for adjuvant osimertinib is 14.4. Of note, we found that 36% of early-stage *EGFR*-mutated cases had current or previous tobacco smoke exposure. This highlights that selection for molecular analysis in the early-stage setting should also not be guided by clinical characteristics such as smoking history. These real-world data provide a rationale for routine testing of early-stage lung cancers for *EGFR* mutations in the West-European population.

Additionally, we provided a descriptive analysis of the characteristics of *EGFR*-mutated NSCLC over disease stages. We found that early-stage *EGFR*-mutated cases differ from late-stage cases with respect to clinical, genomic, and morphological characteristics. The late-stage group harbors more exon 20 insertions and fewer *L858R* mutations, more *TP53* mutations, more patients with previous or current tobacco smoke exposure, and more high-grade growth patterns. Although the *KRAS*-mutated late-stage cases also had a higher prevalence of *TP53* mutations than the early-stage cases, the *KRAS*-mutated cohort seemed more homogeneous over tumor stages. This could imply that the differences between early- and late-stage disease in the *EGFR*-mutated cohort are *EGFR*-specific.

In our *EGFR*-mutated early-stage cases, three patients presented with disease recurrence after an average of 3.2 years. This is longer than the average time to recurrence in NSCLC, as in most post-surgical NSCLC cases occult metastases present within 2 years after surgery. [27, 28] In addition, we found that 27% of all early-stage *EGFR*-mutated cases had been monitored prior to diagnosis because of 'ground glass' lesions. Recent data showed a 5-year overall survival rate of 100% in patients with surgically resected clinical stage 1A *EGFR*-mutated lung adenocarcinoma with ground glass opacity component [29]. In the *KRAS* cohort significantly less patients were followed up prior to diagnosis. This could suggest that some *EGFR*-mutated tumors are 'slow growers', and occult metastases – if present – are only identified after a long follow-up. Therefore, further studies with long survival data could aid in optimizing the timing of resection and surveillance strategies of resected *EGFR*-mutated carcinomas.

In all, these results suggest that *EGFR*-mutated lung adenocarcinoma is not one homogeneous disease, but rather that there are subgroups that could be defined by their different phenotypes. Although we have a limited sample size, it seems that some patients with (high) tobacco exposure, high grade growth pattern, *EGFR* exon 20 insertion and *TP53* mutation often present at a higher TNM stage and often progress to a higher stage. On the other hand, patients who have never smoked, with common *EGFR* mutations without co-mutations and with a low-grade growth pattern are rare in the high TNM stage group and the metastasis group. We should further investigate whether these findings truly indicate a 'high risk' and 'low risk' subtype in larger case series, as this could potentially help clinicians and pathologists identify patients who are at a higher risk of recurrence after surgery than others. It can be hypothesized that 'high risk' patients could derive more benefit from adjuvant TKI treatment than patients who were already at a low risk of recurrence, which could have implications for the prevention of over- and undertreatment.

The main limitation of our study is the sample size. While we screened a substantial number of cases (n = 486), 53 cases harbored an *EGFR* mutation. This is a limited dataset, especially in subset analyses. Consequently, our comparison between, for example, early-stage recurrence and recurrence free disease only included a small number of patients. Therefore, it is possible that our analysis lacked the power to detect smaller differences. However, this did

not limit our primary objective of determining *EGFR* prevalence rates across TNM stages.

In conclusion, the prevalence of *EGFR* mutations in early-stage lung adenocarcinoma in our West-European patient population is 13%, and the prevalence of ADAURA-eligible *EGFR* mutations in stage IB-IIIA is 7%, which constitutes a substantial yield when combining this number with the demonstrated benefit of adjuvant osimertinib. [15] However, we must emphasize that screening for *EGFR* mutations in early-stage lung adenocarcinoma is only a first step. Our data adds to a growing body of evidence that suggests that *EGFR*-mutated lung cancer, although seemingly one homogeneous group, actually consists of several genomic and clinical subgroups, in which we can potentially start to define low-risk and high-risk phenotypes that are correlated to clinical disease behavior. This underlines the intrinsic heterogeneity in NSCLC and the importance of comprehensive tumor characterization in clinical practice, as well as in future research. It would be of interest to investigate potential differences in outcomes between patients with low and high-risk phenotypes receiving adjuvant TKIs such as osimertinib, in order to guide future therapy decisions.

References

1. Le Chevalier T, *Adjuvant chemotherapy for resectable non-small-cell lung cancer: where is it going?* Ann Oncol, 2010. **21 Suppl 7**: p. vii196-8.
2. Kratz JR, He J, Van den Eeden SK, et al., *A practical molecular assay to predict survival in resected non-squamous, non-small-cell lung cancer: development and international validation studies.* The Lancet, 2012. **379**(9818): p. 823-832.
3. Pignon JP, Tribodet H, Scagliotti GV, et al., *Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group.* J Clin Oncol, 2008. **26**(21): p. 3552-9.
4. Wu YL, Tsuboi M, He J, et al., *Osimertinib in Resected EGFR-Mutated Non-Small-Cell Lung Cancer.* N Engl J Med, 2020. **383**(18): p. 1711-1723.
5. Barlesi F, Mazieres J, Merlio JP, et al., *Routine molecular profiling of patients with advanced non-small-cell lung cancer: results of a 1-year nationwide programme of the French Cooperative Thoracic Intergroup (IFCT).* Lancet, 2016. **387**(10026): p. 1415-1426.
6. Cohen D, Hondelink LM, Solleveld-Westerink N, et al., *Optimizing Mutation and Fusion Detection in NSCLC by Sequential DNA and RNA Sequencing.* J Thorac Oncol, 2020. **15**(6): p. 1000-1014.
7. The Cancer Genome Atlas Research Network, *Comprehensive molecular profiling of lung adenocarcinoma.* Nature, 2014. **511**: p. 543-550.
8. Shi Y, Au JSK, Thongprasert S, et al., *A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER).* J Thorac Oncol, 2014. **9**(2): p. 154-62.
9. Ten Berge D, Aarts MJ, Groen HJM, et al., *A population-based study describing characteristics, survival and the effect of TKI treatment on patients with EGFR mutated stage IV NSCLC in the Netherlands.* Eur J Cancer, 2022. **165**: p. 195-204.
10. Gao B, Sun Y, Zhang J, et al., *Spectrum of LKB1, EGFR, and KRAS mutations in chinese lung adenocarcinomas.* J Thorac Oncol, 2010. **5**(8): p. 1130-5.
11. Koopman B, Garcia BNC, Kuijpers CCHJ, et al., *A Nationwide Study on the Impact of Routine Testing for EGFR Mutations in Advanced NSCLC Reveals Distinct Survival Patterns Based on EGFR Mutation Subclasses.* Cancers (Basel), 2021. **13**(14).
12. Soria J, Ohe Y, Vansteenkiste J, et al., *Osimertinib in Untreated EGFR-Mutated Advanced Non–Small-Cell Lung Cancer.* N Engl J Med, 2018. **378**: p. 113-125.
13. Mok TS, Wu YL, Ahn MJ, et al., *Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer.* N Engl J Med., 2017. **376**(7): p. 629-640.
14. Ramalingam SS, Vansteenkiste J, Planchard D, et al., *Overall Survival with Osimertinib in Untreated, EGFR-Mutated Advanced NSCLC.* N Engl J Med, 2019. **382**(1): p. 41-50.
15. Herbst RS, Morgensztern D, Boshoff C, *The biology and management of non-small cell lung cancer.* Nature, 2018. **553**(7689): p. 446-454.

16. Pi C, Xu CR, Zhang MF, et al., *EGFR mutations in early-stage and advanced-stage lung adenocarcinoma: Analysis based on large-scale data from China.* Thorac Cancer, 2018. **9**(7): p. 814-819.
17. Saw SPL, Zhou S, Chen J, et al., *Association of Clinicopathologic and Molecular Tumor Features With Recurrence in Resected Early-Stage Epidermal Growth Factor Receptor-Positive Non-Small Cell Lung Cancer.* JAMA Netw Open, 2021. **4**(11): p. e2131892.
18. van Lier MG, Wagner A, van Leerdam ME, et al., *A review on the molecular diagnostics of Lynch syndrome: a central role for the pathology laboratory.* J Cell Mol Med, 2010. **14**: p. 181-97.
19. Pruis MA, Geurts-Giele WRR, von der Thusen JH, et al., *Highly accurate DNA-based detection and treatment results of MET exon 14 skipping mutations in lung cancer.* Lung Cancer, 2020. **140**: p. 46-54.
20. van Riet J, Krol NMG, Atmodimedjo PN, et al., *SNPitty: An Intuitive Web Application for Interactive B-Allele Frequency and Copy Number Visualization of Next-Generation Sequencing Data.* J Mol Diagn, 2018. **20**(2): p. 166-176.
21. Dubbink HJ, Atmodimedjo PN, van Marion R, et al., *Diagnostic Detection of Allelic Losses and Imbalances by Next-Generation Sequencing: 1p/19q Co-Deletion Analysis of Gliomas.* J Mol Diagn, 2016. **18**(5): p. 775-786.
22. Boyd JA, Hubbs JL, Kim DW, et al., *Timing of local and distant failure in resected lung cancer: implications for reported rates of local failure.* J Thorac Oncol, 2010. **5**(2): p. 211-4.
23. Demicheli R, Fornili M, Ambrogi F, et al., *Recurrence Dynamics for Non–Small-Cell Lung Cancer: Effect of Surgery on the Development of Metastases.* J of Thorac Oncol, 2012. **7**(4): p. 723-730.
24. Li M, Xi J, Zhang H, et al., *Pan-Driver-Negatives versus Epidermal Growth Factor Receptor Mutants for C-Stage IA Lung Adenocarcinoma with Ground-Glass Opacity.* Ann Thorac Cardiovasc Surg, 2022.