General Introduction
1. General background on lung cancer

As the leading cause of cancer-related mortality, lung cancer is a major health problem worldwide. Globally across 185 countries, lung cancer is estimated to account for 12% (2.1 million) of all newly diagnosed cancer patients each year, thereby being the largest contributor. Lung cancer is also the main cause of cancer-related deaths, accounting for approximately 18% (1.8 million) of the total number of cancer deaths in 2018 worldwide [1]. The 5-year survival rate for lung cancer in the United States was only 19% from 2008 to 2014. In contrast, 5-year survival rates for all malignancies increased from 68% to 85% during the same time period [2, 3]. In the Netherlands the incidence of lung cancer was 13,262 cases in 2018; mortality was 10,886 cases in 2017 and the 5-year survival over the period 2011 - 2015 was 19% (https://www.cijfersoverkanker.nl/nkr).

The leading etiological risk factor of lung cancer is tobacco consumption [2]. Other risk factors include air pollution, genetic susceptibility (family history of cancer), chronic inflammatory lung diseases (e.g., interstitial lung diseases, pulmonary tuberculosis), chronic obstructive pulmonary disease, occupational exposures (e.g. asbestos, silica, radon, heavy metals (chromium-6), inhaled chemicals, etc.), and ionizing radiation [2].

Lung cancer is divided into two main histological subgroups, i.e. small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) [3,4]. SCLC constitutes about 12% of lung cancer cases. The major histologic types of NSCLC include adenocarcinoma (AC), squamous-cell lung carcinoma (SCC), large-cell neuroendocrine carcinoma, and pulmonary carcinoid tumour. AC is the most common subtype, making up 40% of all lung cancers and 60% of NSCLC. AC originates from atypical adenomatous hyperplasia, which may develop into adenocarcinoma in situ, minimally invasive AC in a stepwise manner before reaching the invasive AC stage. In general, AC is located in the peripheral parts of the lungs, but AC can also occur near the hilum. SCC is the second largest subtype with about 25-30% of all lung cancers. SCC arises from basal cell hyperplasia, which sequentially develops towards squamous metaplasia, dysplasia and carcinoma in situ before reaching the invasive carcinoma stage. SCC tends to occur in the central part of the lung and near major bronchi. The other histological subtypes represent a minority of lung cancer cases.

Staging of NSCLC according to the 8th edition of the IASLC staging classification is the basis for prognosis prediction and to a lesser extent for the choice of treatment [5]. The staging is based on the tumour node metastasis (TNM) classification, using the size of the primary tumour, extent of lymph node involvement, growth into the pleura and presence of distant metastasis. The T, N and M categories are determined by imaging techniques such as CT and PET. The N category is also determined by endoscopic ultrasound needle biopsies and cervical mediastinoscopy, sometimes along with a biopsy from a suspected metastasized site. Early and localized NSCLC stage I, II and IIIA are treated by surgery with adjuvant chemotherapy. In the very early stages of lung cancer no adjuvant chemotherapy is provided because no survival advantage has been observed. The effect of neoadjuvant chemotherapy and immunotherapy seems to give good results with five-year survival rates of 60% to 80% in a study with a limited number of patients [6]. Patients with regional NSCLC (stage IIIB) with cancer spreading to lymph nodes in the mediastinum, are treated with a combination of chemoradiation therapy in case patients have a good performance score and with sequential chemo and radiotherapy for those with a limited performance score. Recently, this is complemented with adjuvant immunotherapy. The five-year survival rate of regional NSCLC is 33% for the sequential chemoradiotherapy group. Late stage or advanced stage
NSCLC with distant metastasis (stage IIIIB, IIIC, and IV) comprises 60% of all NSCLC cases and this patient group has the lowest survival. Targeted therapy is used for advanced NSCLC patients that have a mutation or translocation that changes one of the tyrosine kinase receptors (RTK) into a constitutively activated state (Figure 1). Tyrosine kinase inhibitors (TKIs) include gefitinib, erlotinib, afatinib, crizotinib, osimertinib, etc. Immunotherapy can be given to patients with PD-L1 positive tumour cells, although some studies show efficacy in those without any PD-L1 expression [7]. Patients with advanced disease who have a “deep” tumour response (showing >90% tumour reduction) and have localized progression of disease – oligometastatic disease - may be candidates for local treatments such as surgery or stereotactic radiation in addition to the targeted therapy. This approach provides a very good prognosis.

2. Molecular pathogenesis

Like many other malignancies, lung cancer develops through an accumulation of distinct genetic and epigenetic alterations leading to activation of oncogenes and inhibition of tumour suppressor genes [8]. Oncogenes are genes facilitating survival, cell growth, proliferation and invasion. They are typically activated by specific (hotspot) mutations, by structural rearrangements leading to fusion genes or by amplifications. In lung cancer, commonly activated oncogenes include KRAS, BRAF, EGFR, ERBB2 (also known as HER2), MET, ALK, ROS1 and RET (Table 1) [9]. Tumour suppressor genes (TSGs) exert their function in regulating cell cycles, promoting apoptosis, and even controlling cell adhesion to prevent invasive growth and migration. Nonsense mutations, out-of-frame INDELs, and even specific non-synonymous mutations can lead to loss of function of these proteins and thereby contribute to cancer. In contrast to mutations in oncogenes, mutations in tumour suppressor genes are usually scattered throughout the entire gene. Commonly inactivated tumour suppressor genes in lung cancer include TP53, RB1, STK11, and PTEN [9].

Activation of RTKs has been proven to act as key drivers of lung cancer development, by activating crucial signaling pathways such as proliferation, differentiation, survival and migration. The molecular structure of RTKs consists of a ligand binding domain on the extracellular part of the protein, a single helix domain facilitating localization through the cell membrane, and an intracellular tyrosine kinase domain. The most commonly activated genes in lung AC are EGFR, ALK, RET, ROS1, ERBB2, MET, and NTRK1; while activating mutations in DDR2 and FGFR1 are commonly reported in SCC. In SCLC, no commonly activated RTKs have been reported so far. SCLC cases originating from EGFR+ transformed lung AC usually retain the activating EGFR mutation and also show genomic loss of Rb1 and TP53 [10].

2.1 Drug Targets

Unlike decades ago, when treatment was based on histological features of the tumour, molecular characteristics have become the standard for guiding management of lung AC. In the past decade, multiple drugs have been generated that inhibit a growing number of activated driver genes (Table 1). RTKs, like EGFR, ALK or RET and ROS1, target the same downstream pathways: RAS-RAF-MEK-ERK, MAPK, PI3K-AKT-mTOR and JAK-STAT pathways (Table 1, Figure 1) [11-13]. The most commonly used approved tyrosine kinase inhibitors (TKIs) target EGFR or ALK and are used in patients with activating EGFR mutations and ALK rearrangements, respectively. Although the recurrent “driver” mutations are less common in SCC patients, several targetable drivers have been identified including amplification of...
FGFR1, mutations of DDR2 and mutations in PIK3CA gene. More recently, blocking of the PD-1/PD-L1 immune checkpoint, has become a first-line treatment option for NSCLC patients with expression of PD-1/PD-L1 receptors [14]. The available checkpoint inhibitors for NSCLC are pembrolizumab as first choice, nivolumab, atezolizumab, and durvalumab. The effect of atezoluzimab acts best in combination with bevacizumab and chemotherapy.

Table 1. Genomic alterations in non-small cell lung cancer.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Aberration</th>
<th>NSCLC</th>
<th>Genomic alterations / fusion partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALK</td>
<td>Fusion</td>
<td>3-13%, more common in adenocarcinoma</td>
<td>EML4, KIF5B, KLC1, TPC, TFG, TPR, HIP1, STRN, DCTN1, SQSTM1, NPM1, BCL11A, and BIRC6</td>
</tr>
<tr>
<td>AKT1</td>
<td>Mutation</td>
<td>1%</td>
<td>E17K</td>
</tr>
<tr>
<td>BRAF</td>
<td>Mutation</td>
<td>1-4%, more common in adenocarcinoma</td>
<td>V600E</td>
</tr>
<tr>
<td>DDR2</td>
<td>Mutation</td>
<td>2.5-3.8% in SCC</td>
<td>S768R</td>
</tr>
<tr>
<td>EGFR</td>
<td>Mutation</td>
<td>10% in US and Europe, 35% in East Asia, more common in adenocarcinoma</td>
<td>L858R, E19 DEL/INS, G719X, L861Q, Exon 20 duplication</td>
</tr>
<tr>
<td>ERBB2</td>
<td>Mutation</td>
<td>2-4%</td>
<td></td>
</tr>
<tr>
<td>FGFR1</td>
<td>Amplification</td>
<td>20-22%, more common in SCC</td>
<td></td>
</tr>
<tr>
<td>KRAS</td>
<td>Mutation</td>
<td>15-30%</td>
<td></td>
</tr>
<tr>
<td>MET</td>
<td>Mutation</td>
<td>3-4%, more common in adenocarcinoma</td>
<td></td>
</tr>
<tr>
<td>MET</td>
<td>Amplification</td>
<td>~2-4% untreated NSCLC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>~5-20% with EGFR positive, TKI resistant patients</td>
<td></td>
</tr>
<tr>
<td>MEK1</td>
<td>Mutation</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>NRAS</td>
<td>Mutation</td>
<td>&lt;1%</td>
<td></td>
</tr>
<tr>
<td>NTRK1</td>
<td>Fusion</td>
<td>3.3%, more common in adenocarcinoma</td>
<td></td>
</tr>
<tr>
<td>PIK3CA</td>
<td>Mutation</td>
<td>1-3%, more common in SCC</td>
<td></td>
</tr>
<tr>
<td>PTEN</td>
<td>Deletion</td>
<td>4-8%</td>
<td></td>
</tr>
<tr>
<td>RET</td>
<td>Fusion</td>
<td>1-2%, more common in adenocarcinoma</td>
<td></td>
</tr>
<tr>
<td>ROS1</td>
<td>Fusion</td>
<td>1-2%, more common in adenocarcinoma</td>
<td></td>
</tr>
</tbody>
</table>

NA: not applicable

2.2 Drug Resistance

Patients inevitably develop resistance approximately one to two years after start of targeted therapy. Resistance to TKI drugs can be categorized as primary (intrinsic) resistance and secondary (or acquired) resistance. Primary resistance can be defined as unresponsiveness to first-line TKI treatment. Acquired resistance to TKIs refers to a relapse or disease progression after a complete or partial response based on imaging according to RECIST criteria [15,16]. For EGFR mutation positive patients, acquired resistance may be arbitrarily defined as progression after ≥6 months since start of treatment. This acquired resistance might reflect outgrowth of a pre-existing minor treatment resistant clone or outgrowth of a
clone with treatment-induced additional genomic aberrations. The resistance mechanisms can be divided into two main categories: alterations in the targeted driver genes and activation of alternative signaling pathways [17]. The most commonly reported examples of additional mutations in the targeted driver genes are secondary mutations in \textit{EGFR} mutant NSCLC, as well as the so-called ‘gatekeeper’ mutations in \textit{ALK}-rearranged NSCLC cases. The reported alterations associated with activation of alternative signaling pathways include amplification of \textit{MET} or \textit{ERBB2} and mutations in \textit{PIK3CA} and \textit{BRAF}. Similarly, downstream activation of signaling through \textit{MAPK1} amplification was reported in \textit{EGFR} mutant patients. Besides the two main categories, phenotypic changes within the cancer cells such as epithelial-to-mesenchymal transformation (EMT) and transition to small cell lung cancer were also reported to be involved in resistance development [17].

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure1.png}
\caption{Schematic representation of the receptor tyrosine kinase (RTK) signaling pathway. The commonly activated RTK members in lung cancer are EGFR, ALK, ROS1, RET and NTRK1. When activated by mutations or translocations, the RTK will active the downstream pathways including RAS–RAF–MEK–ERK, PI3K–AKT–mTOR and JAK–STAT pathways. These networks will promote cell cycle, enhance survival, drive proliferation, apoptosis, differentiation, angiogenesis, invasion and migration. Tyrosine kinase inhibitors (TKIs) are used to block these pathways. Upon resistance, therapy can be switched to other TKIs. The choice of TKI depends on the underlying resistance mechanism.}
\end{figure}

\section{3. Currently applied molecular diagnostic tests}
Several distinct diagnostic tests are applied to identify the above-mentioned aberrations relevant for treatment [18]. These include fluorescence in situ hybridization (FISH) for chromosomal translocations and amplifications, immunohistochemistry (IHC) for protein overexpression and DNA based next generation sequencing (NGS) approaches to detect single nucleotide variants (SNVs), insertions, deletions (INDELs) and amplifications, as well as RNA-based NGS approaches to detect gene fusions [19,20]. All “one-gene-one-test” assays are relatively time- and tissue-consuming. For small biopsies, such as those from advanced
stage lung cancer patients, the amount of tissue is limited and may preclude subsequent testing of all relevant markers.

DNA-based next generation sequencing approaches to detect SNVs and INDELs are now commonly applied in routine diagnostics. These approaches allow simultaneous screening of multiple genes or gene hotspots. NGS-based techniques in general have a high sensitivity even in small biopsies, but depending on the approach used, minimal tumour cell percentages of around 10-20% are required. These tests have been optimized for suboptimal DNA quality due to the dependency on FFPE material in most clinical settings [21-25]. Target gene capturing and enrichment strategies used for NGS analysis can be categorized as multiplex PCR and hybridization capturing methods [26]. A potential risk of PCR approaches are false positive or false negative results, because it is unknown how many unique DNA copies have been analyzed. This can be overcome by using molecular barcoding. In hybridization capturing enrichment strategies, duplicate reads can be recognized and removed according to the sequence start and end positions or by molecular barcoding. Another problem associated with FFPE material is the formalin fixation procedure, which induces C:G>T:A transitions, and to a lesser extend C:G>A:T transitions [27]. Of note, canonical mutations such as the $\text{EGFR}$ c.2369C>T p.(T790M), $\text{EGFR}$ c.2155G>A p.(G719S), $\text{KRAS}$ c.34G>A p.(G12S) that guide therapeutic decisions for NSCLC patients, are identical to these FFPE induced sequence artifacts.

FISH is routinely used to detect chromosomal breaks and amplifications. For lung cancer, FISH is used to identify $\text{ALK}$, $\text{RET}$ and $\text{ROS1}$ breaks and to detect $\text{ERBB2}$ and $\text{MET}$ amplifications. A limitation of FISH for detection of DNA breaks is that the fusion partner remains unknown. Detection of the fusion partner might be clinically relevant as patients with different $\text{ALK}$ fusion partners have been shown to respond differently to treatment [28]. IHC has been introduced as an alternative for $\text{ALK}$ FISH based on a better association with treatment response to ALK inhibition in lung cancer patients [29]. Similar to the FISH test, ALK IHC also does not allow identification of the fusion gene partner. For RET and ROS1, IHC has not been validated yet in relation to treatment outcome. IHC is also applied to assess expression of PD-L1 protein to select patients for immune checkpoint inhibition therapy. Scoring of both FISH for the detection of chromosomal breaks and IHC for protein expression are subjective to interpretation variation by technicians and pathologists. Moreover, for part of these FISH and IHC-based biomarkers, no international guidelines for scoring are available yet.

More recently, RNA-based NGS methods using bait-capturing based library preparation such as the TruSight RNA Fusion Panel (Illumina, Seattle, USA) or target probe hybridization methods such as the Ovation Fusion Panel (NuGEN, San Carlos, USA) have been evaluated. An alternative RNA-based method based on molecular capturing and counting of fusion transcripts is the NanoString nCounter platform (NanoString Technology, Seattle, USA) [30,31]. Compared with NGS platforms, it has the advantage of direct counting of transcripts without the need of an amplification step. All these methods have been shown to be feasible, to a certain extent, on poor quality RNA samples derived from FFPE tissues.

The above-mentioned developments in the field of molecular diagnosis are especially challenging for lung cancer, where a broad spectrum of diagnostic tests is required to screen for all possible therapeutic targets. In advanced lung cancer, tumour tissue is obtained by biopsies resulting in limited amount of material. This is due to the poorly assessable localizations and potential risks related to the procedures. This represents a challenge for
comprehensively monitoring all therapy-guiding biomarkers. Another potential problem is the marked degree of tumour heterogeneity, in combination with the, in general, single tissue biopsy for diagnostic purposes [32-34]. This might especially be important upon resistant to primary TKI, with potential different resistance mechanisms at different tumour sites.

A potential opportunity to limit the number of required tests on the small tissue biopsies is the use of cell-free DNA from liquid biopsies by droplet digital PCR or targeted NGS approaches. Liquid biopsies from plasma, serum and urine are being explored as alternative sources of tumour-derived molecules, i.e. circulating tumour DNA (ctDNA). The most widely studied liquid biopsy sources are cell-free DNA (cfDNA), extracellular vesicles (EVs) from plasma, and circulating tumour cells (CTCs). The analytical performance of the AVENIO cfDNA assay on the Illumina NextSeq 500 reached sensitivities of 96% to 99% [1]. Liquid biopsy-based tests are less invasive and can be used to identify targetable genomic aberrations, and to monitor disease activity and tumour response by tracking known genomic aberrations over time. This may assist in clinical management of patients by early detection of variants with therapeutic significance or variants showing the presence of resistant subclones [19,35,36]. A potential advantage of liquid biopsies might be that these samples better reflect the true heterogeneity of the tumour in comparison to single biopsies [37].
4. Scope of the thesis

The aim of this thesis is to further develop diagnostic methods to identify genomic aberrations with clinical significance in NSCLC patients. We focused on two distinct groups of patients, those with known drug sensitive and resistant mutations and those that developed resistance to targeted TKI therapies.

In Chapter 2, we tested the feasibility of an all-in-one transcriptome-based assay to simultaneously identify different types of genetic variants with clinical significance in NSCLC patients. The single primer enrichment technology (SPET) was applied to capture all for therapy relevant transcripts by using a PCR-based enrichment approach. We included RNA from cell lines and different tissues types, including frozen samples, cells obtained from pleural effusions and FFPE material, to explore the performance of this assay.

In chapter 3, we analyzed molecular signatures of advanced NSCLC patients using targeted DNA sequencing data available from the routine molecular diagnostics in the UMCG. We re-analyzed anonymized data of the IonTorrent platform, which included a total of >3,000 tissue samples, with data of >1,000 lung cancer samples. We explored the feasibility of using the NGS data to identify amplifications in genes relevant for diagnostics. In addition, we explored whether copy number gains of EGFR were associated with tumour response to EGFR TKI.

In chapter 4 we aimed to identify potential novel crizotinib-induced resistance mechanisms in ALK-break positive NSCLC patients. We applied whole exome sequencing on paired pre- and post TKI tumour tissue samples, to identify genomic aberrations associated with resistance.

In chapter 5, we aimed to explore the diagnostic potential of liquid biopsies. For this study we focused on a cohort of esophageal squamous cell carcinoma (ESCC) patients for which paired normal, tumour and cfDNA samples were available. We tested the feasibility of cell-free DNA (cfDNA) analysis as a tool to predict for residual disease after surgery. We performed targeted sequencing for a cancer hotspot panel including 483 genes, using the Illumina NGS platform on matched normal tissue, tumour tissue, pre-surgery and post-surgery plasma in a cohort of 17 patients. The mutational spectrum between the samples within each patient was compared.

In chapter 6, we summarize the main findings of this thesis, discuss the results and present future perspectives.
In chapter 4 we aimed to identify potential novel crizotinib-induced resistance mechanisms. We analyzed anonymized data of the IonTorrent platform, which included a total of more than 3,000 tissue DNA sequencing data available from the routine molecular diagnostics in the UMCG. We re-visited the feasibility of using the EGFR TKI.

In chapter 3, we analyzed molecular signatures of advanced NSCLC patients using targeted therapy relevant transcripts by using a PCR-based enrichment approach. We included RNA from cell lines and different tissues types, including frozen samples, to explore the performance of this assay.

In Chapter 2, we tested the feasibility of an all-in-one transcriptome-based assay to simultaneously identify different types of genetic variants with clinical significance in NSCLC resistance to targeted TKI therapies.

The aim of this thesis is to further develop diagnostic methods to identify genomic alterations in NSCLC.

### Scope of the thesis

- **General Introduction**
- **References**

### References


Chapter 1


36. Tu, M.; Chia, D.; Wei, F.; Wong, D. Liquid biopsy for detection of actionable oncogenic mutations in human cancers and electric field induced release and measurement liquid biopsy (elb). *Analyst 2016*, 141, 393-402.