Aus der Klinik und Poliklinik für
Herzchirurgie, herzchirurgische Intensivmedizin und Thoraxchirurgie der
Universität zu Köln
Direktor: Universitätsprofessor Dr. med. Th. Wahlers

Detection of circulating tumor DNA in early-stage lung cancer with ultra-deep targeted sequencing

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der Universität zu Köln

vorgelegt von
Luca Kuhlmann
aus Lemgo

promoviert am 16. September 2024

Gedruckt mit Genehmigung der Medizinischen Fakultät der Universität zu Köln

2024

Dekan: Universitätsprofessor Dr. med. G. R. Fink

1. Gutachter: Professor Dr. med. H. Alakus
2. Gutachterin: Privatdozentin Dr. med. L. Nogová

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertationsschrift ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes Unterstützungsleistungen von folgenden Personen erhalten:

Frau Dr. med. Asmae Gassa
Herr Prof. Dr. med. Hakan Alakus

Weitere Personen waren an der Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin/eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertationsschrift stehen.

Die Dissertationsschrift wurde von mir bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die Struktur, die Expertise, die Experimente dieser Arbeit unter der Supervision und nach entsprechender Anleitung durch Frau Dr. med. Asmae Gassa und Prof. Dr. med. Hakan Alakus von mir selbst ausgeführt.

Der dieser Arbeit zugrunde liegenden Datensatz wurde durch meine Mitarbeit erstellt. Die vorherige Doktorandin Frau Sarah Schütten hat mit der Erhebung der Datensätze und Gewinnung der Blutproben begonnen und wurde durch sie und Frau Dr. med. Asmae Gassa eingearbeitet. Die Patientendaten wurden aus histologischen Befunden, Arztbriefen, Operationsberichten und weiteren Untersuchungsbefunden, wie der Lungenfunktionsprüfung, zusammengetragen.

Die Blutproben wurden durch die Arzthelfer*innen abgenommen und innerhalb von zwei Stunden durch meine Person abgeholt und wie beschrieben weiterverarbeitet.
Die Räumlichkeiten und Lagerungen wurden durch die AG Tumorgenomik, die durch Herrn Prof. Dr. med. Hakan Alakus geleitet wird, zur Verfügung gestellt.

Die Aufbereitung von sämtlichen DNA-Extraktionen wurden durch mich durchgeführt. Hierfür wurde ich zunächst durch die medizinisch-technischen Assistentin Frau Anke Wienand-Dorweiler angeleitet.
Das Next Generation Sequencing wurde durch Mitarbeiter*innen des Cologne Center for Genomics durchgeführt und die Ergebnisse durch Herrn Dr. med. Sven Borchmann verifiziert und für die weitere Auswertung zur Verfügung gestellt.
Die Ergebnisse der Next Generation Sequencing und der Patientendaten wurden mittels Excel und Access durch mich ausgewertet.

Die in dieser Arbeit angegebenen Experimente sind nach entsprechender Anleitung durch Frau Dr. Asmae Gassa von mir selbst ausgeführt worden.
Erklärung zur guten wissenschaftlichen Praxis:

Ich erkläre hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten (Amtliche Mitteilung der Universität zu Köln AM 132/2020) der Universität zu Köln gelesen habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Köln, den 16.09.2024

Unterschrift:

Danksagung

Hiermit möchte ich herzlich meinem Doktorvater Prof. Dr. med. Hakan Alakus und meiner Betreuerin Frau Dr. med. Asmae Gassa danken, die mir mit ihrer wissenschaftlichen und klinischen Erfahrung nicht nur bei der Promotion mit neuen Ideen, Anreizen und Erfahrung halfen, sondern mir dieses auch viel für meinen weiteren Berufsweg mitgaben.

Einen nicht in Worte zu fassender Dank geht an meine Familie und insbesondere an meine Eltern, die mich bei jeder Idee und jedem Problem bedingungslos unterstützten.
Meiner Schwester, die mich als kleinen Bruder ertragen musste.

Vielen Dank!

Table of Contents

Abbreviations

Abbreviations	Acronyms
A	Adenine
AMWF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V.
BioMaSOTA	Biologische Material Sammlung zur Optimierung Therapeutischer Ansätze
C	Cytosine
ccfDNA	Circulating cell-free DNA
CNS	Central Nervous System
COPD	Chronic obstroctive pulmonary disease
CRC	Colorectal carcinoma
CT	Computer tomography
CTC	Circulating tumor cell
CTEC	Circulating tumor vascular endothelial cell
ctDNA	Circulating tumor DNA
DNA	Deoxyribonucleic acid
ddPCR	Digital droplet PCR
EDTA	Ethylenediaminetetraacetic acid
EGFR-TKI	EGFR tyrosine kinase inhibitor
EMT	Epithelial-to-mesenchymal transition
exoNa	Exosomal nucleic acid
FFPE	Formalin-fixed paraffin-embedded tissue
FISH	Fluorescence in situ hybridization
G	Guanine
gDNA	Genomic DNA
HE	hematoxylin-eosin
IASLC	International Association for the Study of Lung Cancer
ICC	Immunocytochemistry
i.e.	Latin 'id est' - 'that is', Immunocytochemistry
MET	Mesenchymal-to-epithelial transition
miRNA	MicroRNA
MRI	Magnetic resonance imaging
NGS	Next-Generation Sequencing
NSCLC	Non-small cell lung cancer
PCR	Polymerase chain reaction
PER-CT	Positron emission tomography computed tomography

RNA	Ribonucleic acid
rpm	Rounds per minute
SCLC	Small cell lung cancer
T	Thymine
tDNA	Tumor DNA
TEP	Tumor-educated blood platelets
TMB	Tumor mutational burden
UICC	Union internationale contre le cancer

Kurzfassung der Dissertationsschrift

Detection of circulating tumor DNA in early-stage lung cancer with ultra-deep targeted sequencing

von Luca Kuhlmann

Aus dem Herzzentrum der Universität zu Köln
Klinik und Poliklinik für Herzchirurgie, herzchirurgische Intensivmedizin und Thoraxchirurgie
Direktor: Universitätsprofessor Dr. med. T. Wahlers

Ausführliche deutsche Zusammenfassung:

Insbesondere in der Früherkennung von Lungenkrebs liegt eine große Herausforderung, da Symptome häufig erst im fortgeschrittenen Stadium auftreten und die Erkrankung zu weit fortgeschritten ist, um einen kurativer Therapieansatz zu verfolgen.

Die Liquid Biopsy ist eine neu entwickelte Methode, die ein möglicher Ansatzpunkt für die Krebsfrüherkennung sein könnte. Hierbei werden Tumorzellen oder Tumorzellbestandteile im Blut der Patienten nachgewiesen. Durch neue Sequenziermethoden, wie der digitalen droplet PCR und Next Generation Sequencing, ist es möglich Mutationen in der zirkulierenden zellfreien DNA festzustellen, der sogenannten zirkulierende Tumor DNA. Die Isolation ist aufgrund der Fragilität und der kurzen Halbwertszeit von 2,5 h schwierig. Des Weiteren zeigte sich, dass gefundene Mutationen in der zirkulierenden zellfreien DNA nicht zwangsläufig von dem Tumor stammen. Zellen, die bei der Mitose geschädigt wurden, sterben durch die Apoptose ab und die DNA tritt in die Blutbahn ein. Dieses wurde insbesondere bei Blutzellen beobachtetet. Dieses Phänomen nennt sich Clonal Hematopoiesis und ist insbesondere bei Menschen in höherem Lebensalter oder mit Entzündungen zu beobachten. Die Clonal Hematopoiesis ist eine der größten Herausforderungen der Liquid Biopsy. Um die Problematik der Clonal Hematopoiesis zu lösen, ist es notwendig die mutierten Gene und ihre Eigenschaften zu kennen. Hierbei stellt sich die Frage, ob Mutationen in Patienten mit benignen Erkrankungen nachzuweisen sind, ob nachgewiesene Mutationen aus dem Tumor stammen und ob der Nachweis von zirkulierender Tumor DNA einen Einfluss auf die Rezidivrate und eine kürze Überlebensdauer der Patienten hat.

Um diesen Fragen nachzugehen, wurde 74 Patienten mit NSCLC im Stadium I-IIIB im Zeitraum vom März 2017 bis November 2019 unmittelbar vor der Operation Blut abgenommen. Zusätzlich wurden bei neun Patienten mit benignen Erkrankungen ebenfalls Blut abgenommen und untersucht. Innerhalb von 2 Stunden wurde das Blut zentrifugiert und das gewonnene Plasma und die Interphase bei – 80 ° C eingefroren. Die DNA aus den Leukozyten wurde analysiert und diente als Baseline DNA. Anschließend wurde die zirkulierende zellfreie DNA aus dem Plasma gewonnen und auch durch Next Generation Sequencing und ein Genpanel, welches 85 Gene umfasst, untersucht und mit der Baseline DNA verglichen. Hier konnten Mutationen herausgearbeitet werden, die als zirkulierende Tumor DNA bezeichnet werden.
Um zu überprüfen, ob die gefundenen Mutationen aus dem Tumorstammen wurde 9 Patienten DNA aus Tumorzellen extrahiert, sequenziert und mit den im Plasma gefundenen Mutationen verglichen. Anschließend wurde ein Follow up über zwei Jahre durchgeführt.
Es wurden die Proben von 74 Patienten mit NSCLC untersucht. Hiervon wurden 6 Patenten aus der Studie ausgeschlossen. 32 Patienten befanden sich in Stadium I, 18 in Stadium II, in Stadium IIIA 17 und einer in Stadium IIIB. Insgesamt konnte bei 73% der untersuchten Proben zellfreie Tumor DNA nachgewiesen werden. 43% bei den Patienten in Stadium I, 72% in Stadium II und 88% in Stadium III. Es konnte gezeigt werden, dass das T-Stadium und somit das metabolische Tumorvolumen einen starken Einfluss auf den Nachweis zellfreier Tumor DNA aufweist. Dies war für das N-Stadium nicht möglich. Ein Erklärungsansatz hierfür ist der damit verbundene höhere Zellumsatz, vermehrte Nekrosen und die stärkere Infiltration von Blutgefäßen.
Das Gen *KMT2D* ist das am häufigsten detektierte Gen in unserer Studie. Dieses Gen codiert Methyltransferasen und es besteht eine Interaktion zu *P53*. Mutationen des Gens *TP53* konnten vermehrt bei Rauchern und bei Patienten mit einem Plattenepithelkarzinom festgestellt werden. In frühen Stadien konnte wesentlich weniger nachzuweisen werden (9 % in Stadium I), als bei Patienten im Stadium III (37%). 69% der *TP53* Mutationen, die im Tumor gefundenen wurden, waren im Plasma nachweisbar. Insgesamt 29% der im Tumor gefundenen Veränderungen waren im Plasma detektierbar und 74,6% der im Plasma gefunden Mutationen im Tumor nachweisbar. Dies lässt sich wahrscheinlich dadurch erklären, dass die zirkulierenden Tumor DNA sehr fragil ist und gleichzeitig wurde die DNA nicht aus allen Zellen des Tumors gewonnen, sondern lediglich aus kleinen Abschnitten des Tumors. Diese Ergebnisse untermauern die Annahme, dass sich einzelne Gene nicht zur Tumorfrüherkennung eigenen.
Eine Mutation in *FGFR*-Genen konnte nur in 5 Fällen gezeigt werden. In 3 dieser 5 Fälle kam es zu einem Tumorrezidiv innerhalb der Follow up Zeit. Dies könnte ein Hinweis darauf sein, dass Mutationen in diesem Gen mit einer erhöhten Rezidivrate und einer damit verbundenen kürzeren Überlebensdauer, einhergehen. Das 2-Jahres Follow up konnte bei 76% der

Studienteilnehmer durchgeführt werden. 14% der Patienten ohne zellfreie Tumor DNA-Nachweis erlitten ein Rezidiv und 24% der Patienten mit einem positiven Nachweis. Dies stützt die Ergebnisse vieler Studien, die einen positiven Nachweis mit einer erhöhten Rezidivrate in Verbindung bringen.

Bei zwei der neun Patienten mit benignen Erkrankungen hatten keine Mutationen in der zirkulierenden zellfreien DNA. Insbesondere bei Patienten mit entzündlichen Prozessen war die Anzahl der Mutationen im Plasma überdurchschnittlich hoch. Dieses unterstreicht die Hypothese der Clonal Hematopoieses.

Abschließend lässt sich sagen, dass sich im Rahmen dieser Studie zeigt, dass die Liquid Biopsy wichtige Ansätze für eine individualisierte Therapie liefert, wie der Früherkennung, dem Therapie Monitoring und als prognostischer Marker. Gleichzeitig wurde deutlich, dass der bloße Nachweis oder einzelne Gene nicht ausreichend sind, um sichere Aussage treffen zu können, sondern in weiteren Studien tiefergehend auf verschiedenen Einflussfaktoren eingegangen werden muss.

English summary:

Liquid biopsy is a newly developed method that could be a possibility for the early detection of lung cancer. Tumor cells or tumor cell components are detected in the patient's blood. New sequencing methods make it possible to detect mutations in the circulating cell-free DNA, the circulating tumor DNA. Mutations found in the circulating cell-free DNA do not necessarily originate from the tumor. Cells that have been damaged during mitosis die through apoptosis and the DNA enters the bloodstream. This phenomenon is called clonal hematopoiesis and can be observed particularly in people of advanced age or with inflammation. This raises the question of whether mutations can be detected in patients with benign diseases, whether detected mutations originate from the tumor and whether the detection of circulating tumor DNA has an influence on the recurrence rate and shorter survival time of patients.

To investigate these questions, blood was taken from 74 patients with stage I-IIIB NSCLC in the period from March 2017 to November 2019 immediately before surgery. In addition, blood was also taken from nine patients with benign diseases. The DNA was examined using next generation sequencing.

To check whether the mutations found originated from the tumor, DNA was extracted from tumor cells in 9 patients, sequenced and compared with the mutations found in the plasma.

A two-year follow-up was then carried out.

Samples from 74 patients with NSCLC were examined. Of these, 6 patients were excluded from the study. 32 patients were in stage I, 18 in stage II, 17 in stage IIIA and one in stage IIIB. Overall, cell-free tumor DNA was detected in 73% of the samples examined. 43% of the patients in stage I, 72% in stage II and 88% in stage III.

The *KMT2D* is the most frequently detected gene in our study. Mutations of the *TP53* were detected more frequently in smokers and in patients with squamous cell carcinoma. A total of 29% of mutations found in the tumor were detectable in the plasma and 74.6% of the mutations found in the plasma were detectable in the tumor. This can probably be explained by the fact that the circulating tumor DNA is very fragile and at the same time DNA was not extracted from all cells of the tumor, but only from small sections of the tumor.

A mutation in *FGFR* could only be shown in 5 cases. In 3 of these 5 cases, the tumor recurred within the follow-up period. This could be an indication that mutations in this gene are associated with an increased recurrence rate and an associated shorter survival time. 14% of patients without cell-free tumor DNA detection suffered a recurrence and 24% of patients with positive detection.

In conclusion, it can be said that this study shows that liquid biopsy provides important approaches for individualized therapy. But it became clear that the more detection or individual genes are not sufficient to be able to make a reliable statement. Further studies need to look more closely at various influencing factors.

1. Introduction

1.1. Lung cancer

Lung cancer is an epithelial malignancy that has developed primarily in the lung. It is divided into two major entities: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Additionally, NSCLC can be further subdivided according to histological, genetic, and immunohistochemical parameters. Adenocarcinoma and squamous cell carcinoma are the two main subdivisions of NSCLC. Since the lung is one of the main sites of metastases for numerous types of cancer, anamnestic or pathohistochemical, a distinction must be made as to whether it is a metastasis of another malignancy or a primary lesion. In many cases, this differentiation is not easy to make and requires further analysis. Lung cancer is one of the deadliest and most common cancers worldwide. In Germany, it is the most common cause of cancer death and the second most common type of cancer in men. Prostate carcinoma is diagnosed more frequently in men. In women, lung cancer ranks third behind breast cancer and colorectal cancer in terms of diagnosed cancer cases and second among the most common causes of cancer death. Breast cancer is the most common malignant disease from which most women in Germany die. In figures, this means that there are about 50,000 newly diagnosed lung cancers every year. In 2010/2011, the Robert Koch Institute reported that 33,000 men and 17,000 women were newly diagnosed. The absolute 5-year survival rate is 13% for men and 18% for women. When the mortality of the general population is considered, that is, the relative 5-year survival rate, the percentage increases slightly. It is 15% for men and 20% for women. Although the risk of illness and death is decreasing, the number of new cases increases by 0.5% and deaths by 0.3%. This contradiction can be explained by the structural change in the population structure and the ageing society. The median age of onset of the disease is between 69 and 71 years. This is also the age at which most cases occur. However, if this is considered in relation to the size of the population, the age-specific risk of disease is highest among people over 80 years of age. As the population gets older, the number of cases of lung cancer also increases [1].

Lung cancer has a particularly high mutation rate [2]. This makes them one of the malignant diseases with the highest number of genetic changes of all, further complicating therapy. Therefore, the tumor mutational burden (TMB) is particularly high in this type of tumor. However, targeted drugs can also be used for some of the central oncological driver mutations. For example, specific drugs are available for *EGFR, PD-L1,* or *BRAF* mutations. For this reason, genome-based classification is becoming increasingly important and complements histological classification [1].

1.1.1. Risk factors

Risk factors for lung cancer can basically be divided into endogenous and exogenous factors. Endogenous, that is, genetic, risk is more likely if one or more first-degree relatives suffer from lung cancer. The most important exogenous risk factor is smoking. 90% of lung cancer are related to it. Passive smoking is also an important risk factor. Other risk factors include ionizing radiation, particulate matter, diesel engine exhaust, asbestos, quartz dust, and chronic infections that can cause scar carcinoma [1].

1.1.2. Screening

The time between the first imaging change and the clinical diagnosis of lung cancer varies greatly and is between 1.38 and 3.86 years for computed tomography [3, 4]. For this reason, the AMWF S3 guideline states that people at high risk between 55 and 74 years of age should receive a low-dose annual CT of the thorax if there are risk factors. The risk factors are as follows. Persons between 55 and 74 years of age and having at least 30 pack years or less than 15 years of nicotine withdrawal, or people aged at least 50 years, having 20 pack years or more, and having at least one other risk factor, such as post-lung cancer status, a positive family history, status after other smoking-related malignancies, post-lymphoma status, asbestos exposure, COPD, and/or pulmonary fibrosis [1].

1.1.3. Clinical presentation

Symptomatic, which can occur in lung cancer, can be divided into general, local, and metastatic symptoms. Clinical symptoms often occur only in the metastatic stage, which is why many cases are diagnosed in stage IV. First symptoms are often very unspecific, for this reason they are often not associated with lung cancer and are underestimated.

Local symptoms produced by the tumor include coughing, dyspnea, chest pain, hemoptysis, and bloody-tinted sputum. Vena cava superior syndrome, dysphagia, stridor, hoarseness due to vocal fold paresis caused by nerve recurrence infiltration, arm weakness resulting from brachial plexus infiltration, and Horner syndrome, which causes infiltration of the ganglion stellatum, are less frequent symptoms.

Metastatic symptoms are often painful because the nerves are infiltrated by metastases, and brain metastases can cause dizziness, headaches, and neurological deficits, but also confusion or seizures. In addition, swelling of the supraclavicular lymph nodes or an icterus can occur, which occurs in liver metastases.

General symptoms do not only occur in patients with lung cancer, but can be caused by other malignant or benign diseases. B-symptomatic with weight loss (10% in 6 months), night sweats (sweating through several tops per night) and fever (≥39 ° C) is a general symptom. Physical weakness, autoimmune, endocrine, skin, and hematologic symptoms are also reported [1].

1.1.4. Diagnostics

The diagnosis of new symptoms that raise suspicion of lung cancer should consist of a blood test and a computer tomography of the thorax and upper abdomen. Different biomarkers with poor sensitivity and specificity are measurable in the blood, such as carcinoembryonic antigen (CEA), carbohydrate antigen 19-9 (CA19-9), neurone-specific enolase (NSE), fragment of cytokeratin 19 (CYFRA 21-1), and carbohydrate antigen 125 (CA 125) [5]. If suspicion is confirmed, a biopsy-free bronchoscopy or transthoracic biopsy can be performed. Staging must then follow when the diagnosis of lung cancer is confirmed to allow the evaluation of the spread of the disease and thus the clinical classification. A PET-CT, a cerebral magnetic resonance imaging, and an abdominal CT should be performed. Stages have enormous therapeutic relevance for the patient, as therapy is chosen depending on cTNM. The problem is that metastases to lung cancer can occur almost everywhere in the body. Frequently, space-occupying tumors occur in lymph nodes, ipsilateral or contralateral in the lung, skeletal bones, liver, adrenal glands, and CNS [1].

1.1.5. Tissue biopsy

Currently, the gold standard for tumor tissue samples is transthoracic or bronchoscopic biopsies, which is not possible in approximately 20-30% of all cases due to an unfavourable location of the tumor and poses high risks to the patient at the same time, such as suffering from pneumothorax, bleeding, infection, or pain [6] [7]. Furthermore, biopsies do not represent all tumor heterogeneity. Formalin fixation can show a high level of C>T or G>A transitions in genetic material. In 1-25% of all cases, this phenomenon is absorbable and influences downstream analyses. It has great potential for false positive results and the high costs of a biopsy are also mentionable. For these reasons, a biopsy or cytological tissue is available in only 80% of patients with advanced NSCLC.

1.1.6. Classification

To classify NSCLCs into stages, the IASLC/UICC8 stage classification is used. This is based on the TNM and UICC8 criteria. TNM is the abbreviation for T= tumor size, N= lymph node involvement, and M= metastasis. This classification is of great therapeutic relevance, since up to UICC stage IIIA curative therapy is sought, from IIIB to IVA it is a very individual decision, and in stage IVB palliative therapy of the patient is pursued.

First, a clinical classification is performed based on the results of the staging. After a possible operation a further classification takes place. This is based on the pathological examination. A distinction is made between histological, immunohistological, and molecular biological aspects. Histological categorization has been the most important so far, but is currently being complemented by molecular analyses [8].

In squamous cell carcinoma, adenocarcinoma, large cell tumors, and neuroendocrine tumors are distinguished by histological classification. Neuroendocrine tumors are carcinoid, small cell lung cancer, and large cell neuroendocrine carcinomas. Platy epithelial carcinomas can also be differentiated further into keratinized, non-keratinized, and basal carcinomas. Adenocarcinomas are differentiated according to their invasiveness. They can be para-invasive, minimally invasive, or invasive. In immunohistology, *PD-L1* positive cells are examined, which can further influence the tumor proportion score. Investigation has a great influence on therapy, because target therapy is available. As already mentioned, molecular biological testing using the NGS gene panel is becoming increasingly important, as first-line drug therapies or specific therapy concepts for genetic modification are available [9]. First-line therapies are available for *EGFR* exon 18-21 mutations, *ALK* translations, *ROS1* translocations, *BRAF V600* mutations and *NTRK* fusions, and specific therapeutic concepts exist for *non-V600 BRAF* mutations, *HER2* amplifications, *KRAS* mutations, *cMET* alterations, *NRG* fusions, and *RET* translocations [1].

1.1.7. Therapy

Since the goals and desires of therapy are very individual and thus very diverse, here is presented only a reduced form of all of them. The patient must always be involved in the therapy decision, as he or she has to support it, and this is only possible if the doctor informs the patient well so that he or she can make an autonomous decision for himself. Objectively, the decision is influenced by the clinical stage of the disease and the general condition of the patient, functional lung reserves, and comorbidities. Furthermore, when an operation has been performed and the pathological stage classification has been completed, especially the N classification, the therapy planning has been reviewed and adjusted. In principle, therapy can be derived from the UICC8 or TNM classification. In stage IA and stage IB, surgery is sought and no subsequent adjuvant therapy is required. If there is inoperability, stereotactic radiation should be performed. Patients with stage IIA and IIB should undergo surgery and adjuvant chemotherapy. Stage III consists of a very heterogeneous group. For this reason, stage IIIA is further classified according to Robinson. This classification is based on the distribution of lymph node metastases and whether they are unilevel or multilevel. In the case of stage IIIA with cT3, N1, or T4, N0/1, the standard treatment is the same as for stage IIA/B. If stage IIIA is unilevel, then surgery is performed, followed by adjuvant chemotherapy and subsequent radiation of the mediastinum. At the multilevel stage IIIA definitive radiation chemotherapy is sought. The same procedure is performed in stages IIIB and IIIC. Clarification of curative therapy in stages IIIB and IIIC must also be considered to determine whether surgery can be performed after neoadjuvant therapy. A curative approach should also be pursued in stage IVA if solid organ metastases are present. Then, local therapy of the primary tumor and metastases followed by adjuvant chemotherapy should be performed. Palliative systemic therapy should be

administered at stage IVB, as multiple metastases are present. Once again, it should be emphasised that therapy planning must be done very individually and closely in consultation with the patient. Lung cancer cases should always be discussed in an interdisciplinary tumor conference to use critical considerations of differential therapeutic options as a basis for the physician's recommendations. During recent years, more potential treatments, such as targeted and immunotherapy, have been developed for adjuvant or neoadjuvant therapy. Furthermore, an optimal surgical strategy is still discussed [10] [1].

1.2. Liquid biopsy

In 1948, Mandel and Meatais described circulating cell-free DNA (ccfDNA) [9], but for decades scientists did not know what to do with this information until new sequencing methods were developed [11]. Digital droplet PCR (ddPCR) and next-generation sequencing (NGS) were a breakthrough. Circulating tumor DNA (ctDNA) could be detected in the ccfDNA. In cancer research, the circulation of acids in body fluids attracted attention. The potential to create a minimally invasive diagnostic and prognostic tool by detecting ctDNA in plasma or serum generated great hope [12]. Different components could be found in body fluids for liquid biopsy: circulating tumor DNA, circulating tumor cells (CTC), plasma microRNA, tumor-educated blood platelets (TEPs), circulating tumor vascular endothelial cells (CTECs), and exosomes, which contain DNA, RNA, and proteins [13]. Furthermore, in different body fluids, these components can be verified. It can be detected in blood (plasma and serum), urine, cerebrospinal fluids, and sputum. The most noteworthy case is the capture of cancer-derived material in blood samples. [14]. Liquid biopsy can be used to detect different carriers of DNA or RNA in blood samples.

1.2.1. Structures to be analyzed in liquid biopsy

Liquid biopsy can be used to examine various structures such as circulating tumor cells (CTC), exosome, microRNA, methylation DNA, tumor-associated antigens, tumor-associated autoantibodies, or circulating tumor DNA [15]. The most studied liquid biopsy-derived biomarkers are CTC and ctDNA. During hematogenous spread, CTCs are released into the bloodstream and could be subdivided into three different subgroups: intact CTC, apoptotic CTC, and clusters of CTC. These clusters are also called tumor microemboli [16]. The size range of CTCs is between 8-20 µm, just slightly larger than that of erythrocytes, and CTCs are extremely rare. 1 to 10 CTCs come from 10^6 blood cells. Some CTCs undergo a transition from epithelial to mesenchymal (EMT), become invasive into the bloodstream, go through the mesenchymal to epithelial transition (MET) in secondary organs, and become extravascular. During EMT, cells lose epithelial markers, such as cytocheratin filaments and E-cadherin, and gain mesenchymal markers such as vimentin for MET [17]. In this way, they are involved in metastasis and correlate with poor patient outcome [18] [19]. Furthermore, in patients with

stage IV lung adenocarcinoma, CTC detection is more common than in earlier stages, and counts were higher in stage IV. Only DNA from a few cells is represented in the CTC analyzes, not the entire heterogeneity of the tumor. Furthermore, it is not clear whether CTC is caused by the primary tumor or by metastases. Metastases often have genetic alterations different from the primary lesion.

But CTCs offer the possibility to analyze the morphology, perform functional analysis of cells with FISH or ICC, and perform analysis at subcellular levels. Only CTCs have these possibilities. Furthermore, functional studies in vitro and in vivo, such as patient-derived xerography models, could be used for CTC. Problems are caused by the nonstandardized isolation and recovery of CTCs and the heterogeneity of individual cells, which are very fragile to handle and rare in the blood. In addition, monitoring and identifying drug treatment in real time is one opportunity. It is possible ex vivo (CTC culture) or in vivo (CTC-derived xenograft). In this way, CTCs can be an alternative to patient-derived tissue xenograft for drug screening [20]. For these reasons, they are not well suited to serve as a diagnostic tool. Due to their involvement in metastasis, they are not suitable for early detection [13] [21] [22].

Exosomes are membrane-derived vesicles and contain exosome nucleic acids (exoNA), which means that DNA and RNA are included, but not the whole genome. In exosomes, the genetic material contained could be functionally active and play an important role in the cancer microenvironment. Cancer immunology exosomes are an important part. For example, microRNA is transferred by exosomes and regulates the expression of the target gene [10]. Tumor cells release exosomes actively into the bloodstream, but not only tumor cells. Furthermore, thrombocytes, mast cells, dendritic cells, nerve cells, and astrocytes can be a source of exosomes. However, tumor cells produce a greater number of exosomes than healthy cells [23]. The size range is between 5-150 nm, which is comparable to the size of viruses. Like CTCs, exosomes appear to be involved in metastasis and capable of transferring genes to other cells [24]. Studies showed that plasma exoNA NGS had higher sensitivity compared to plasma ctDNA testing, but only hotspot mutations such as *BRAF, KRAS*, and *EGFR* were tested. As a prognostic factor for longer survival, a low level of exoNA was investigated. Furthermore, exosomes are released actively and not passively as ctDNA. The release of living cells, such as exosomes, might better correspond to tumor dynamics than ctDNA, which apoptotic or necrotic cells release. But exosomes represent only a part of the coding and noncoding DNA / RNA, and not the whole genome. CcfDNA does that, and many genes could be examined in addition [25]. Frydrychowicz et al. showed that lung cancer exosomes could be a useful diagnostic marker and highlight their value in early diagnosis and prediction of the prognosis [26].

CcfDNA contains DNA from apoptotic or necrotic cells. Some of these cells were tumor cells, which released their DNA into the bloodstream. The DNA released includes alterations,

epigenetic information, and can be detected with new sequencing methods such as ddPCR and NGS. Information on the genetic and epigenetic aspects of tumor tissue could be generated [10]. However, it should not be forgotten that circulating DNA is very fragile and has a half-life of 2.5 h, making it difficult to isolate ctDNA. The length of the ctDNA fragments is between 90-150 bp [9]. Other cells also contribute to ccfDNA and contaminate ctDNA probation in this way. Therefore, a wide range of probation of ctDNA in ccfDNA was measured from 0.01% to 90%. It is not yet clear whether apoptotic and necrotic cells are the only way cffDNA enters the bloodstream. For detection of ctDNA in blood only 50 million malignant cells are necessary, and, in contrast, generally no less than 7-10 mm or 1 billion cells are necessary to detect a tumor in PET-CT. Furthermore, the amount of ccfDNA is higher in cancer patients (5-1500 ng x ml $^{-1}$) than in healthy patients (1-5 ng x ml $^{-1}$) [27] [28]. The potential for full representation of spatial and temporal tumor heterogeneity has a liquid biopsy with ctDNA [9]. Additionally, the first FDA-approved liquid biopsy test (Food Drug Association of America) liquid biopsy test for NSCLC has been available since 2016.

In this way, early-stage NSCLC detection and whole-genome mutations can be indicated with ccfDNA analyzed [29] [9].

1.2.2. Possibilities of liquid biopsy

As described above, cell-free DNA is released into the bloodstream not only from tumor cells but also from healthy body cells or misproduced cells. Blood cells play a major role here. This is called clonal hematopoiesis. The personal microenvironment of the patient strongly influences this. It depends on the activities of the patients and on the time of day the samples are collected. This severely deteriorates the validity of the liquid biopsy. Because of that repeating results in quality and quantity becomes extremely difficult because of that. Increased levels of ccfDNA have been shown to be associated with malignant diseases, but in benign diseases such as inflammation, it also occurs. Therefore, no standard is defined for the concentration level of ctDNA / ccfDNA as a cancer biomarker. Another challenge is that quantity is associated with advanced stages of the disease. This is especially true for CTCs but also for ctDNA and exosomes. Together with technical challenges, early detection of lung cancer or relapse becomes difficult. It also reduces the monitoring of response to therapy and tumor heterogeneity [30].

Currently, liquid biopsy is an alternative type of sample when tissue is unavailable. It is a non-invasive tool, easy to repeat, has a shorter turnaround time, and is cheaper than a biopsy. It has the potential to gain information on prognostic, recurrence, and response to therapy. Offers a more comprehensive picture of the disease and real-time monitoring [11] [6] [30]. Enhanced liquid biopsy is an important step to improve drug response and resistance monitoring. An early identification of therapeutic resistance by detecting clonal dynamics and tumor heterogeneity is necessary in personalized targeted therapy. For example, in 60% of patients with *EGFR*

tyrosine kinase inhibitor (*EGFR*-TKI) a secondary mutation and drug resistance appear. Before a relapse is detectable by radiological examination, different *EGFR* mutations can be provable by liquid biopsy. Although there are good isolation procedures available, ccfDNA is very fragile. Intact biological material can be extracted, which is not necessarily given in formalin-fixed tissue. The aim is to take advantage of liquid biopsy in the early detection of lung cancer and to develop a diagnostic test. Right now, the most important advantage is the representation of tumor heterogeneity, which cannot be achieved by biopsy. Oxnard et al. showed that 31% of the patients tested negative for *EGFR T790M* in tissue were positive in plasma. However, on the other hand, around 30% of patients with *T790M* are missed in plasma and have a positive tissue test result. That makes these two approaches a commentary [31]. Liquid biopsy is gentler to the patient and can be repeated as required; it also offers advantages in practical application. The combination of CT and liquid biopsy is promising for screening. But for that goal, a major limitation factor is the lack or very low amount of ctDNA in the early stages. There is also a discussion about the array of genes detected. It must be a wide array, but still focused [9].

The great potential offers this methodology, which must be further improved and promoted to fulfil the great hope of a great cancer screening and a new class of biomarkers, which improve patient outcomes.

1.3. Liquid biopsy of lung cancer

Current lung cancer screening is inadequate and the recommendations of the current AWMF S3 guideline describe that low-dose CT can be administered once a year to asymptomatic people at risk between the age of 55 and 74 years. The reduction in lung cancer mortality can be achieved in radiological screening was described in the New England Journal of Medicine in 2015. In the National Lung Screening Trail (NLST) study, 53,454 people at high risk for lung cancer participated in 33 medical centers from August 2002 to April 2004. One group (26,722) received annual low-dose CT and the other group (26,732) received single-view posteroanterior chest radiography. Data obtained up to December 31, 2009, were correlated with lung cancer cases and deaths. 90% of the probands participated in all screenings, reflecting a high level of adherence. 24.4% of the low-dose CT group had a positive screening and only 6.9% of the radiography group. Furthermore, the tumor stage group where most tumors were found was stage IA and IB, which have a good prognosis. In stage IV there were fewer cases in the low-dose group than in the radiography group. Low-dose CT showed a recurrent reduction in mortality of 20%. On the contrary, mortality in the radiography group could only be reduced by 6.7% and a false positive rate of 96,4%. False positive results lead to unnecessary surgical treatment or invasive examinations. NLST shows that low-dose CT screening can reduce lung cancer mortality [32]. Even more impressive results were generated in the NELSON study. At 10 years of follow-up, lung cancer mortality was reduced by 25% in

high-risk (ex)smokers between 50 and 75 years of age. Additionally, a false positive rate of 64.3% was decreased compared to the NLST study. They achieved this by 3D-based lung nodule management in which the volume of nodules was classified and compared with low-dose CT before [33]. Guibert et al. discuss a combined CT and liquid biopsy screening. They have several ideas that could be possible. The idea is that clinical and blood monitoring alone replace imaging and that if the patient develops symptoms or increases in ctDNA, a CT scan [9] should be performed.

As previously described for colorectal cancer and breast cancer, liquid biopsy can also be used as a diagnostic, productive, or prognostic biomarker for bronchial carcinoma [34]. Since lung cancer is often diagnosed in advanced stages and shows a high degree of heterogeneity within the tumor, which rapidly develops new mutations, this tool is very promising and urgently needed to improve patient outcome. In addition, targeted therapy is becoming more and more important in cancer therapy [8] [9]. This can be seen in a significant increase in survival of advanced NSCLC caused by *EGFR* tyrosine kinase inhibitors and *ALK* receptor tyrosine kinase inhibitors compared to chemotherapy. Drug resistances often occur in patients, leading to a deterioration in health status and would have to be detected by repeated biopsy. 20% of these patients do not have the opportunity to successfully re-biopsied, and heterogeneity is not properly reflected in a biopsy [9]. 79.7% of *EGFR* mutations have been shown to be detectable in plasma, and therefore the FDA has been recommending a liquid biopsy procedure since 2016, which detects *EGFR* mutations. This test is called the 'cobas® *EGFR* Mutations Test v2', based on a ddPCR assay that can detect 42 mutations in exons 18, 19, 20, and 21 for *EGFR*. It was developed exclusively for *EGFR* mutations and contains *T790M* mutations. The test can be used to monitor the effects of the medication and develop resistance. In 117 patients with NSCLC treated with *EGFR*-TKI, 47% of them were identified to have a *T790M* mutation. Patients who had a positive detection of *T790M* ctDNA in the ddPCR analysis had a significantly shorter overall survival than patients with a negative result. Additionally, the test was compared with tumor biopsies. Sensitivity to exon 19 *EGFR* deletions was 81.82% and for exon 21 *L858R* 80%. Therefore, the FDA recommends cobra® if no biopsy is available [29] [9].

Razavi et al. conducted a study of 124 patients with metastatic cancer patients and 47 cancer-free probands. An NGS with a 508 large gene panel was performed in 41 patients with NSCLC, 44 patients with castration-resistant prostate cancer, and 39 patients with progressive metastatic breast cancer. They revealed some very interesting points. They showed that the amount of ccfDNA in the blood correlates with the burden of the disease, the site, and the biological characteristics of the tumor. This includes histology, vascularization, proliferation, and the rate of adaptation. Somatic mutations that accumulate in hematopoietic stem cells and pass them on to their progeny are a form of somatic mosaicism called clonal hematopoiesis

[35] [36]. Clonal hematopoiesis is more common in older patients and therefore increases with age. They are found in 31% of older people. Without knowing the mutations in tumor tissue, the concordance of ccfDNA/ctDNA and white blood cells for de novo characterization could be reported. The broad spectrum of the gene panel led to the discovery of many different mutation variants that were previously unknown, but were of somatic origin and did not originate from the tumor. However, mutations that occurred in tumor tissue were also identified. The five main mutated genes that carry variants matched to WBC for NSCLC were *DNMT3A, TET2, PPM1D, NF1* and *TP53*. Razavi et al. point out that it can be misleading to sequence ccfDNA alone, as clonal hematopoiesis mutations, which affect cancer genes, may be interpreted as tumor-derived mutations. They also showed that a high intensity ccfDNA assay with such a wide spectrum of genes analyzed could provide a more comprehensive landscape of tumor mutation profile than tumor sequencing alone [35].

Most previous studies focus on advanced NSCLC, but Chen et al. analyzed ctDNA from patients in operable stage (stage I-IIIA). Plasma and matching tumor samples were analyzed in 76 patients. Overall concordance was found in 52 of the 76 cases (68.4%) with a panel of 50 NGS gene. 24 patients did not have concordance, but only in 10 cases were discordant mutations in plasma and tissue found. In the other 14 cases, plasma or tissue analysis was unable to be performed. In addition, the stage had a huge impact on the concordance and increased with the advanced stages. In stage I, 57.9% were detectable, 66.7% in stage II, and in stage IIIA 90%. Chen et al. show that multiple mutations in ctDNA can be detected in an early stage of the disease and have potential clinical value [37]. In the TRACERx study, necrosis, lymph vascular invasion, and more aggressive and faster growth are associated with early detection in patients with lung cancer. All these characteristics are combined in nonadenocarcinoma, making it more difficult to detect adenocarcinomas at an early stage. These findings are reflected in 100 patients with TRACERx with early-stage NSCLC, who were analyzed with whole exome sequencing. In 97% of squamous cell carcinoma and 71% of other NSCLC subtypes, they found at least one mutation. In contrast, only 19% of patients with adenocarcinoma. The 10 mutations most frequently found in the context of this study were *TP53* (57.5%), *TTN* (46.8%), *CSMD3* (34.7%), *RYR2* (34.5%), *LRP1B* (29.1%) *ZFHX4* (24.4%), *KRAS* (23.0%), *USH2A* (21.5%), *MUC16* (20.4%), and *RYR3* (19.9%) [38] [39].

Despite this, about 2/3 of patients are diagnosed in an inoperable stage, which often means that no curative therapy can be sought. Therefore, a non-invasive screening procedure is a medical need that does not require harmful ionizing radiation to the patient. Improvements in early detection of relapses and disease monitoring have also been shown to improve the possibilities of targeted therapy and respond to individual development of the disease [6]. This study should contribute to the use of liquid biopsy to achieve these goals and identify biomarkers that have predictive value and could influence the intensity of therapy.

1.4. Liquid biopsy in other entities of tumor

Not only lung cancer patients have high hopes for liquid biopsy. Other tumor entities, such as colorectal and breast cancer, also face major problems in diagnosis, early detection of recurrence, and monitoring of therapy. Liquid biopsy could also be a solution to this problem and help identify new biomarkers. A brief overview of the possibilities that have been created so far for these other tumor entities is described below.

1.4.1. Liquid biopsy of colorectal cancer

As part of the early detection of cancer by public health insurance, you are entitled to an annual stool test between the ages of 50 and 54. The probability of contracting the disease increases with age. Men are therefore recommended to have a colonoscopy at 50 years of age, women at 55. If the findings are inconspicuous, it only needs to be repeated after 10 years. Those who do not wish to have a colonoscopy can have a stool test every year (between 50 and 54) or every 2 years (from 55). Precancerous lesions, such as adenomas, can be removed during colonoscopy. Here, 92.3% of 10-mm polyps are detected and removed for optical coloscopy and 87.5% for 6 mm [40]. But this depends on the experience of the examiner. An American study showed that colon cancer mortality can be reduced by 48% if colonoscopies with removal of adenoma are performed regularly [41]. However, this investigation has risks, such as intestinal perforation, and aims to increase early diagnosis. Approximately 25% of all colon cancer patients already have a lymph node or remote metastases. In the case of an advanced tumor, it is important to be able to make statements about the success of therapy and the prognosis of the patient. There is still potential for improvement here, which could be improved by liquid biopsy.

In patients with colorectal carcinoma (CRC), the level of ccfDNA was higher [9]. On average 25-50 times higher than in healthy test persons. Furthermore, the concentration of ccfDNA was higher in colon cancer than in rectal cancer, which was analysed by several authors. Yang et al. showed that tumor size and cancer stage correlate with an increasing level of ccfDNA [42]. Bettegowda et al. provable clinically relevant *KRAD* gene mutations in 87% of patients with CRC in stage IV and only in 47% of patients with CRC in stage I [43]. 200 patients who underwent a preventive colonoscopy were taken blood samples. They performed a liquid biopsy to detect tumor-associated DNA mutations with a panel of 96 NGS genes. 88% of the patients had wild-type DNA, 6% mutations and 6% indeterminable results. Of the 6% with a positive liquid biopsy result, only 20% had polyps and 80% had a colonoscopy without findings [44]. These results show that liquid biopsy as a diagnostic biomarker for CRC could be a possibility, but there is still a great need for improvement.

Additionally, as a predictive biomarker, a liquid biopsy could be used. Associated with a decrease in plasma ccfDNA level is a response to therapy, and no change or even an increase is found in patients without a response. Cituximab, a monoclonal antibody against *EGFR*,

which is used for cancer immunotherapy, shows a reduction in the level of ccfDNA and *KRAS* from baseline to cycle 3 treatment and an increase in the case of progression [30].

Higher levels of ccfDNA as a prognostic biomarker are associated with a higher risk of recurrence and a reduced overall survival in CRC patients. The monitoring of ccfDNA with shorter survival and disease recurrence was correlated in patients with stage I, II and III. This was also observed by Lecomte et al.. Patients with detectable ctDNA had a 2-year decrease in survival. Another study showed that more than 80% of patients with CRC at the time of recurrence were positive for ctDNA [45].

In summary, it can be said that liquid biopsy for CRC shows a great need for improvement in diagnostics, but as a predictive and prognostic biomarker has achieved promising results.

1.4.2. Liquid biopsy in breast cancer

Currently, mammography is the gold standard as a preventive tool for the early diagnosis of breast cancer. However, several concerns about this golden standard are strongly discussed and generate large debates in specialist circles. 20% of women with breast cancer are under 50 years of age and mammography is not recommended for such young women. Furthermore, all mammography is connected with ionizing rays, which have their own oncogene properties. The experience of the examiner has a great influence on the results, and false positive and false negative diagnoses have a huge impact on patients, because tissue biopsies with risks such as infections, bleeding, and pain are unnecessary, or the diagnosis is made too late. These concerns underscore that mammography is a suboptimal screening method [46].

Zubor et al. has two ideas to improve mammography by combing it with liquid biopsy. The first idea is to perform the screening based on genetic risks, which seems to be more effective than age-based screening. A multilevel diagnostic is the second idea. Mammography and liquid biopsy could be performed in parallel. Blood-based biomarkers for breast cancer screening are not sufficiently sensitive, especially in the early stages, and would be more expensive. Zang et al. observed that for breast cancer, the sensitivity was 74% and the specificity 65.7% for miRNA-30a [47].

As already described in detail for lung cancer and colorectal carcinomas, tumor heterogeneity is also a major problem in breast cancer therapy. In different subtypes of advanced breast cancer, the landscapes of somatic mutations are unknown, but knowing the landscapes allows precise targeted therapy, which is reconciled with the heterogeneity of the tumor. With 100 women with advanced breast cancer, a liquid biopsy was performed with a 25-hotspot gene panel. 96% of the patients had a somatic genomic alteration in the study by Yi et al. Furthermore, the number of somatic mutation members in the endocrine therapy line increased by multivariate regression analyzes. In addition, patients with the target therapy line were positive for fractions of trunk mutations [48].

Another study by Ahlborn et al. showed a lower level of ctDNA in breast cancer than in lung cancer or prostate cancer. The explanation that Ahlborn et al. gives for this fact is that it depends on the type of cancer or / and treatment during blood collection [6].
In conclusion, liquid biopsy is not yet an alternative to mammography, but provides long-term support and takes disease monitoring to a new level. Compared to the expected potential of liquid biopsy in the area of breast cancer and lung cancer, it appears more promising for lung cancer according to the current study situation.

1.5. Sequencing methods available in liquid biopsy

In 1948 ccfDNA was detected in plasma [9]. However, for decades, nothing could be done with this information. Only new sequencing methods have allowed us to examine ccfDNA with more attention. Sanger sequencing, which made DNA sequencing possible in the first place, was replaced by new methods such as ddPCR and NGS.

1.5.1. Sanger sequencing

In 1975, the English biochemist Frederick Sanger invented a chain termination method that made it possible to sequence DNA and find the exact base sequence. Today, Sanger sequencing, also known as the chain termination method or dideoxy method, is widely used as the Maxam-Gilbert method. It is a more advanced form of sequencing; reagents are readily available, and it is automatic. The determination of the sequence is based on four parallel procedures. Four different incubations are carried out, each containing a template primer duplex, a suitable DNA polymerase, an equal amount of the four 5'-phosphate deoxyribonucleosides, and a small amount of the 2',3'-dideoxy derivative of one of the four nucleotide bases. Now, a PCR is performed with all four incubations. If the 2-',3'-hydroxide derivative analog is added to the growing polynucleotide chains instead of a 2-'-deoxynucleotide, the chain extension stops because a 3-'-hydroxyl group is missing. When repeated PCR runs are performed, the 2 -3 -hydroxide derivatives are randomly inserted, and chains of different lengths are created, reflecting all the positions of the bases. The four resulting sets are separated into parallel tracks on a polyacrylamide sequencing gel, producing a specific band pattern. For example, the resulting band patterns can be determined by autoradiography [49].
This method was improved and varied in various ways, but the basic framework remained the same. The applicability and more convenient handling of the vectors in which DNA is sequenced have been improved. Furthermore, the method of labelling the 2',3'-dideoxribonucleotide terminated chains has been changed. 2'-deoxyribonucleotide-5'-triphosphate analogues have been introduced to eliminate the so-called "compression" of the bands on the sequencing gel. The DNA polymerases used were modified, and the ability to sequence both double- and single-stranded DNA matrices was developed.

With this method, Frederick Sanger revolutionized the methodology of genetics about 45 years ago. The Human Genome Project, which was launched in 1990, took 13 years and cost 3 billion US dollars, shows the possibilities that were created. It is the project that employed most laboratories around the world and took international cooperation to a new level. At the same time, it also shows the problems that the Sanger method has. It produces short DNA fragments, is very expensive, takes an extremely long time, and is therefore too labor intensive. By comparison, in 2014, sequencing the human genome using NGS cost less than 1000 US dollars [50].

1.5.2. Digital droplet PCR

Digital droplet PCR (ddPCR) is a modified and further developed form of the PCR method. Basically, it is a PCR divided into 20,000 individual droplets, each droplet containing no DNA molecule or one. Each of these droplets is considered and evaluated individually. A digital result can be generated for each reaction vessel by means of a florescent marker and comparison with wild-type DNA from the tested gene. The most striking difference from PCR is the separation of DNA molecules by boundary dilution and microfluidics. The volume of a droplet is in the femtolitre range [51]. Reaction vessels with successful amplification are proportional to the amount of DNA used. Furthermore, the distribution of DNA molecules follows the Poisson distribution (l= ln(1-p); l= average number of target DNA molecules per replicate reaction; p= fraction of positive end-point reactions), which calculates the number of target DNA fragments per droplet. In each droplet, an amplification with isolated DNA molecules takes place. Using a florescence method, an individual digital result can be determined for each reaction vessel: yes or no amplification of the tested gene. These results have a lower sensitivity to PCR inhibitors and a higher detection limit. High-throughput screening generates a statistically significant result, since about 20,000 reaction vessels are in one square millimeter. The problem that, for example, two positive DNA fragments are in one drop and count as one is compensated by a statistical program.

In the implementation of ddPCR, the production and stability of droplets is a challenge. Monodisperse droplets are generated by a flow-focusing junction at a rate of 1000 droplets per second. In a collection well, the surfactant stabilized droplets flow and form a packed bed above the excess oil because of density differences. This is transferred to a PCR plate and was thermally cycled. After that, the plate is submitted to a droplet reader. Here, they are aligned for single-file simultaneous two-color detection. Each droplet can be positive or negative because each has an intrinsic fluorescence result that allows the detection of negative droplets, despite imperfect querying of the fluorogenic probes. The proportion of positive ones is then offset against the total droplet volume to obtain the absolute concentration of the target sequence. As already mentioned, the Poisson correlation is an important part of getting the

correct result, so that the randomly distributed templates across the droplet partitions can be determined, and the dynamic range extends into the realistic realm.

DdPCR has many advantages over similar procedures. One advantage is that the assay precision can be improved by increasing the number of PCR replicants. In addition, complex mixtures can be examined, which also results from a high tolerance to inhibitors. This is the main advantage of ddPCR. The disadvantage is that one or more specific targets must always be known and that the entire genome can be examined. This means that diagnostics cannot be performed in such a wide range and that target areas are limited compared to NGS.

However, ddPCR is very well suited for the detection of rare mutations and pathogens and for their quantification and determination of the variation in copy number. Also, for the identification of genetically modified organisms. It can also be used in liquid biopsy to perform NGS library quantifications and data confirmations, to analyze specific genes of ccfDNA or ccfDNA for quantification [52].

1.5.3. Next-Generation Sequencing

Next-Generation Sequencing (NGS) emerged in 1994 to 1998 and has been commercially available since 2005. It is also called massive parallel sequencing. This name can be explained by one million to 43 billion short reads, each 50-400 bases long per instrument run. The platforms used are miniaturized and parallelized for sequencing. The technique can be further classified by differentiating between amplicon capture and hybrid capture. This will be described in detail later. First, the basic features of the NGS are described. The NGS process can be divided into four phases: first, the DNA must be prepared, the second phase is the cluster generation, the sequencing is the third, and the last one is the data analysis. The preparation begins with fragmentation into small pieces and denaturation into single strands. Specialized oligonucleotides are ligated with a covalent bond on both sides of DNA fragments. These steps are done simultaneously, which increases stability. Oligonucleotides serve during cluster generation as primers for PCR polymerase and also bind to DNA adapters in flow cells. Now comes a PCR to multiply the DNA fragment. For this purpose, the oligonucleotides bind to DNA fragments that are attached to the top of the flow cell. The polymerase can now work and complete the strand. Once this is done, the double strand can form a bridge with other nucleotides that are attached to the top of the flow cell and the strands are separated again. Another round of PCR continues. After several rounds, two complementary DNA sequences form at the location of each original DNA fragment at the end of each round. One strand of them is cut and rejected. Small islands form, each containing many individual copies of one strand. The third phase is the sequencing. This begins with annealing a primer for DNA-polymerase, so that the polymerase can begin to add nucleotides. These nucleotides are modified with a fluorescently labelled group that can distinguish them. In addition, they have a blocking group, so no addition of other nucleotides is possible. The other nucleotides are

washed away. Now, a laser induces nucleotides to fluoresce, and the color is recorded. Each type of nucleotide fluoresces in a different color, and thus the color can be assigned to a specific nucleotide. At that point, the fluorescence part and the blocking group are removed, and a three-prime hydroxyl group, required to add additional nucleotides, is appended. On the basis of this ability, these nucleotides are called reversible chain-terminating nucleotides. This process is repeated until the DNA fragment is complete. In the last step, the data are analyzed. The colors of the fluorescence sensors are translated into the corresponding base. A computer program detects overlapping sequences and uses this information to reproduce the order of individual fragments. This correlation allows for the reproduction of the entire genome. In addition, it can visualize the various DNA sequences and compare them with a reference DNA/ reference genome to determine alterations [53].

There are different variants of NGS that are based on the basic principles just described. The amphibian and hybrid capture methods have different advantages and disadvantages. In amplicon-based NGS, a primer is first bound to DNA and then PCR is performed. This already results in determined and identical nucleotide strands. The adapters are then ligated and sequencing takes place. In contrast, in capture hybridization-based NGS, a random fragmentation is performed. The adapters are then ligated and only after that PCR is started. A sample is added that binds to the selected nucleotide strands, which are sequenced. This slightly different implementation results in the advantages and disadvantages of the two variants. The on-target rates are higher with the amplicon-based NGS, but the hybrid capture-based NGS performs better uniformity. The sensitivity is lower by the amplicon-based, and that is why the potential for false positive or negative results appears from limited coverage, low variant frequency, vicinity to read starts/ends, or the need for platform-specific variant calling algorithm. In comparison, amplicon-based NGS has a simplified workflow, as it has a shorter preparation time and requires less DNA input, but hybrid capture has a higher test accuracy, which is reflected in greater uniformity of coverage and fewer mismutations. In providing deep uniform coverage and yielding higher sensitivity to variant calling, hybrid capture is superior to amplicon sequencing [9]. For example, Newman et al. showed that the CAPP-seq method, which is hybrid capture-based, is the most sensitive assay. In stage I, they detected ctDNA in 50% of cancer patients and in stages II and III in 100% of patients [54]. Based on the weighing of the advantages and disadvantages, a hybrid capture method was used in this study [55].

NGS is generally inferior to ddPCR in specificity and sensitivity. However, NGS does not require a target gene, and whole genome sequencing can be performed, which would not be possible with ddPCR. Additionally, whole genome sequencing can detect expected oncogenic driver mutations or resistance mechanisms. In this way, new molecular mechanisms of resistance could be explored [9]. In addition, this method is very fast and cost-effective. Therefore, it is used for whole genome sequencing, which is important for the diagnosis of

unknown genetic diseases. NGS is also used for other diagnostic purposes, such as tumor DNA sequencing or liquid biopsy, to detect somatic mutations.

1.6. Questions and aim of the work

Based on the literature described above, these questions arise:

- Is ctDNA detection already successful in early stages lung cancer?
- What characteristics of the extracted free circulating tumour DNA can we identify?
- Are the genetic characteristics of certain groups of patients reflected in the ctDNA and what does this mean for the patient's outcome?
- Can we find indicators of a good or poor prognosis in our patient population?
- Are we able to differentiate certain groups, such as smokers, patients receiving neoadjuvant therapy, or patients who were operated on with suspected lung cancer but had benign disease on histology, based on ctDNA or do they have certain characteristics?
- Whether the mutations found really originated from the tumor or from other cells of the organism and whether all tissue derived DNA mutations are also found in plasma?

2. Materials and Methods

2.1. Patient cohort

In this study, 74 patients with NSCLC and nine patients without cancer were included. Of these 74 patients with lung cancer, 51 (69%) had adenocarcinoma, 22 (30%) squamous cell carcinoma, and one patient had metastases of an NSCLC adenocarcinoma that he had already two years ago.

In all cases, a routine staging was contacted before surgery. Spiral contrast-enhanced computer tomography [3] of the thorax and abdomen and cerebral magnetic resonance imaging (MRI) were performed. Furthermore, lung function was determined by a forced expiratory volume of 1 second. If the clinical tumor stadium was I-IIIB, the patients were included in the study. Patients with tumors of a clinical stage greater than IIIB, neoadjuvant therapy, metastases, or other cancer entities other than NSCLC were excluded from this study. Furthermore, the selection of primary treatment was based on the clinical level of the tumor. The patients underwent primary resection. The standard surgical procedure was a thoracotomy with anatomic pulmonal resection that included a two-field lymphadenectomy of the mediastinal lymph nodes. Before the operation, periphery blood samples were taken. Retrospectively, the histopathological results were analyzed and at least 12 months of follow-up were performed for each case.

2.1.1. Study design

The detection of circulating tumor DNA by NGS and the correlation of results with patient outcomes is a necessary multidisciplinary cooperation.

73 patients with NSCLC were included in Stadium I-IIIA and one patient with IIIB, who underwent surgery between March 2017 and November 2019 at the University Hospital of Cologne, and nine patients without malignant entity. Previously, 30 patients from the cohort collected were examined for a pilot study, which has already been published. Before surgery, all patients were able to give their consent and sign the BioMaSOTA enrolment (Ethics-No. 13-091, BioMaSOTA). Right before surgery, 20 ml of periphery blood was collected in EDTA tubes from the patient. Within two hours, the blood sample had to be processed to avoid becoming unusable due to cellular decay. After centrifugation, plasma and interphase were giant out of the periphery blood samples and biologic material was stored at -80 ° C until DNA isolation followed.

Subsequently, a tumor resection was performed, and the tumor tissue was transferred to the Pathological Institute of the University Hospital of Cologne for histology, genomic DNA sequencing by digital droplet PCR, and pathological TNM classification. The tumors were stored as formalin fixated; paraffin embedded (FFPE) tissue.

Depending on the clinical course, the patients were discharged after one week and on the next course they were treated and examined in the outpatient thoracic surgery clinic or Bethanien Hospital Solingen. For this study, 12 months of follow-up were performed to identify postoperative recurrence. Every six months, Doctor's letters were evaluated, or telephone interviews with the patients' attending general physicians were conducted. The primary target variable was recurrence and the secondary target variable was the survival of the patients.
Parallelly, in Access, a patient database was created with information on pretreatment, type of procedure, histological examination results, lung function parameters, staging results, previous malignant diseases, and follow-up data.
In this way, it should be determined whether preoperative detection of ctDNA is associated with an increased postoperative recurrence rate.

2.1.2. Biobank

In principle, it is extremely important that the biological material is stored properly so as not to jeopardize the success of the experiments carried out. For these reasons, the collection, processing and storage of plasma, interphase, and tumor tissue is an essential step to ensure the adequate performance of the study. The required materials are at least 7 ml of plasma, 1200 µl white blood cells (WBC), and tumor tissue with more than 10% tumor cells. Plasma contains circulating cell-free DNA (ccfDNA). Plasma variant allele frequency is predicted by tumor volume, and only a small amount in ccfDNA of around 0.002% to 3.1% is circulating tumor DNA [38]. WBCs are blood cells that include genomic DNA (gDNA), which serves as a baseline for the interpretation of the sequencing results. This baseline can be compared with the ccfDNA to identify the ctDNA. Furthermore, tumor tissue is necessary to get tumor DNA (tDNA) out of the tissue. Comparison of tDNA and ctDNA is performed to identify mutations stemming from the tumor and/or to determine whether the liquid biopsy reflects a wider proportion of the tumor's heterogeneity.
One to three days before surgery, 20 ml of EDTA (S-Monovetten ®, Sarstedt) blood was taken from the patient in the admission laboratory. Venous blood was obtained using a sterile needle through a puncture of the brachiocephalic vein. Up to two hours after collection, EDTA blood was further processed in the research laboratory. During centrifugation, three layers were created. On the top is plasma (55%), which contains no cells at all, in the middle is a thin white layer, the interphase (4%), which contains lymphocytes and the lowest are erythrocytes (41%). Therefore, the EDTA blood was first centrifuged at 2100 rpm (rounds per minute) for 10 minutes at 4 ° C. The plasma was then pipetted into two Falcon tubes (Tube 15ml®, Sarstedt). This produces 10 to 14 ml of plasma. To remove excess cell debris, Falcon tubes containing the transferred plasma were centrifuged again at 4000 rpm at 4 ° C for 10 minutes together with the EDTA blood from which the plasma was obtained. Plasma was pipetted again into 1.8 ml tubes (CryoPure Tube 1.8 ml yellow®, Sarstedt) and at least 7 ml of plasma should be

generated. A total of 1200 µl interphase (Micro tube 1.5 ml®, Sarstedt) was obtained from the cellular component of the EDTA blood. The remaining cellular supernatant is discarded. Plasma and interphase were stored in appropriate storage containers at -80 ° C until DNA extraction was completed.

Tumor resection was obtained during the operation and is immediately fixed in formalin to avoid falsifying cell decay and pathological examination results. The formalin-fixed tissue was transferred to the Institute of Pathology of the University Hospital of Köln. In pathology, tumor tissue was further processed. Formalin-fixed tissue was embedded in paraffin and processed into FFPE tissue. From this, hematoxylin-eosin (HE) slices were made that were used for histological examination by the pathologist to determine the tumor cell content, the type of tumor, and the classification of TNM. Additionally, gDNA was extracted and examined by ddPCR in the pathology. The gene panel covers 17 genes (*ARAF, BRAF, CTNNB1, DDR2, EGFR, ERBB2/HER2, FGFR2, FGFR3, KEAP1, KRAS, MAP2K1, MET, NFE2L2, NRAS, PIK2CA, PTEN and TP53*) and was designed for genomic analyses of lung cancer. After these examiners, the FFPE tissue and HE slices are stored in the archive of the Institute of Pathology.

2.2. Extraction

DNA extraction is a critical part of liquid biopsy, because sufficient cffDNA yields are necessary for meaningful NGS results, and that is why ccfDNA purification is a bottleneck in liquid biopsy. The ctDNA concentration has to be high enough for NGS feasibility. Qiagen shows a high amount of DNA and read length profiles after sequencing and is well designed for extraction [56].

2.2.1. Extraction of tDNA from FFPE tissue

For the isolation of genomic tumor DNA from FFPE tissue sections, the QIAamp® DNA FFPE tissue kit was used and the QIAamp® DNA tissue kit Handbook of 07/2012 was followed. Especially in FFPE tissues, nucleic acids are often present in fragmented form. Therefore, more care must be taken to avoid further damage to the nucleic acids and to create the best possible basis for subsequent experiments. To obtain tumor DNA from FFPE tissues, HE slices prepared from tumor blocks were examined in the Institute of Pathology of the University Hospital of Cologne, the tumor was marked, and the tumor cell content was determined. Therefore, DNA isolation could begin.

First, eight sections, which were 5 µm thick, were cut and then only marked tumor tissue was carefully scraped with a scalpel. The tissue was placed in a 2 ml microcentrifuge tube, 1 ml of xylene was added to the dissolved paraffin and removed. The mixture was vortexed for 10 seconds and centrifuged for 2 minutes at full speed and room temperature. By pipetting, the supernatant was removed. In this and the next step, it is important that none of the pellets is

pipetted out so that no DNA is released. Now 1 ml of 96-100% ethanol was added, vortexed again for 10 seconds, and at full speed for 2 minutes at room temperature centrifuged. After ethanol is discarded, the pellet is incubated at 37 ° C for 10 minutes or until all the ethanol has evaporated. This is important to avoid endangerment of lysis, which happens in the next steps. Therefore, 180 µl buffer ATL and 20 µl proteinkinase K were resuspended in the pellet. To mix it, it is briefly vortexed and then the mixture is intubated at 56 ° C for one hour to ensure complete lysis, as the conditions are denaturing. After one hour, incubation was continued at 90 ° C for another hour to obtain reserve formalin crosslinks. To remove drops from inside the lid, the tube was briefly centrifuged. Then two similar steps were performed. First, 200 µl AL buffer was added and vortexed, and then 200 µl 96-100% ethanol was added and vortexed to prepare DNA for binding. The microcentrifuge tube was briefly centrifuged to remove drops from inside the lid. To bind DNA, the entire lysate was carefully transferred to the QIAamp® MinElute® column and centrifuged at 6000 g (8000 rpm) for one minute. The collection tube was discarded and replaced with a new one. The DNA was then washed to remove contamination. Carefully pipet 500 µl Buffer AW1 on top of the QIAamp® MinElute® column and centrifuge at 6000 g (8000 rpm) and discard the collection tube. This step was repeated exactly as described above with buffer AW2. The washing has now been completed, and drying follows. To dry QIAamp® MinElute®, it was centrifuged at room temperature for 3 minutes at full speed. Finally, DNA bound to the QIAamp® MinElute® column must be dissolved to obtain it. The QIAamp® MinElute® column was placed in a 1.5 ml sterile microcentrifuge tube. On the membrane 20 µl Buffer ATE is applied, which serves as an eluate. This was incubated for 5 minutes and then centrifuged at full speed for 1 minute. The eluate obtained is frozen at -20°C until NGS is performed.

2.2.2. Extraction of ccfDNA from plasma

For ccfDNA extraction, QIAamp® MinElute® ccfDNA was used and the QIAamp® MinElute® ccfDNA Handbook protocol of 08/2018 was followed for DNA purification. DNA fragments of <100 bp are present in human plasma, and fragments as short as 20 bases of nucleic acids can bind to QIAamp® [9]. To avoid further destruction of the fragments, the plasma should not be frozen more than once prior to extraction. Furthermore, the concentration of ccfDNA is typically 1 to 100 ng / ml in human plasma and for NGS it requires a concentration of 50-100 ng/ml and at last 2 µg. In this study, 7 ml of plasma was used to increase the concentration of ccfDNA and ctDNA and only 40 µl of ultra-clean water to elute the sample. The elution was stored at -20 ° C until NGS was performed.

By a high-force g-force centrifuge, plasma was prepared to remove cellular debris and to reduce gDNA and RNA. Plasma was stored at -80 ° C. At the beginning of extraction, 1050 µl Bead Binding Buffer, 385 µl Proteinase K and 210 µl Magnetic Bead Suspension were added to 7 ml of plasma, 385 l of proteins K and 210 l of magnetic beads suspension. Magnetic Bead

Suspension contains magnetic beads that bound ccfDNA during 10 minutes of end-to-end rotation. Proteinase K generates the complete release of nucleic acids from the bound proteins and lysed the sample. After 10 min of incubation by end-to-end rotation, the samples were briefly spun (30 s at 200 x g), so that no solution remained in the cap. During the next step, magnetic beads were collected on a magnetic rack, where the tubes were placed for at least one minute. The supernatant was discarded. To remove bounded DNA from the magnetic beads, 200 l of beads 200 µl of beads was added, vortexed and pipetted up and down to remove residual beads from the wall. After transferring the solution to beadsolution evaporation tubes, the mixture was incubated on a shaker for 5 min at room temperature and 300 rpm. During that time, DNA was released from the magnetic beads, and in the next step, the beads and the solution were separated by a magnetic rack until the solution was clear. The clear solution was pipetted into a new beadsolution buffer and the beads (with no DNA bounded) were rejected. To allow optimal binding of ccfDNA to the QIAamp® Mini Elute membrane, 300 µl of Buffer ACB was added and vortexed. The mixture from the last step was then transferred onto a QIAamp® Mini Elute membrane. Circulating DNA was adsorbed onto a silica membrane because the solution was spun at 6000 x g for 1 minute in a centrifuge. Contaminations cannot bind to the membrane. The salt and Ph conditions prevent it. In this way, the downstream enzymatic reaction cannot be inhibited. The following QIAamp® UCP Min Elute column was placed in a new 2 ml collection tube and the flow was discarded. 500 µl of ACW2 was added to the membrane and centrifuged at 6000 xg for 1 minute. To achieve the removal of residual contaminations. In this step, contaminants were washed away efficiently, and nucleic acids remain bound to the membrane. To dry the membrane, it was spun at full speed for 3 minutes and then for 3 minutes at 56 ° C and incubated in a shaker with the lid open. This ensures that the residual ethanol evaporates and is not carried over. The last step was to carefully pipet 40 µl of ultra-clean water into the membrane as an elution. The ultra-clean water was then incubated for 1 minute at room temperature and centrifuged at full speed for 1 minute. Directly after the last step, the solution was frozen at -20 ° C.

2.2.3. WBC gDNA extraction

The interphase is the upper layer of the cellular part, which was created during centrifugation because leukocytes are the lightest cells in blood. It retains lymphocytes, which represent immune cells, and are the only blood cells that contain genomic DNA. To detect mutations, gDNA is needed to compare it with ccfDNA and tDNA. For the extraction of gDNA from leukocytes and blood, we used QIAamp® DNA Mini and Blood Mini and followed the Handbook 05/2016. A mixture of 200 µl interphase and blood per patient was used as a sample. On average 200 µl of blood contains 6 µg DNA or between the values of 4-12 µg. Additionally, more DNA is included in the 5x106 lymphocytes included. On average 50 µg DNA are detectable in 200 µl. Blood treated with citrate, heparin, or EDTA QIAamp® DNA Mini and

Blood Mini is suitable for use. In this study, blood samples were collected in EDTA tubes and frozen at -80 ° C. In the end, 50 µl Buffer AE buffer was used as elution.

During the first extraction step, 20 l of extraction, 20 µl QIAGEN protease, was added to the 200 µl interphase sample. The protease lyses the sample and records a complete release of DNA, which was bound to proteins. Free of DNase or RNase activity, QIAGEN protein does not produce a loss of valuable and required gDNA. 200 µl Buffer AL was added and at least vortexed for 15 seconds. The AL buffer supports enzymes of the QIAGEN protease. Subsequently, the mixture incubates at 56 ° C for 10 minutes, so protease can fully perform its function. After incubation, 200 µl ethanol was pipetted into the sample and mixed for 15 seconds by pulse vortexing. To purify DNA, we carefully applied the solution to the QIAamp Mini spin column, which includes a membrane. This membrane has optimal conditions for binding to DNA due to the Ph value and the silica that binds to the membrane. To achieve that, the tube was centrifuged at 6000 xg for 1 minute. Furthermore, the conditions that rule the membrane ensure that no proteins or other contaminants are able to bind. Thus, downstream enzymatic reactions cannot be inhibited. If the membrane was nevertheless contaminated, it was washed twice in the next two steps. First, 500 µl Buffer AW1 was added and centrifuged at 6000 xg for 1 minute. In the next step, 500 µl Buffer AW2 was carefully transferred to the membrane, centrifuged at full speed for 3 minutes and washed a second time. Thus, residual contaminates were removed without affecting DNA binding. The purpose was to detach the DNA from the membrane and perform an elution for that reason. The elution was carried out with 50 µl of Buffer AE. The AE buffer was stored at 5-10 ° C and should first be brought to room temperature (15-20 °C), so that this step works as desired. Furthermore, to increase the level of DNA, the membrane was incubated in AE buffer for 5 minutes before centrifugation. The length of eluted gDNA is predominant 20-30 kb, but the longest fragments are up to 50 kb in the elution. After that, storage is recommended at -30 to -15 ° C. In this study, we freeze the samples at -20 ° C until the sequencing was performed.

2.3. NGS of gDNA, tDNA, and ccfDNA

NGS was performed at the Cologne Centre for Genomics. Using Illumina sequencing platforms (MiSeq or NextSeq; Illumina, San Diego, CA, USA). For target enrichment, the TruRisk gene panel based on custom hybridization capture was used for target enrichment (manufactured by Agilent, Santa Clara, CA, USA; or Illumina). An overview of the targeted genes is found in Figure 1. In European Molecular Genetics Quality Network (EMQN) schemes, the diagnostic pipelines of the laboratories involved have been successfully tested. The genes analyzed here are based on the published gene panels used and successfully detected on cBioPortal. NGS-based data with predictions of large genomic rearrangements are known to produce false positive results and, therefore, validation is required. Furthermore, intrinsic and synonymous mutations were excluded from our study. Also, all SNPs located 5 bases around an InDel were

excluded because SNP calling around InDels proved to be difficult. In the *CNTN1* gene and in the *UTR3* region artefacts suspicious anomalies were found, which were present with low allele frequency and partially clustered. Because nothing was known about these mutations in the literature and their combination with the low allele frequency, the suspicion of the artefact is very close and they were also excluded.

Figure 1: Listing of all genes tested by NGS:

ARAF, BRAF, CTNNB1, DDR2, EGFR, ERBB2, FGFR2, FGFR3, KEAP1, KRAS, MAP2K1, MET, NFE2L2, NRAS, PIK3CA, PTEN, STK11, NF1, SETD2, RBM10, MAG, JAK2, ARID1A, SMARCA4, RB1, CDKN2A, U2AF1, RIT1, ALK, ROS1, RET, ATM, DIS3, ERCC2, FANCG, FANCA, FLT3, NBN, NRP2, TP53, BRIP1, BRCA2, POLQ, NKX2-1, PTPRD, HGF, EPHA3, PPP3CA, KDM6A, NOTCH1, HRAS, RAF1, AKT1, DNMT3A, USH2A, NID1, PLB1, MAP4K3, MSH2, AFF3, KCNH7, SOX2, SMAD4, TERT, MDM2, FBXW7, FOXP1, CUL3, MPPED1, ARFGEF2, BPIFB1, NLRP4, LTBP4, STAT5B, NOS1, CNTN1, HLA-A, APC, CREBBP, TP63, BAI3, CDKN2A, FGFR1, CDK4.

2.4. Characteristics of the patients

In the whole study, 83 patients were examined. 74 had lung cancer in an operable stadium and nine had no cancer during blood collection. The most common cancer type was adenocarcinoma with 50 cases and the second squamous cell carcinoma with 22 cases. One patient had metastases, and another had stage IIIB. After the UICC stage, all included cases were within I-IIIB and received an R0 resection. Furthermore, we separated patients with performed neoadjuvant therapy and therapy naive patients. For these reasons, patients who did not meet these criteria were excluded from the study. Before surgery, all patients received a routine staging consisting of computer tomography of the thorax and abdomen, cerebral magnetic resonance imaging (MRI), a check of lung function, and laboratory blood results. No one of the included patients was in a metastatic stage.

Figure 2 provides an overview of the clinical facts of the lung cancer cohort of patients.

Figure 2: Baseline characteristic of the patient of all 74 cancer patients

Patient cohort. In this figure, patients with resectable NSCLC of both histopathological subtypes (adenocarcinoma and squamous cell carcinoma) and patients treated with neoadjuvant were included. The patient's blood was sampled prior to surgery.

Characteristics	Value
Females	35 (47%)

Age (median and range) years	68.5 [34; 86]
Smoking history	
Current smoker	32 (43%)
Former smoker	24 (33%)
Never smoke.	8 (10%)
Pathological stage (TNM 8th ed)	
IA	26 (35%)
IB	8 (12%)
IIA	4 (6%)
IIB	14 (19%)
IIIA	16 (22%)
>IIIA	4 (6%)
Unclear result	2 (3%)
T Stadium	
T1	35 (47%)
T2	14 (19%)
T3	14 (19%)
T4	10 (14%)
N-Stadium	
N0	52 (71%)
N1	10 (14%)
N2	11 (15%)
N3	0 (0%)
Histology	
Adenocarcinoma	51 (69%)
Squamous cell carcinoma	22 (30%)
Metastases	1 (1%)
Tumor in history	
Yes	14 (19%)
No	60 (81%)

2.5. Follow-up

To ensure sustainable verification of the clinical study and scientific investigation, a follow-up was performed with study patients. Follow-up meets the holistic demand and creates a practical approach to daily hospital life. To obtain information on circulating tumor DNA as a possible biomarker, clinical follow-up is necessary. In this way, the long-term effects on test subjects can be verified.

Most of the patients received their aftercare at Bethanien Hospital Solingen, the Department of Internal Medicine III of the University Hospital Cologne or the Outpatient Thoracic Surgery Clinic of the University Hospital Cologne. At Bethanien Hospital and the Department of Internal Medicine III, computer tomography of the thorax and abdomen was performed every 3 months for 2 years, possible adjuvant (radio-) chemotherapy was administered, and holistic care of the patients was sought. On the contrary, in the outpatient thoracic surgery clinic, surgical follow-up was performed about two weeks after surgery. This was primarily aimed at controlling wound healing, pain therapy, and breathing difficulties that could arise, for example, from pleural effusions. A written declaration of consent of the patients is essential for the exchange

of data between Bethanien Hospital and the University Hospital of Cologne. This is not only necessary for this study, but also for normal patient therapy. Every six months, follow-up was performed and, if the situation allows, for more than 24 months. In general, medical letters were evaluated for this purpose and made available to the study. The target variables or the so-called dependent variables, which were in the foreground of the evaluation, were defined as follows. Relapse was used as the primary variable and patient survival was considered the secondary variable. Disturbing factors, also called independent variables, are factors that influence the dependent variable and affect the dependent variable. This presents a challenge that makes it difficult to obtain information. Some patients discontinue therapy and change their appointments in clinics, move to a different location, live too far away to receive follow-up care in the mentioned clinics, or die for other reasons, such as comorbidities, to name some examples of these external influences. For these reasons, follow-up was possible in 77% of cases over 24 months. To keep losses as low as possible, patients who were not cared for at Bethanien Hospital or the Department of Internal Medicine III were information collected by telephone interviews with general physicians. These telephone conversations were about regular tumor aftercare in a clinic or a resident oncologist, about suffering a relapse, survival of patients, or other physical complaints that may arise from surgery, such as chronic wound pain.

2.6. Database

In the period from 03.02.17 to 08.01.20, samples from 200 patients were collected in this study. 83 of these patients were examined for circulating tumor DNA in their plasma. In addition, a database was created about all cases. Access is the program in which the database was created. To collect the data, Orbis doctors' letters were used. The database was divided into five categories, which were called "Master List", "Diagnosis", "Pathology", "Post-OP", and "Therapy". Patients can be assigned cases using an identification number. However, this list with the name and ID number was always stored separately from the database and the password was protected. Therefore, this information cannot be taken together.

The content of the Master List is data about the general case, such as date of birth, age, sex, height, weight, referring clinic, and whether they had signed the BioMaSOTA enrolment. Information about the diagnostic performed was collected in the “Diagnosis” table. For example, information about the performance of computer tomography of the thorax and abdomen, cerebral magnetic resonance imaging (MRI), lung function check-up, positron emission tomography, cTNM, UICC stage, preexisting malignant conditions, and what kind of it. Furthermore, in "Pathology" information was collected on cancer subtype, pTNM, UICC stage, R0 situation, tumor size, degree of malignancy, percentage of *PD-L1* positive cells, Cologne score, and Ki67 expression. Data on neoadjuvant therapy, date, and type of operation were collected in the 'Therapy.' In the last table, postoperative information was collected on

the postoperative course, such as adjuvant therapy, the date of the last follow-up, relapse and location, and patient survival.

It was not always possible to collect all data from the patients. Sometimes important information, such as follow-up data, was not possible to collect.

2.7. Ethic Votum

The study was approved by the Institutional Ethics Committee of Cologne University Hospital (Ethics-No. 13-091, BioMaSOTA). All patients signed up for the study prior to surgery. Furthermore, all experiments were performed according to relevant guidelines and regulations.

4. Results

3.1. Demographic information

Of the 74 cancer patients, 51 (69%) had adenocarcinoma, 22 (30%) squamous cell carcinoma, and one patient had metastasis of an NSCLC adenocarcinoma that he had already had two years ago. Tumor tissues from nine tumor patients were examined using the same gene channel. Additional ccfDNA was sequenced from nine patients who suffered benign diseases (anthro histiocytosis, lymphoid interstitial pneumonia, necrotizing chronic granulating fibrosing inflammation, granulomatous inflammation, lipomatous tissue, schwannoma, tissue excidate, haemorrhage, and fibroadipose connective tissue) was sequenced. Only one patient was excluded because his neoplasia was a metastasis. Five patients were treated neoadjuvantly and excluded. Therefore, they were analyzed separately. 32 patients were in stage I, 18 in stage II, 17 patients in stage IIIA, and one patient in stage IIIB. All patients received surgical treatment that achieved resection R0 (Figure 3). The age distribution at the time of surgery can be clearly seen in Figure 4. On average, the patients were 68.5 years old. 34 years was the youngest patient, and the oldest 86 years. 39 patients (53%) were men and 35 women (47%). The smoking status was recorded. This information could not be determined for 12% of the patients. 41% were smokers, 35% had smoked in the past and 12% had never smoked. Among patients with squamous cell carcinoma, all had a history of smoking (for one patient, no information was determinable). An overview of smoking behavior, classification by stage of UICC and tumor histology is shown in Table 2. Seven out of nine patients with benign disease had a smoking history (2 smokers, 5 ex-smokers). In the past, 14 of all patients (17%) had already been treated for malignant disease.

A follow-up was performed for two years. In 76% of the cases, complete 2-year follow-up was obtained. Adjuvant chemotherapy, radiotherapy, and / or immunotherapy was performed in 39%. In 54%, the patient did not recommend or reject adjuvant therapy. Unfortunately, no information could be obtained for 7%. The occurrence of recurrences was documented. In total, 11 patients suffer recurrence. 2 patients relapsed in half a year, 3 others within 6-12 months, another within 12 to 18 months, and 5 others within 18 to 24 months. In this period, the survival of the patients was also observed. It shows that seven patients died within 24 months after follow-up. Four died in the perioperative process in the first six months after surgery. Three more patients in the first year. Only two patients had relapsed and died during the follow-up time.

Figure 3: Study patients' subdivision

Workflow of the study with the respective sample numbers for each analysis step comparing of positive or negative detections of ctDNA in patients with proven lung cancer and patients without malignant disease.

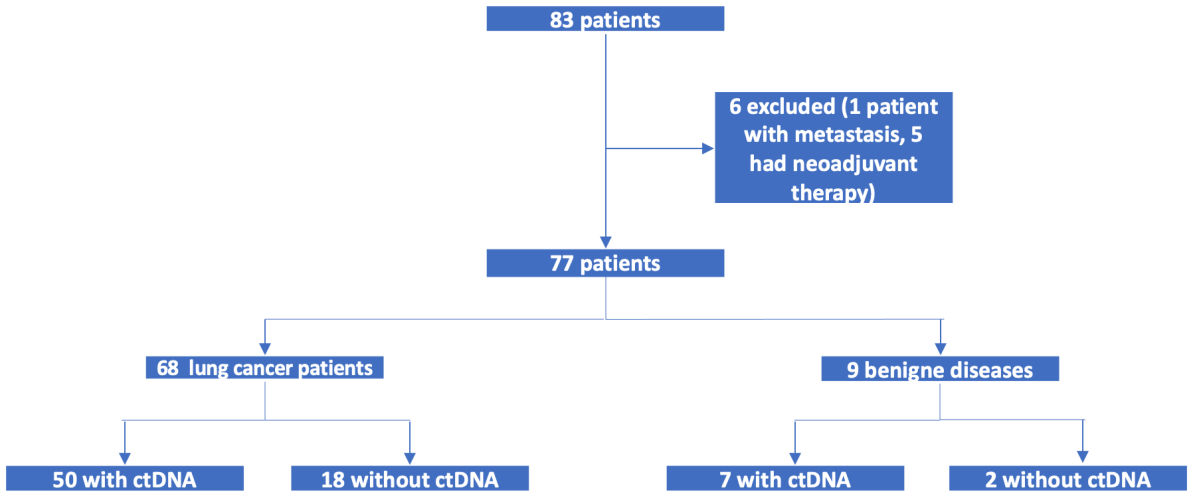
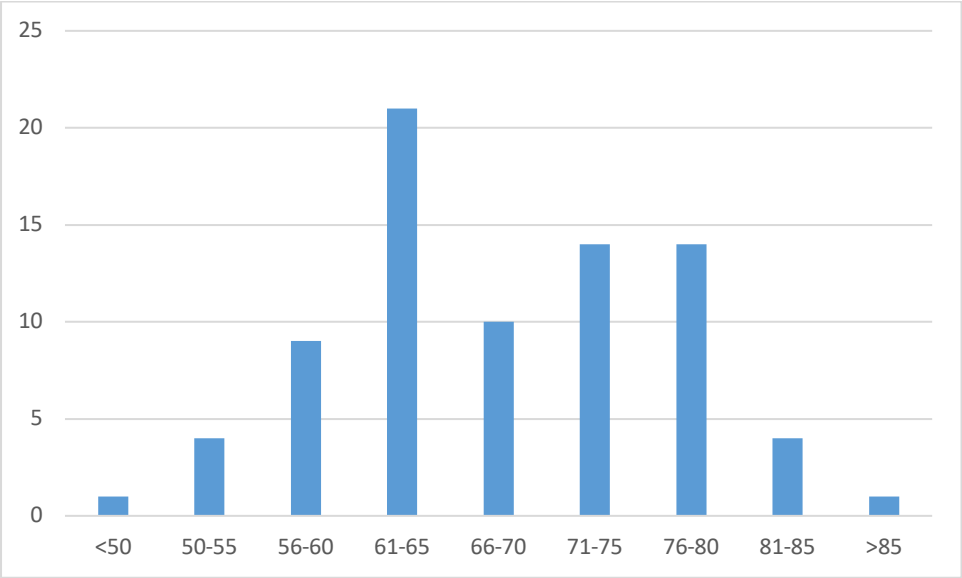


Figure 4: Patients age during surgery

Age distribution of operated patients. The largest group consists of patients between 61 and 65 years of age. The case in which the patient was under 50 years of age stands out. This case is also included.



3.2. De novo detection of tumor-derived ctDNA mutations

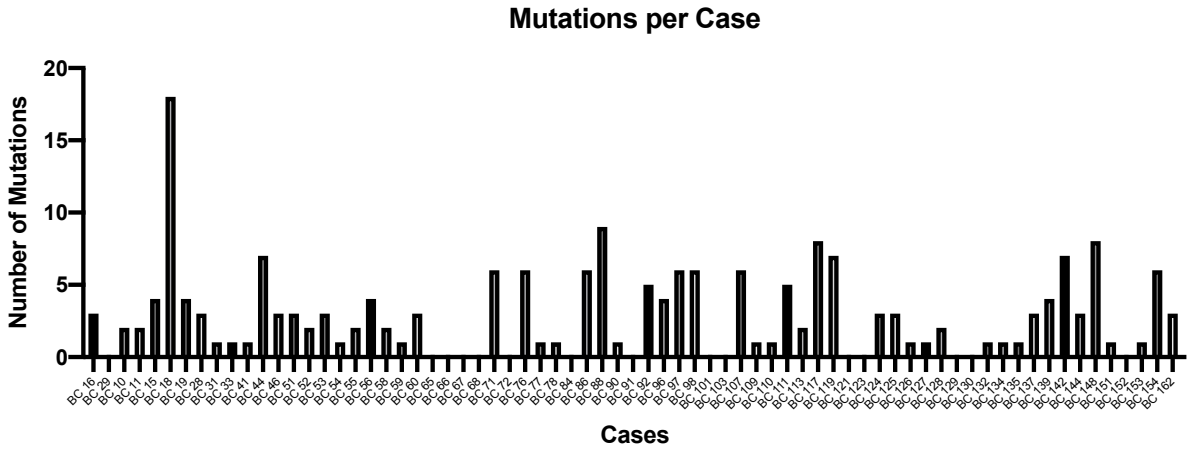
In 73.5% (50/68) of the cases, ctDNA was detectable. 26.5 % of the cases remained without detection. *KMT2D* is the gene most frequently detected. In 18 cases, which is 26.5% of all patients, *KMT2D* was identifiable by mutation. The second gene detected the most frequently is *TP53* with 17.5% (n=12). *CREBBP* (n=8; 12%) and *LTBP4* (n=6; 9%) follow. These results refer to patients with NSCLC. Seven genes were found five times (7%). *FANCA, KEAP1, NF1, NOTCH1, POLQ*, and *USH2A* are these genes. It was possible to detect *ARID1A, DNMT3A, FOXP1*, and *PTPRD* four times (6%) and *APC, BRCA2, ERCC2*, and *HGF* three times (4.5%). We detected 2.58 mutations per lung cancer patient using this method (Figure 5).

Based on tumor histology, the differences between ADC and SCC can be shown. In patients with ADC, ctDNA was detected in 72% (34/47) and 2.14 mutations per patient. Especially a patient with 17 mutations increases the average. The patient with the second most mutations has only eight. *KMT2D* was the mutated gene most often found in patients with ADC. 26% (n=12) of all patients with ADC had a mutation in this gene. In 11% (n=5) NF1 and *TP53* could be detected. Compared to the total collective, there are 7% fewer detections in patients with ADC. *LTBP4* and *POLQ* were shown in 9% of the cases (n=4).

In patients with SCC, more mutations were detectable. Here, 3.00 mutations per patient were found. Compared to ADC, this is on average nearly one more mutation per patient. 76% (16/21) of the SCC cases had a ctDNA finding. About 4% more than in patients with ADC. *TP53* is the gene most often found in patients with SCC. In 32% (n=7) of the cases, a mutation was proven, which is 14% more than in the total. *KMT2D* is the second most common here, with 27% (n=6). *CREBBP* is the third in 23% (n=5) of the cases. *CREBBP* was found only eight times in total and five times in patients with SCC. This shows that 67% of all *CREBBP* mutations are found in SCC patients, and SCC patients represent only 31% of the cohort. *ARID1A, KEAP1* and *NOTCH1* could be detected three times in patients with SCC.

Figure 5: Mutations per Case

Somatic mutation landscape ctDNA samples. The bars represent the mutational number of each sample.



3.3. Comparison with the UICC stage

The study was carried out in patients with early-stage lung cancer. 32 patients in stage I, 18 in stage II, 17 in stage IIIA, and one patient in stage IIIB were included. The proportion of patients in stage I is 48%. In stage I, it was possible to detect ctDNA in 68% (n=21) of the cases. On average, 1.47 mutations per patient were detected, which is below the overall average of 2.58 mutations per patient. Furthermore, the maximum number of mutations per patient was six. In 26% (n=8) *KMT2D* could be detected. This is the gene that is found most frequently. Four genes were detected three times each in stage I. *CREBBP, DNMT3A, POLQ*, and *TP53* were detected in 9.5% of the cases. DNMT3A was found only four times in total. Three-quarters of all detections occurred in stage I. The detection rate in stage II is 4% higher than in stage I and is 72% (13 out of 18). Mutations per patient are also higher, with 3.61. A patient had 17 mutations that increased this value. This case stood out. The patient with the second most mutations had only eight. If this patient is excluded from the calculation, the average number of mutations per patient is 2.82. This would be 1.35 more mutations per patient compared to stage I or 2.14 if the case is included. Stage II differs in terms of the gene that is detected most frequently. *KEAP1* was found to have five mutations and was the most common gene. This is a share of 28% and at the same time, *KEAP1* was only found in stage II. *FGFR1-3, LTBP4*, and *TP53* are the second most common genes and were found in 22% of cases (n=4). 80% of *FGFR1-3* (4 out of 5) detections occur at this stage. *KMT2D* and *PTPRD* were found three times. *KMT2D* was not the most common gene in stage II patients. In stage IIIA, 15 of 17 patients had positive detection of ctDNA. This corresponds to 88% of all patients. Furthermore, 3.18 mutations per patient were observed. The average number of mutations per patient would decrease from 3.61 to 3.18 from stage II to stage IIIA. It should be mentioned here that in stage II a case with 18 mutations was included in the calculation. In stage IIIA, the highest number of mutated genes per patient was eight. Six of the 15 patients with positive detection of ctDNA in stage IIIA had between six and eight mutations. In the other stages, a large proportion of

high mutations per patient could not be achieved. In the 17 patients in stage IIIA, *KMT2D* was detected most frequently. With a rate of 35% (n=6). *TP53* is second with a detection rate of 29% (n=5) and is followed by *CREBB4* with four mutations found and a rate of 23%. The development of *KMT2D* and *TP53* should be considered. The detection rate of both genes increased from stage I to stage III. *KMT2D* was found in stage I in 26% of the patients. In stage III in 35% of the cases. 9.5% of the cases in stage I had a positive *TP53* result, which increased in stage II with 22% and in stage III with 29%.

3.4. Correlation of the T-stage and the N-stage

To look at the stage of TNM in a more differentiated way, patients are based on tumor size, that is, into stages T 1, 2, and 3, and then into the lymph node stage, N0, N1 and N2.

Regarding the number of mutations per patient, a clear increase can be observed from T1 to T3. Patients in stage T1 have an average of 1.44 mutations and in stage T2 3.08. This is an increase of 1.64 mutations per patient. Furthermore, between T2 and T3 there was an increase in mutations with 1.13. From 3.08 mutations per patient in stage T2 to 4.21 in patients with T3. An overview of the genes found is in Table 3. In 8 cases, 23.5%, in the 34 T1 patients *KMT2D* could be detected. In 9% (n = 3), *TP53, POLQ*, and *LTDP4* could be identified. In patients with a T2 tumor, *HGF, KMT2D*, and *PTPRD* are genes that were detectable three times and were shown in 23% of cases. Tumors in stage T3 were found in 20 patients. Both *KMT2D* and *TP53* were detectable in 37% (n=7) of these patients. For TP53, this is an increase of 28% compared to T1. *FGFR1-3* were detectable only in T3. The five mutations were detected here and therefore were conspicuous in 26% of the cases. *CREBB4* was also more detectable in T3 than in earlier stages. In 21% (n=4) a mutation of this gene was detectable. *NOTCH1, APC*, *ARID1A,* and *KEAP1* were also found in 16% (n=3).

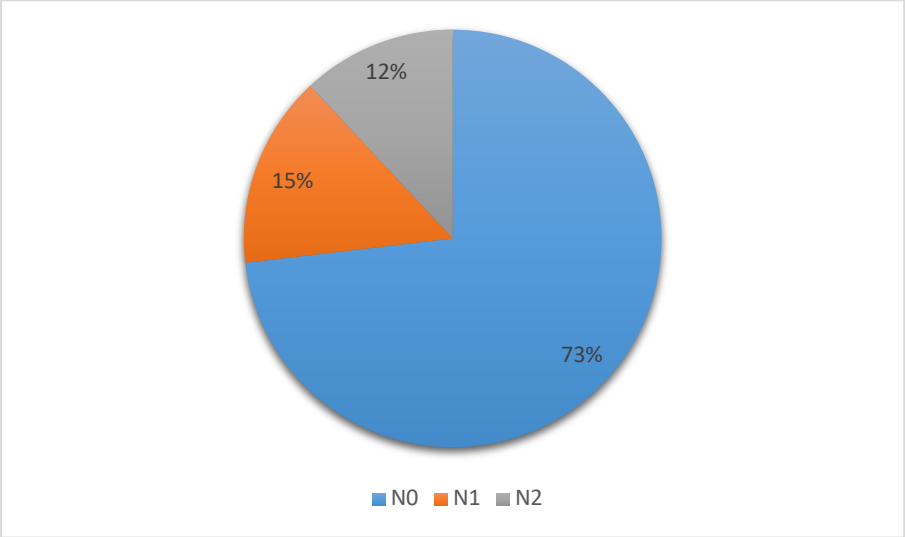
When considering the different stages of the lymph nodes, it is noticeable that only 10 patients showed stage N1 and eight patients N2. Due to this low number, it is difficult to draw conclusions (Figure 6 and Table 4). Patients in stage N1 showed only one mutation and one mutation in *CREBBP* (n=4), *NF1, ROS1* and *TP53* (n = 2 each). *CREBBP* is the most notable of these, as it was found in 40% of patients. *KMT2D* and *LTBP4* were the only two genes that were found multiple times in N2 patients. Both could only be found twice.

In previous parts of the results, it was shown that the average number of mutations per patient is indicative of an advanced stage. This could not be confirmed for lymph node involvement. Patients with N0 showed 2.59 mutations per patient, N1 patients 2.6 and only 1.8 mutations per patient could be detected in patients with the N2 stage. Many patients had stage N0 according to pathological findings (n=49). This group shows results as the entire cohort. *KMT2D* was the gene most frequently mutated with 28.5% (n=14). It was followed by *TP53* with 18% (n=9). *NOTCH1, POLQ*, and *USH2A* were detectable in 10% (n=5).

In summary, the number of mutations increases with stage T, but not with stage N. Furthermore, the T stages are close to the UICC stages in the type of mutations. The N-stages show a different pattern.

Figure 6: Lymph Node Involvement – included Patients

The pie chart reflects the stages of lymph nodes of included patients and shows that most patients do not yet have lymph node metastases.



3.5. Consideration of Smoking Behavior

In our patient population, it was not possible to collect information on smoking behavior for each patient. About 62 patients were informed. Eight patients never smoked in the past, twenty patients were former smokers, and 32 patients were active smokers at the time of surgery. Therefore, 87% of the information patients had a smoking history. In patients with squamous cell carcinoma, all patients had at least a history of smoking (in one case no information was available). An overview of smoking behavior, tumor histology, and stage of UICC can be found in Figure 7.

Smokers, former smokers, and non-smokers slightly differed in the number of average mutations per patient. Patients who never smoked had about half a mutation less on average. For non-smokers had 2.13 mutations per patient, smokers 2.53 and former smokers 2.55 mutations per patient. In eight non-smoking patients, 11 different genes were mutated. Four genes occurred more than once. *KMT2D* was the only gene found three times (37.5%). *EGFR, FANCA,* and *NOTCH1* were found twice. *EGFR* was found only in non-smokers (Table 5).

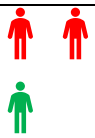
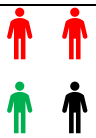
In smokers, genes repeat more frequently. *KMT2D* and *TP53* are the most frequently found genes. Both could be detected eight times, which corresponds to a share of 25%. In 15.5% (5/32) of the cases, *LTBP4* and *NF1* could be detected, and in four cases, *FGFR1-3* could be found in smokers. This corresponds to 80% of the *FGFR1-3* mutations found.

In former smokers, *CREBPP, TP53*, and *PTPRD* were found in 20% (n=4). Interestingly, *PTPRD* was found only in former smokers and *TP53* only in smokers or ex-smokers and not in non-smokers.

In summary, *EGFR* was found only in non-smokers, and *TP53* only in smokers or ex-smokers. *PTPRD* is also found only in ex-smokers.

Figure 7: Overview of smoking behavior, tumor histology, and stage of UICC

Smoking behavior of the patient is reflected and subdivided into lung cancer subtypes (ADC, SCC) and benign patients. In addition, the UICC stages are differentiated.

	IA (n=19)	IB (n=5)	IIA (n = 3).	IIB (n=10)	IIIA (n=10)	>IIIA (n=4)
Adenocarcinoma						
	IA (n=6)	IB (n=2)	IIA (n=1)	IIB (n=5)	IIIA (n=7)	
Squamous cell carcinoma						

Benign (n = 9);	

 Never Smoker

 Smoking in History

 Smoker

 Unknown

3.6. Comparison of tissue-derived DNA and plasma-derived DNA

DNA extracted from tumor cells was sequenced from 9 patients and compared with ctDNA. To compare tissue-derived DNA and plasma-derived DNA, tissue-derived DNA was done from cases BC 15, BC 18, BC 55, BC 68, BC 71, BC 88, BC 97, BC 117 and BC 119.

In the case of BC 15, 16 mutations were detected in the tumor and three in the plasma. None matched each other. In this case, the tumor was very small with T1c with advanced lymph node involvement of N2. In case of BC 18 31 mutated genes were detected in tissue-derived DNA and 17 in plasma-derived. A match of 15 genes was found. In percentage terms, this corresponds to a 48% match of mutated genes in the tumor and an 88% match of plasma ctDNA. Case BC 55 shows 16 mutations in the tumor and only one in the plasma. This matches the tumor. Expressed as a percentage, this means that only 6% of the mutations found in the tumor were detected in plasma and 100% of the mutations found in plasma matched tumor mutations. 11 mutations could be detected in tissue-derived DNA in case BC 68. No mutation was found in plasma. Thus, no match could be shown. As already described in case BC 15, a small tumor size is present here (T1c) and an advanced stage lymph node with N2. In BC 71 14 mutations were detected in the tumor and six in the plasma. Three of these six matched mutations in the tumor. Therefore, 21.5% of the tumor mutations matched those in plasma, and 50% of the plasma mutations matched those in the tumor. BC 88 showed 19 mutations in the tumor and eight in the plasma. Of these, six overlapped. In percentages, this means that tumor mutations matched 31.5% and plasma mutations 75%. In the case of BC 97, 12 mutated genes were detected in the tumor and six in the plasma. Of these six genes, five genes from the tumor matched. 41.5% of tumor genes were found in plasma and 83% of the plasma genes were matched to those of the tumor. 14 mutated genes were detected in the tumor of case BC 117 and seven in plasma. Five genes were overleaped. This means that 36% of tumor genes matched the detected ctDNA and 71% of the detected ctDNA matched tumor mutations. In the last case, BC 119, nine genes were found in the tumor and seven genes in the plasma. Six genes were matched. In percentage terms, this means that 67% of tumor mutations are found in the plasma and 83% of plasma mutations are found in the tumor.

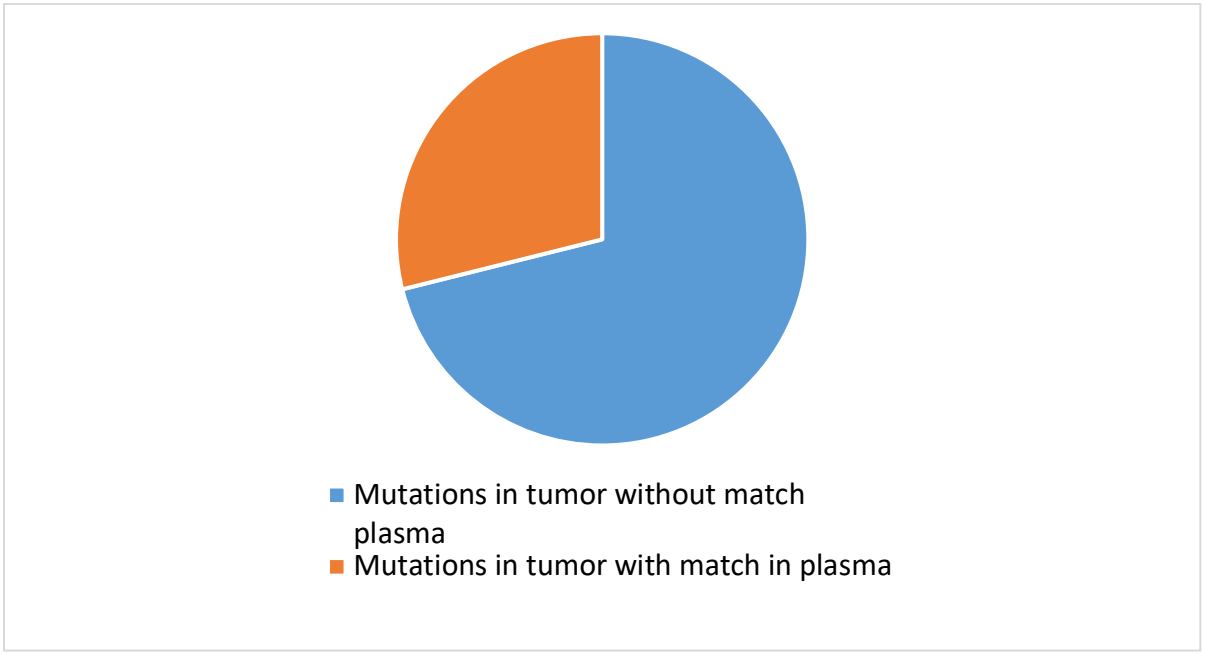
If these cases are considered as a whole, the average number of mutations per sequenced tumor is 15.78 and 6.11 mutations per patient are found in the plasma. 4.56 of these genes

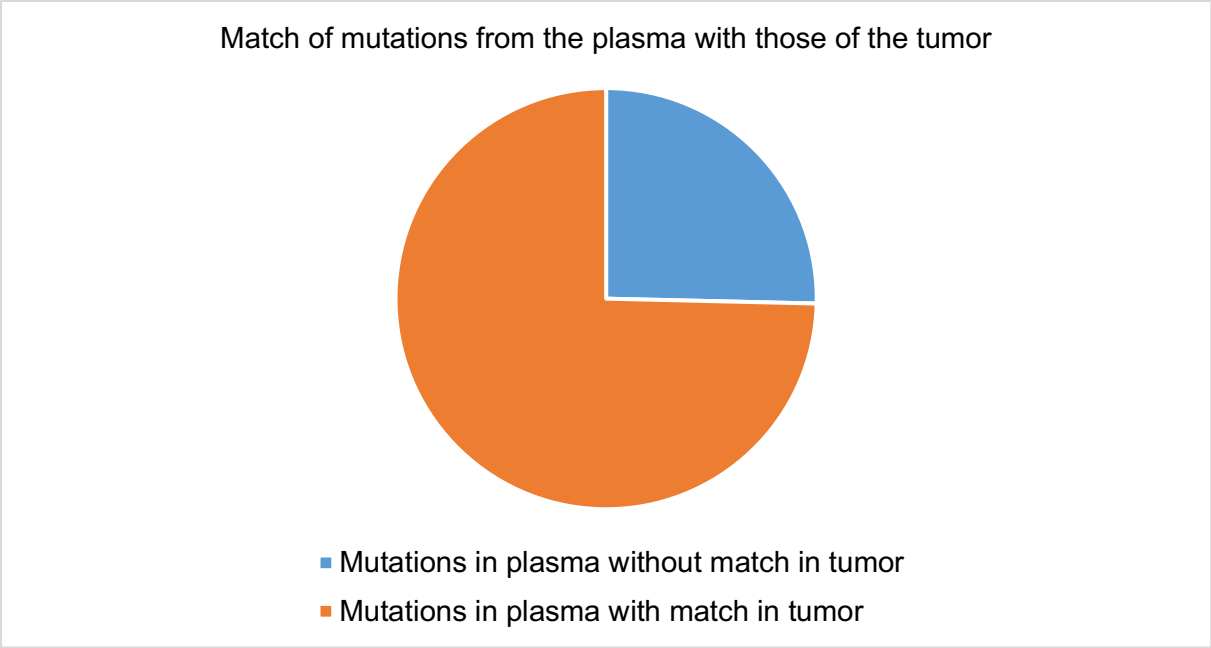
are on average consistent. On average, 29% of tumor mutations match those of plasma, and 74,6% of plasma mutations match those of tumor (figure 8).

In the following section, focus is placed on the genes found in the tumor, in the plasma, and on their overlap. The most frequently mutated gene found in the tumor is *PTPRD*. This was detected eight times in nine patients. Four of these eight (50%) were also found in plasma. *KMT2D* and *USH2A* were detected seven times. Both were present in plasma three times, representing 43% of the total. *AFF3, PTEN*, and *TP53* were found six times in the tumor. In the case of *AFF3* and *PTEN,* it was only possible to find them twice in plasma (33%). This is different for *TP53*. Mutations found in the tumor could also be found in plasma in 67% (n=4). The mutations *NLRP4* and *STK11* were shown to be present in the tumor five times. None of the five *NLRP4* mutations was found in plasma. This was similar for the *FLT3* and *HLA-A* gene mutations. These were detected three times in the tumor and none in the plasma. The genes that were mutated four times in the tumor were *APC, ARID1A, ATM, BRCA2, KRAS* and *POLQ. BRCA2* and *KRAS* are highlighted here as examples. *BRCA2* was detected twice (50%) in plasma, and *KRAS* was detected only once (25%).

None of the frequently mutated genes found in the tumor could be 100% detected in plasma. *TP53* with 67% has the highest overlap among the genes found frequently. *NLRP4* (n = 5), *FLT3*, and *HLA-A* (n=3 each) did not have a matching mutation in plasma.

Figure 8: Match of mutations from the tumor with those of the plasma:





3.7. A closer look at *KMT2D, TP53, FGFR1-3, KRAS*, and *EGFR*

Frequently detected genes, their characteristic evidence and the position of the mutation are worth looking at in more detail to highlight regularities. First, the most frequently detected gene *KMT2D* is considered. This was detected 18 times in our cohort of tumor patients. Histone lysine N-methyltransferase 2D (*KMT2D*) is part of six Set1-like H3K4 methyltransferases and a large protein with more than 5,500 amino acids. In 26.5% of all cases (18/68), it was possible to detect a mutated form of the gene. Especially in UICC stage I, it was by far the most frequently found gene with 26% and the second most frequent gene was only detectable in 9.5% of the cases. Despite this high percentage detection rate, there was progress in the other stages. In stage III, *KMT2D* was observed in 35% of the cases. This is a 9% increase in the detection rate from stage I to stage III. Interestingly, *KMT2D* was equally common in adenocarcinomas and squamous cell carcinomas (ADC: 26%; SCC: 27%), but in squamous cell carcinomas, *TP53* was the most detectable gene. A similar trend is observed when considering the T stages. Here, too, there is an increase in the detection rate from stage I with 23.5% to stage III with 37%. *KMT2D* is also the gene that is most often found here. However, *TP53* reaches stage I with 9% and stage III with 37%. Many of the *KMT2D* mutations appear to originate from tumors, as we detected seven *KMT2D* alterations in the nine tumors analyzed. Again, *KMT2D* was the gene that mutated most frequently with *USH2A*. Three of these seven mutants could also be detected in plasma. Now, we look at the relative allele frequency (AF). To do this, all patients were considered. It could be shown that on average 1.7% of the relative AF originated from mutated *KMT2D*. 14.72% was the highest result and 0.31% the lowest. In T1 an average of 0.75% (n=8), in T2 0.56% (n=4), in T3 1.44% (n=7), and in T4 6.04% (n = 3). relative AF was detectable. This suggests that relative AF also increases with tumor growth.

Next, the tumor suppressor gene *TP53* is the focus. This gene was detected a total of 12 times in tumor patients. It is interesting to note that it was not detected either in patients with neoadjuvant treatment or benign disease, but only in tumor patients. Shows a strong increase in the detection rate in the progressive stages of UICC. In stage I, it is detectable in only 9% of cases and in stage IIIA in 29.5%. The increase of 20.5% is the largest increase in the detection rate of all genes. As in *KMT2D,* a similar trend can be observed in terms of stages T. When the patients were in stage pT1, a mutation was detected in 9% and in pT3 in 37%. A marked difference was also observed in patients with adenocarcinoma and squamous cell carcinoma. In patients with adenocarcinoma, a gene change is observed in only 11% of cases and 32% of patients with squamous cell carcinoma. In patients with squamous cell carcinoma, *TP53* is also the most frequently mutated gene. These are two indications that *TP53* mutations are more strongly associated with squamous cell carcinoma. Smoking behavior also appears to have an influence on the frequency of the *TP53* mutation, as all mutations were shown in smokers or patients with a smoking history. Unfortunately, only in one patient with a *TP53* mutation is information about smoking behavior available.

In patients with relapse, it could not be shown that *TP53* was disproportionately mutated. In the sequencing of the nine tumors, a *TP53* mutation was detected six times. Four of these six mutations were found in plasma. 66% of tumor mutations were found in plasma. Such a good overlap was only seen by *TP53*. When looking at relative AF, on average 3.48% of the *TP53* alleles were mutated. In patients with T1, it was 0.6% (n=3), in T2 7.31% (n=2), in T3 2.1% (n = 4) and in T4 5.6% (n=3). No clear trend in relative AF can be observed here.

Fibroblast growth factor receptor 1-3 (*FGFR1-3*) was detected only five times in the entire cohort, but the distribution of the proto-oncogene is interesting. Two *FGFR 1* mutations were located on chromosome 8, the *FGFR2* mutation on chromosome 10, and the two *FGFR3* mutants on chromosome 4. For repeat mutations, the localizations in the gene differ. None of these mutations could be detected in a patient with benign disease. However, four of the five mutations were detected in patients with stage II UICC and all patients in whom an *FGFR* mutant was detected showed stage T3 in pathological findings and were therefore in an advanced stage. It is also striking that only patients with a positive smoking history (4 smokers and one patient with a smoking history) showed alterations.

60% (n=3) of the *FGFR1-3* mutations were found in patients who relapsed. No other gene was detected in this number of patients with recurrence, although there were only five *FGFR* mutations in total. This is a unique finding. If we look at the allele frequencies, we see that patients with recurrence show an average AF of 0.76% and the overall average is 1.6%. There is a patient who had a very high percentage of 4.54% and died shortly after surgery. Without this patient, the average relative AF would be 0.8%, which would still be higher on average than in patients with recurrence.

The last two genes, which will be considered in more detail, were detected very rarely. *KRAS* was only found once, and *EGFR* was found twice. The *KRAS* mutation was detected in a patient with stage IIB (pT3, pN0) and had an extremely high AF of 18.72%. The reason why *KRAS* is highlighted is that it was found only once in the entire cohort. In the nine tumor samples that were sequenced, four *KRAS* mutations alone were detected and one of them was also detectable in plasma (25%).

The epidermal growth factor receptor (*EGFR*) has only been detected twice in patients with lung cancer. Another mutation was shown in a patient with benign disease. This patient suffered from pneumonia. When looking at the mutations found in patients with lung cancer, both patients were non-smokers and suffered from adenocarcinoma. In the tumor samples that were sequenced, no *EGFR* mutation could be detected and therefore no statement can be made about the percentage overlap of the detection rate from detected mutations in the tumor and the overlap detection in plasma.

Genes described in the literature as very frequent were not found in part in this study with the described frequency. For this reason, we reduced the filter criteria for *KRAS, EGFR*, and *TP53* and reduced the absolute AF to 9 to avoid filtering and to check whether our filter criteria were too strict. During the process, 23 *TP53* mutations became apparent, 3 *KRAS* mutations, and 7 *EGFR* alterations. Of the 23 *TP53* mutations, 11 were already known and have reappeared. Therefore, 12 mutations remained, 2 of which were in benign patients and 2 in pretreated patients. As a result, 8 new *TP53* mutations were found with the lower filter. With a total of 20 mutations throughout the cohort, *TP53* would be the gene found the most frequently, 24.5%.

KRAS 3 mutations were detected even after the lower filter. One of these three has previously been found. Thus, two new ones were added; both were found in tumor patients. However, *KRAS* was detected in only 4.5% of the patients. For *EGFR*, it turns out that 7 mutations appeared with the lowered filter. Three of these seven were previously detected and a mutation was found in a benign patient. This leaves three newly detected variants and increases the number of *EGFR* mutations found to 5 (7%).

3.8. CtDNA detection in pretreated patients

Five of the 73 patients were pretreated with chemotherapeutic therapy or tyrosine kinase inhibitors (TKI) prior to surgery. In two cases, it is a personalized treatment. These cases were BC 104 and BC 133. The neoadjuvant patients were cases BC 41, BC 54 and BC 65.

BC 54 and BC 65 are very similar cases. Both had stage IA1 with ypT1a, ypN0 after neoadjuvant therapy. They also had blood drawn immediately prior to surgery and tested for ctDNA. No mutations were detectable in either of them. The two cases differ in their follow-up. Patient BC 65 survived the 24-month follow-up without recurrence. Patient BC 41 relapsed after nine months. This patient survived 24 months of follow-up. There is no information about the follow-up in case BC 54. In this case, surgical planning and a subsequent pathological

examination are available. A stage IIIA was found after neoadjuvant therapy. A still large tumor with ypT4 was found, but it does not appear to have metastasized lymphogenically (ypN0) or hemotogenically (cM0). In liquid biopsy, *NID1* was found to be mutated.

In case BC 104, the patient initially had a stadium IV. After neoadjuvant therapy, stage IIIC was still present. Therefore, no surgical therapy with a curative approach was attempted. Nine mutations could be detected in ccfDNA. These mutations are *USH2A, NID1, KEAP1, SMARCA4, LTBP4, POLQ, SETD2, ROS1,* and *FGFR1*. The patient was treated neoadjuvantly with chemotherapeutic agents, as no targetable mutations were found in biopsy analyzes. Additionally, in liquid biopsy, no such mutations could be found. Follow-up in case BC 104 shows that the patient relapsed after nine months and died after 12 months.

Patient BC 133 was also treated without a curative approach. Here, too, the patient had stage IV disease. An *EGFR* mutation was found in the biopsy and TKIs were used. After the tumor stage was reduced, a lobectomy was performed. This was also an alternative therapy attempt that did not meet the guidelines. In this liquid biopsy, a gene was also detected. *CNTN1* was found. Unfortunately, in this case, also, a recurrence was detected after 22 months.

NID1 was found in two of the five pretreated patients (Table 6). There is only one other *NID1* mutation in the entire cohort, and that was found in a patient with benign disease. *FGFR1* was also associated with recurrence with death during the follow-up period. The three patients with neoadjuvant therapy showed fewer plasma mutations. If the average number of mutations per patient for neoadjuvant patients is 0,33, that is by far the lowest number of mutations per patient. Only one of the three has a mutation in plasma.

3.9. Surgical patients with benign diseases

Nine patients with benign disease were clouded in this study as a control group. They were treated with surgery because lung cancer was suspected and could not be ruled out during staging. Pathological examination showed that it was a benign disease. Blood samples were taken before surgery (Table 7). These patients were classified into categories such as inflammatory processes, proliferation of benign tissues, and nonclassification.

Patient BC 13 showed in the pathological examination findings of lymoid interstitial pneumonia. In the liquid biopsy, four mutations (*SMARCA4, MPPED1, EGFR, NOTCH1*) were found. Also suffering from an inflammatory process was patient BC 17, who showed chronic necrotizing granulating fibrotizing inflammation and a mutation of the *NID1* gene in liquid biopsy. Another patient was found to have inflammation. In case BC 176, pathologists detected granulomatous inflammation. Four mutations were also detected in plasma (*USH2A, FLT3, TP63, PTPRD*). The three patients with inflammatory syndromes showed an average of three mutations per patient with an average relative AF of 0.51% (minimum value: 0.25%; maximum value: 1.05%). Three mutations per patient are on average more than in the tumor patient cohort.

Patients who showed benign tissue proliferation on pathological examination were cases BC 39, BC 43, BC 64, and BC 87. Patient BC 39 showed lipomatous tissue on histological examination. Examination showed two mutations in the sequencing of ccfDNA. In this case, *BRCA2* and *ADGRB3* were detected. A schwannoma was found in patient BC 43. Furthermore, no mutations were found in the plasma. A mutation in plasma could be detected in BC 64. In this case, an *ALK* mutation could be identified. Histologically, due to the anatomical proximity, it could not be conclusively clarified whether it was a cyst or a diverticulum of the trachea or bronchi. A clearer diagnosis was made in the case of BC 87, where fibroadipose tissue was found. Two mutations were detected in plasma (*ARID1A* and *NOS1*). On average, patients showed 1.25 mutations per patient. The average AF was 0.66% (minimum value: 0.31%; maximum value: 0.78%).

Two cases could not be classified into the two previous categories. In case BC 62, no mutation could be detected in plasma. This patient had antrohistocytosis. The last case examined with benign disease is patient BC 82. Histologically, haemorrhage was detected. In plasma, the patient was identified to have five mutations. This patient had the highest number of mutations in the plasma of all patients with benign diseases. In this case, *USH2A*, two *KMT2D, NF1* and *SETD2* were found.

17 different genes were detected in only nine subjects examined. Only two genes were repeated (*KMT2D* and *USH2A*). NID1 was only found in benign or pretreated patients in our study. Only two of the nine patients had no detection of ctDNA in plasma. In terms of the average number of mutations per patient, benign patients had half a mutation less than tumor patients (benign: 2.11 mutations per patient; lung cancer patients: 2.61 mutations per patient). Patients with an inflammatory process showed 3 mutations per patient, and those operated for benign tissue proliferation showed only 1.25 mutations per patient.

3.10. Patient outcome compared to ctDNA detection

A follow-up of the patients was carried out over two years. The primary objective was recurrence, and the secondary objective was survival. Follow-up could not be completed for all 74 patients. In 18 cases (24%), no follow-up could be achieved for the entire 24 months. In 56 patients (76%), information could be collected. Seven of the 56 patients died in two years (12.5%). In four cases, this occurred in the perioperative setting, in two cases a recurrence had been previously diagnosed before death, and in one case only the information about the death was collectable (figure 9). The four patients who died in the perioperative setting and the case in which no further information was obtained on the type of death were not included in the percentages indicating whether recurrences were observed. For this reason, complete follow-up was achieved in 51 patients.

Of 51 patients, 11 patients (21.5%) experienced a recurrence, and 40 patients (78.5%) survived this period without a diagnosed recurrence. 9 of the 11 patients with relapse had a positive liquid biopsy. CtDNA was detected in 82%.

24% of patients with ctDNA detection had recurrence (9/37). In 14 patients with negative ctDNA, 2 suffered recurrence. This corresponds to 14% of patients with negative ctDNA. This indicates that in our study, 10% more recurrences occurred in patients with positive ctDNA detection. 12.5% (n=7) of the patients died during the period considered. Two of the seven patients relapsed before dying. The first case is patient BC 104. This patient was no longer given a curative therapy approach because the tumor stage had progressed too far. As there was a very strong wish for therapy, chemotherapy was initially performed, followed by surgery. A more detailed description of the case can be found in the results section of the pretreated patients. The second case is BC 162. The patient was in stage IIIA and shows two mutations in cffDNA analysis (*MAP2K1, ACGRB3*). After 4 months after surgery, recurrence occurred, and after 6 months, the patient died. The other four deceased patients died in the perioperative setting or in one case no further information was available. Only one of the deceased patients had a negative ctDNA result. This means that only 7% of patients without plasma ctDNA findings died during the follow-up period (1/15). Compared to each other, 14.5% of patients with ctDNA detection died (6/41).

The mutated genes found in those patients with relapse were considered in more detail as the next step. Only three genes are repeated in the eleven patients. *KMT2D* and *MAP2K1* were detected twice. This is not more frequent than in the overall cohort. *FGFR1-3* was found three times and in all included subjects only five times in total. 60% of the *FGFR1-3* mutants were found in patients with relapse.

The amount of ccfDNA was also evaluated in relation to the recurrence rate (figure 10). Here, the average amount of the measured Qubit HS ccfDNA rate was used as a parameter. This was 1.94 ng/µl. The first group consisted of cases who had less than the average amount of ccfDNA and the second group consisted of patients who had more than the average amount in plasma. In the low ccfDNA group, 25 patients achieved complete follow-up and 21 patients in the high ccfDNA group. Patients who had received pretreatment were not included in this group. 19% (n=4) in the high ccfDNA group relapsed and 16% (n=4) in the low ccfDNA group. Patients in stage I of UICC were also evaluated separately. Follow-up was available for 25 patients. No differences in follow-up could be observed for patients with and without a positive liquid biopsy, because none of the patients had relapsed or died in the period considered.

Figure 9: Outcome of the patients

In A, the survival rate of patients with and without positive detection of ctDNA after 24 months is shown. In B, the recurrence rate is shown for 24 months and 24% of patients with positive detection suffered a recurrence and 14% without detection.

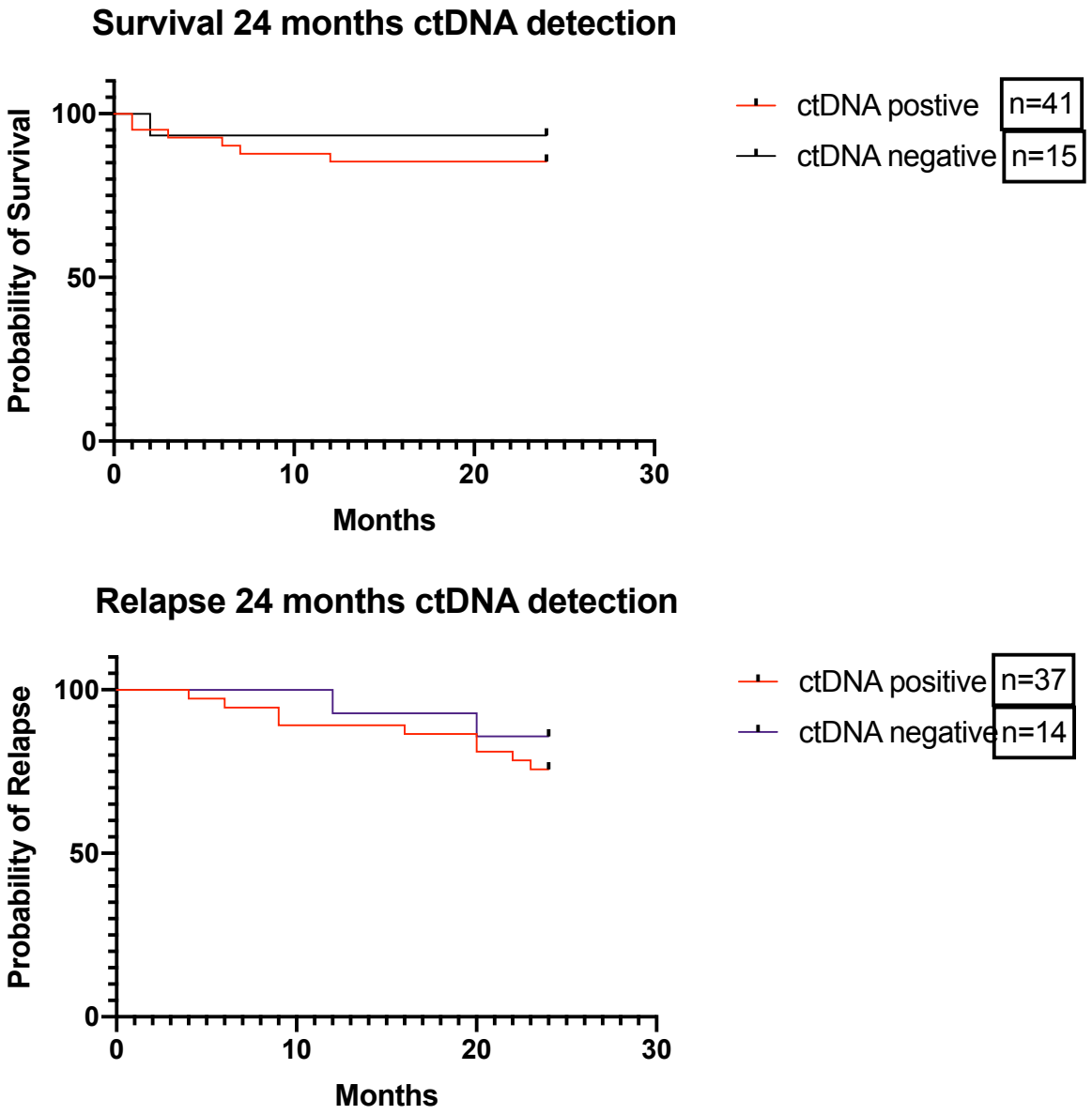
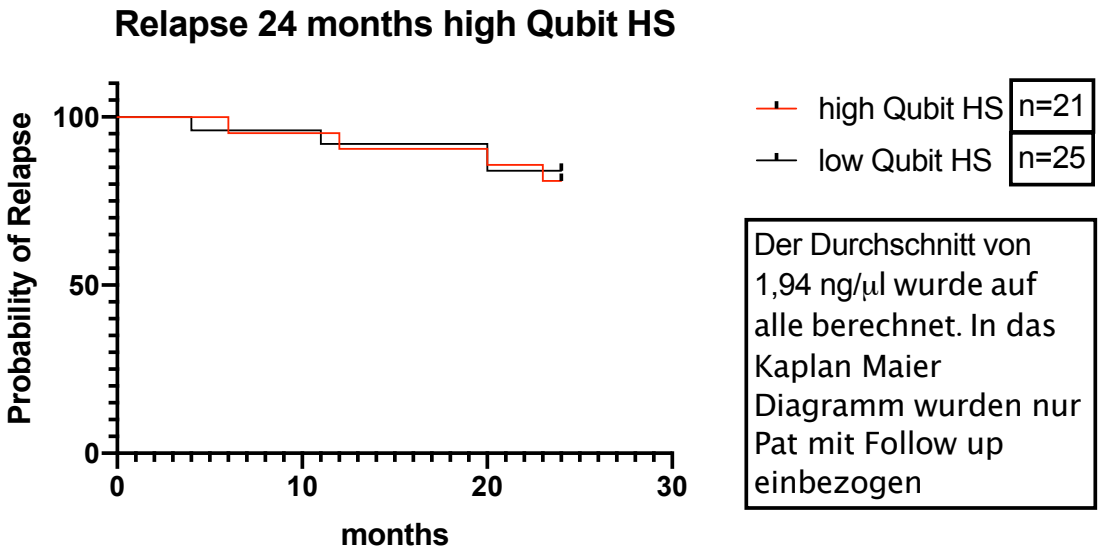


Figure 10: Relapse rate after 24 months compared to high and low proportion of ctDNA

The recurrence rate is shown in relation to the average quantity of ctDNA obtained from all samples using Qubit HS.

Comparison groups are made up of patients who have completed 24 months of follow-up and are above or below the average amount. There is no relevant difference between the two groups.


Relapse 24 months high Qubit HS
Probability of Relapse
100
50
0
0
10
20
30
months
high Qubit HS
n=21
low Qubit HS
n=25
Der Durchschnitt von 1,94 ng/µl wurde auf alle berechnet. In das Kaplan Maier Diagramm wurden nur Pat mit Follow up einbezogen

4. Discussion

Lung cancer screening and prognostic statements have significant potential for development. Therapy and follow-up are based on the UICC classification, as currently there is no better way to assess the outcome. In our study, we searched for early detection characteristics and prognostic markers in liquid biopsy.

In our cohort, the average age and smoking history fit the data of the average lung cancer patient. We have a disproportionate number of women, which will fit the perspective that the percentage of women will increase among patients. Furthermore, the proportion of 30% SCC and 69% ADC reflects other studies [34] [33]. Critically, the type of adjuvant therapy received by the patients was not considered and PY was not recorded for smoking status.

The detection rate of mutant genes is satisfactory compared to other highly published studies, such as the TRACERx study [57]. Compared to this study, we were able to identify at least 6% more patients with liquid biopsy mutations at each stage and this without large genes such as MUC in our gene panel, which generally remain without pathogenesis consequence. On average, we were able to detect at least one mutation in 73% (I: 68%, II: 72%, IIIA: 88%) of the patients. In more detail, we would like to highlight the influencing factors that make detection possible.

One factor is the histological history of NSCLC. SCC have considerably more mutations per sample than ADCs. This allows for statistically better detection of ctDNA. Furthermore, in our patient population, more *TP53* mutations are found in patients with SCC than in patients with ADC. *CREBBP* mutations should be of particular importance in this context, as 2/3 of SCC had a mutation here. Other studies also mentioned stage and tumor histology as an influence factor. Furthermore, metabolic tumor volume (MTV) is described as a factor that influences the detection rate [58]. The number of detected mutations per patient increases with UICC stages, as described. If the UICC stages are now considered in a more differentiated way. The influence of the T-stage is clearly greater than that of the N-stage. The number of mutations found continuously increases from T1 to T3, as in the UICC stages. This is different for N0 to N2. The N-stage does not seem to have any influence. Patients with a high T stage but a small N stage offer more mutations that can be observed than patients with a low T stage but an advanced N stage. This is probably due to the metastatic pathway. Lymphogenic metastasis is not reflected in ctDNA in the same way as a large tumor, which has more connections to blood vessels and possibly metastasizes to hematomas. This is supported by the study by Chabon et al. showing that MTV and sensitivity of 16%, 52% and 80% were detected for tumors with volumes of 1 ml, 10 ml, >100 ml, respectively. Therefore, tumor volume correlates with detection sensitivity [58].

In terms of the potential interaction of smoking and gene mutations, Halvorsen et al. conducted a study on the association of smoking and tumor driving gene mutations. There, more *TP53*

mutations were described in patients with a smoking history. Furthermore, more mutations were detected in the *TP53* gene in patients with SCC than in patients with ADC. These results were also found in our study and further contribute to this hypothesis [59]. Furthermore, some studies [60] presented increased mutations per patient, both more point mutations and a higher number of mutations/Mb, in smokers than in non-smokers. This result was not reflected in our data. Furthermore, pack years were also reported to have influence, which was not considered in our study.

Currently, clonal hematopoiesis is one of the biggest difficulties in liquid biopsy. Because differentiation of mutations originating from tumor cells or from hematopoietic stem cells that make blood remains is one of the greatest challenges. To illustrate this, we also sequenced cells from the tumor of nine patients with the same panel and compared the detected mutations with those from plasma. This showed a wide range of matches. On average, 29% of the tumor mutations matched those found in plasma, and 74,6% of the mutations detected in plasma were found in the tumor. This also means that on average 25,4%% of the mutations found in the plasma did not originate from the tumor. In this context, it is important to note that in our study the entire tumor was not sequenced. This means that mismatched mutations could still originate from the tumor, since cells that have this mutation are located at an unsequenced tumor site. At the same time, gene alterations were sequenced in 78% of patients with benign diseases and many mutations were found in patents with inflammatory processes.

This makes it clear that the pure detection of mutations is not a guarantee of malignant processes. This result has already been shown in several studies [35, 37, 58]. This raises the question whether individual genes show a particularly high degree of concordance and therefore are particularly well suited for differentiation. In our data*, TP53* has the highest match rate of mutations found in tumor and plasma. 67% of the mutations detected in plasma were also detected in the tumor. 29% of the mutations found in the tumor were also detectable in plasma and 69% of those found in plasma originated in the tumor. This clearly shows that single genes cannot be a solution if liquid biopsy is used as a screening method. The results would be far too inaccurate and variable. Chabon et al. have also addressed this issue and compared them with tumor-derived mutations and clonal hematopoiesis mutations. They considered several factors, for example, that clonal hematopoiesis mutations occur in longer cfDNA fragments. They combined these findings with other molecular features and integrated them. The result was a machine learning method termed lung cancer likelihood in plasma (Lung-CLiP), which can robustly discriminate early-stage lung cancer patients from risk-matched controls. This represents a possible way to solve the problem of HC and to enable differentiation [58].

In clinical practice, the detection of lung cancer in the early stages remains a major problem. In this area, liquid biopsy is expected to lead to many new possibilities. Liquid biospy is discussed as a screening method and if the result is positive, the patient should receive a chest CT. In this context, the problems of HC, which have been mentioned above, must be taken into account, and which genes are particularly suitable. *KMT2D* is a gene that was found in our data with 26% in stage I. It was found to be the most common gene by a large amount. Rao et al. explicitly addressed *KMT2* in their study and investigated the effects of a *KMT2* mutation in different tumor entities. They found that *KMT2C* and *KMT2D* are among the most frequently mutated genes in human tumor diseases. *KMT2D* is also described in lung cancer. *KMT2* encodes methyltransferases. Subsequent misregulated methylation has been discussed for some time as a major factor in the dysregulation of enhancer activity and thus the destruction of regular cell development. *KMT2C* and *KMT2D* are likely very important factors in tumor suppression. They also state that there is a direct interaction between *P53* and *KMT2D*, a promoter of the *P53* gene. Another statement is that a *KMT2D* mutation has a poor prognosis for breast cancer. This is not applicable to lung cancer, but it gives an incentive to look closely at this gene and consider it. Although *KMT2* encodes methyltransferases, it remains a very large gene and is therefore prone to many spontaneous mutations [61]. These statements are also supported by the study by Ardeshir-Larijani et al. [62]. They detect mutations in 17.5% of NSCLC tumors and 32.8% of SCLC tumors. This detection rate in their cohort is not as high as ours, but they demonstrate an association with a poor prognosis. Unfortunately, we cannot demonstrate a reduced survival rate for patients with a *KMT2D* mutation. These statements support our findings and show that alterations in the *KMT2D* gene are not only explicable by HC, but are likely to originate frequently from the tumor itself and should continue to be considered in research.

A much-discussed gene in liquid biopsy is *TP53*. Both as a prognostic factor and as a diagnostic element. In our study, the detection rate increased with the UICC stage and the T stage, and thus with the tumor volume. In stage I, a *TP53* mutation was detected in only 9% of patients and 37% of patients in stage III. Therefore, it remains to be questioned whether *TP53* is suitable for early detection. It will have to be considered that in our study it has the highest concordance of tumor mutations found in plasma. 67% of tumor *TP53* mutations could be found in plasma. *TP53* could be one of the genes that supports differentiation between HC and tumor-originating mutations. The statement by Chabon et al. that *TP53* is one of the most frequently mutated genes in cancer but is also frequently found in CH must be considered here. Furthermore, it establishes a strong link to smoking behavior, which we support with our data, as all mutations were detected in active or former smokers [58]. Razavi et al. also found that *TP53* mutations were found in 89.5% of cancer patients, but at the same time in 83% of non-cancer patients. *TP53* was among the top 15 genes associated with CH in their study [35].

These results further support our hypothesis that individual genes cannot be used to identify potential tumor patients, but rather an interplay of various factors. Another characteristic associated with an increased detection rate of *TP53* in our study was the type of tumor. In patients with adenocarcinoma, we detected a gene mutation in only 11% and in patients with squamous cell carcinoma in 32% of the cases, representing the mutation most frequently found. This result is described in the study by Ojlert, among others [63].
The *FGFR1-3* genes also deserve special consideration based on the results. In our cohort, mutations were only found in the advanced stage, which means T3. This is the stage at which the prognosis of curatively treated patients worsens. Furthermore, mutations in these genes in our results were associated with a positive smoking history and squamous cell carcinoma, which has already been described in the literature. For example, Kelleher et al. described that *FGFR2* was almost only detected in squamous cell carcinomas and *FGFR1* was described in 9-22% of cases of SCC and 73% of these patients were smokers [64]. In general, we detected only five *FGFR1-3* mutations in our cohort. Three of these five mutations were associated with recurrence within the follow-up time. Furthermore, 60% of the *FGFR* mutations were found in patients who were still in the curative stage, but relapsed. That *FGFR* mutations could be a possible prognostic factor has already been described in the review by Kelleher et al*. FGFR3* is associated with an increased risk of recurrence [64]. The study by Kim et al. reported a significantly higher risk of recurrence and shorter overall survival for *FGFR* mutations (51.2 vs. 115 months). They reached this conclusion after adjusting for gender, smoking status, pathological stage, and adjuvant therapy [65]. These previous results support our findings and should be investigated in more detail in future studies.

KRAS and *EGFR* have received a lot of attention in the literature and were detected very little in our cohort (n=2). It is not clear why the detection rate for *KRAS* and *EGFR* was so low. In particular, *KRAS* mutations were detected in 23-26.8% of cases in many studies. These studies are also patients in curative stages and a European cohort [39] [66]. With these viewpoints, we find no starting point to explain the low detection rate. The situation is somewhat different for *EGFR*. Especially in Asian cohorts, high detection rates are described, as in the study by Chen et al. in which detection was achieved in 47.4% of patients [67]. In studies with European cohorts, detections were significantly lower, as in the TRACERx study. Here, detection was only achieved in 9.4% or in the study by Ding et al. detection was 18.4% [39] [68]. Although these detection rates are lower and, therefore, not as different from our results as with *KRAS*, they are clearly different from our results. We do not have a proper explanation for this phenomenon for both genes. *NID1* could only be detected in patients with benign disease or pretreated patients in our study. This is also not in agreement with the literature, where it is also described that *NID1* mutations were detected in patients with NSCLC, were

significantly upregulated in some tumor entities, and even in patients with low-grade gliomas, *NID1* expression was associated with shorter overall survival [69].

Nine patients who were operated due to suspected lung cancer but have suffered benign diseases received a liquid biopsy. Only in two of these nine patients could no detection be achieved. It was striking that patients with inflammatory disease had far more mutations than patients with a benign tumor. This shows that the type of benign disease probably has an impact on the number of mutations, but not necessarily that mutations are found at all. It underscores the findings of other studies such as those of Chabon et al. [58]. Most somatic mutations in patients with lung cancer or patients at a similar risk of lung cancer are not caused by the carcinoma itself, but by representational mutations from clonal hematopoiesis.
Again, these findings are consistent with previous literature and underscore the importance of differentiating between mutations originating from lung cancer and clonal hematopoiesis in order to make a meaningful and reliable use of liquid biopsy in the future.

In addition to patients with benign disease, we also have five patients with neoadjuvant therapy in our study and received a personalized treatment. These were considered separately. Two of these five patients received a palliative treatment. In both cases ctDNAs could be detected in the liquid biopsy and both died during the follow-up period. In the other three patients, few somatic mutations could be detected. On average, only 0.33 mutations per patient. This is by far the lowest detection rate and shows, in our view, that neoadjuvant therapy complicates liquid biopsy. This hypothesis is only an idea due to the small number of patients and could be further investigated in further studies.

The two-year follow-up was completed successfully in 76% of the study subjects. The primary objective was recurrence in the patients and the secondary objective was survival of the patients. Based on these data, we analysed whether there were indicators of an earlier relapse or whether we found indications in our data.
One result, which was expected, is that the higher the UICC stage, the more likely relapse becomes. Regardless of whether ctDNA was detected or not. When positive ctDNA detection is now considered, only a 10% higher relapse rate is recorded in patients with positive ctDNA detection. In negatively tested patients 14% suffered a relapse and in patients with positive detection of ctDNA 24%. In stage I there is even no difference, as not a single patient suffered a recurrence. Other studies have identified the positive detection of ctDNA for all stages of UICC as a highly prognostic factor [58]. We cannot fully support these findings, but we show that patients with ctDNA have more frequent relapses than patients without it.

Kelleher et al. wrote in their review that the detection of the *FGFR 3* mutation in tumors was associated with an increased risk of recurrence. Furthermore, *FGFR amp*+ alteration was associated with shorter disease-free survival (26.9 vs. 94.6 months) and shorter overall survival (51.2 vs. 115 months). In general, after adjusting for gender, smoking status, pathological stage, and adjuvant therapy, they found that there was a significantly higher risk of relapse and death in patients with *FGFR1 amp*+ [64]. We also find evidence for these findings in our results, as *FGFR1-3* was detected only five times and three of these patients relapsed within the follow-up interval. FGFR mutations could be indications of shorter disease-free survival and shorter overall survival. This would need to be further investigated in further studies.
Unfortunately, we were unable to calculate the proportion of ctDNA from ccfDNA in a reliable way. Therefore, it is not possible to conclude whether a higher proportion has an influence on disease-free survival and/or overall survival. This has also been implicated as an influencing factor in some studies [58].

To surmise the discussion, highlight most important points, and topics which should be clarified in future shall be seen in the following.
It has been shown successfully that ctDNA detection is already successful in UICC stage I. In our results, tumor size, and thus UICC classification, plays an important role in the probability of detection. The higher the stage, the more likely the detection will be. Furthermore, the detection of the *KMT2D* mutation was particularly successful in the early stages. However, they did not have the highest concordance with the mutations we found in the tumor tissue itself. Here, the mutations of the *TP53* gene were able to provide the highest match. Even here, the match rate is too low to rely on *TP53* alone as a diagnostic tool.
Clonal hematopoiesis poses major problems in this field. In the future, differentiation between clonal hematopoiesis and a mutation originating from the tumor itself will be an important research approach. Some promising approaches have already been described during the discussion. Liquid biopsy has a lot of potential not only in the field of diagnostics. Liquid biopsy could also open new possibilities as a prognostic marker and could be included in future therapy planning.
Based on our results, the *FGFR* gene group offers a possible approach. Mutations in this gene family seem to have an influence on the prognosis. But also, here it will probably come down to a multimodal approach, which takes different factors into account.
Liquid biopsy will contribute to the early diagnosis of lung cancer and promises to play an important role in early detection and treatment, as well as precision medicine, individual treatment, and prediction of prognosis.

5. Literaturverzeichnis

1. Frank Griesinger, W.E., Martin Früh, Oliver Gautschi, Wolfgang Hilbe, Hans Hoffmann, Rudolf Maria Huber, Sonja Loges, Robert Pirker, Christoph Pöttgen, Ron Pritzkuleit, Martin Reck, Niels Reinmuth, Martin Sebastian, Dieter Ukena, Cornelius Waller, Jürgen Wolf, Martin Wolf, Bernhard Wörmann In Kooperation mit der AIO, *Lungenkarzinom, nicht-kleinzellig (NSCLC).* Oktober 2019.
2. Burrell, R.A., et al., *The causes and consequences of genetic heterogeneity in cancer evolution.* Nature, 2013. **501**(7467): p. 338-45.
3. Mohr, A., et al., *Dyes for Eyes: hydrodynamics, biocompatibility and efficacy of 'heavy' (dual) dyes for chromovitrectomy.* Ophthalmologica, 2013. **230 Suppl 2**: p. 51-8.
4. AL, V.H., M. Eggel, and N. Biller-Andorno, *Is selecting better than modifying? An investigation of arguments against germline gene editing as compared to preimplantation genetic diagnosis.* BMC Med Ethics, 2019. **20**(1): p. 83.
5. Gould, M.K., et al., *Evaluation of individuals with pulmonary nodules: When is it lung cancer?: Diagnosis and management of lung cancer: American College of Chest Physicians evidence-based clinical practice guidelines.* Chest, 2013. **143**(5): p. e93S-e120S.
6. Ahlborn, L.B., et al., *Application of cell-free DNA for genomic tumor profiling: a feasibility study.* Oncotarget, 2019. **10**(14): p. 1388-1398.
7. Yeow, K.-M., et al., *Risk factors of pneumothorax and bleeding: multivariate analysis of 660 CT-guided coaxial cutting needle lung biopsies.* Chest, 2004. **126**(3): p. 748-754.
8. Devarakonda, S., D. Morgensztern, and R. Govindan, *Genomic alterations in lung adenocarcinoma.* Lancet Oncol, 2015. **16**(7): p. e342-51.
9. Guibert, N., et al., *Current and future applications of liquid biopsy in nonsmall cell lung cancer from early to advanced stages.* Eur Respir Rev, 2020. **29**(155).
10. Chen, K., et al., *Liquid biopsy in newly diagnosed patients with locoregional (I-IIIA) non-small cell lung cancer.* Expert Rev Mol Diagn, 2019. **19**(5): p. 419-427.
11. Kaisaki, P.J., et al., *Targeted Next-Generation Sequencing of Plasma DNA from Cancer Patients: Factors Influencing Consistency with Tumour DNA and Prospective Investigation of Its Utility for Diagnosis.* PLoS One, 2016. **11**(9): p. e0162809.
12. Van Paemel, R., et al., *The pitfalls and promise of liquid biopsies for diagnosing and treating solid tumors in children: a review.* Eur J Pediatr, 2020. **179**(2): p. 191-202.
13. Li, W., et al., *Liquid biopsy in lung cancer: significance in diagnostics, prediction, and treatment monitoring.* Mol Cancer, 2022. **21**(1): p. 25.
14. Ilie, M. and P. Hofman, *Pros: Can tissue biopsy be replaced by liquid biopsy?* Transl Lung Cancer Res, 2016. **5**(4): p. 420-3.
15. Liang, W., et al., *Liquid biopsy for early stage lung cancer.* J Thorac Dis, 2018. **10**(Suppl 7): p. S876-S881.
16. Hofman, V., et al., *Detection of circulating tumor cells from lung cancer patients in the era of targeted therapy: promises, drawbacks and pitfalls.* Current molecular medicine, 2014. **14**(4): p. 440-456.
17. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-715.

18. Zeinali, M., et al., *High-Throughput Label-Free Isolation of Heterogeneous Circulating Tumor Cells and CTC Clusters from Non-Small-Cell Lung Cancer Patients.* Cancers (Basel), 2020. **12**(1).

19. Kulasinghe, A., et al., *Circulating tumour cells in metastatic head and neck cancers.* Int J Cancer, 2015. **136**(11): p. 2515-23.

20. Hodgkinson, C.L., et al., *Tumorigenicity and genetic profiling of circulating tumor cells in small-cell lung cancer.* Nature medicine, 2014. **20**(8): p. 897.

21. Kang, B.J., et al., *Circulating Tumor Cell Number Is Associated with Primary Tumor Volume in Patients with Lung Adenocarcinoma.* Tuberc Respir Dis (Seoul), 2020. **83**(1): p. 61-70.

22. Pierga, J.-Y., et al., *Circulating tumor cell detection predicts early metastatic relapse after neoadjuvant chemotherapy in large operable and locally advanced breast cancer in a phase II randomized trial.* Clinical Cancer Research, 2008. **14**(21): p. 7004-7010.

23. Hannafon, B.N. and W.-Q. Ding, *Intercellular communication by exosome-derived microRNAs in cancer.* International journal of molecular sciences, 2013. **14**(7): p. 14240-14269.

24. Nogues, L., et al., *The influence of tumour-derived extracellular vesicles on local and distal metastatic dissemination.* Mol Aspects Med, 2018. **60**: p. 15-26.

25. Mohrmann, L., et al., *Liquid Biopsies Using Plasma Exosomal Nucleic Acids and Plasma Cell-Free DNA Compared with Clinical Outcomes of Patients with Advanced Cancers.* Clin Cancer Res, 2018. **24**(1): p. 181-188.

26. Frydrychowicz, M., et al., *Exosomes–structure, biogenesis and biological role in non-small-cell lung cancer.* Scandinavian journal of immunology, 2015. **81**(1): p. 2-10.

27. Mouliere, F., et al., *Circulating Cell-Free DNA from Colorectal Cancer Patients May Reveal High KRAS or BRAF Mutation Load.* Transl Oncol, 2013. **6**(3): p. 319-28.

28. Thierry, A., et al., *Origins, structures, and functions of circulating DNA in oncology.* Cancer and metastasis reviews, 2016. **35**(3): p. 347-376.

29. Oellerich, M., P. Kanzow, and P.D. Walson, *Therapeutic drug monitoring - Key to personalized pharmacotherapy.* Clin Biochem, 2017. **50**(7-8): p. 375-379.

30. Vymetalkova, V., et al., *Circulating Cell-Free DNA and Colorectal Cancer: A Systematic Review.* Int J Mol Sci, 2018. **19**(11).

31. Oxnard, G.R., et al., *Association between plasma genotyping and outcomes of treatment with osimertinib (AZD9291) in advanced non–small-cell lung cancer.* Journal of clinical oncology, 2016. **34**(28): p. 3375.

32. National Lung Screening Trial Research, T., et al., *Reduced lung-cancer mortality with low-dose computed tomographic screening.* N Engl J Med, 2011. **365**(5): p. 395-409.

33. Zhao, Y.R., et al., *NELSON lung cancer screening study.* Cancer Imaging, 2011. **11**(1A): p. S79.

34. Passiglia, F., et al., *Monitoring blood biomarkers to predict nivolumab effectiveness in NSCLC patients.* Ther Adv Med Oncol, 2019. **11**: p. 1758835919839928.

35. Razavi, P., et al., *High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants.* Nat Med, 2019. **25**(12): p. 1928-1937.

36. Hu, Y., et al., *False-positive plasma genotyping due to clonal hematopoiesis.* Clinical Cancer Research, 2018. **24**(18): p. 4437-4443.

37. Chen, K., et al., *Comparison of plasma to tissue DNA mutations in surgical patients with non–small cell lung cancer.* The Journal of thoracic and cardiovascular surgery, 2017. **154**(3): p. 1123-1131. e2.

38. Abbosh, C., et al., *Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution.* Nature, 2017. **545**(7655): p. 446-451.

39. Abbosh, C., N.J. Birkbak, and C. Swanton, *Early stage NSCLC—challenges to implementing ctDNA-based screening and MRD detection.* Nature Reviews Clinical Oncology, 2018. **15**(9): p. 577-586.

40. Pickhardt, P.J., et al., *Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults.* N Engl J Med, 2003. **349**(23): p. 2191-200.

41. Zauber, A.G., et al., *Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths.* N Engl J Med, 2012. **366**(8): p. 687-96.

42. Yang, Y.C., et al., *Circulating tumor DNA detectable in early- and late-stage colorectal cancer patients.* Biosci Rep, 2018. **38**(4).

43. Bettegowda, C., et al., *Detection of circulating tumor DNA in early- and late-stage human malignancies.* Sci Transl Med, 2014. **6**(224): p. 224ra24.

44. Fleshner, P., et al., *Tumor-associated DNA mutation detection in individuals undergoing colonoscopy.* Cancer Med, 2018. **7**(1): p. 167-174.

45. Lecomte, T., et al., *Detection of free-circulating tumor-associated DNA in plasma of colorectal cancer patients and its association with prognosis.* Int J Cancer, 2002. **100**(5): p. 542-8.

46. Qin, Y., et al., *Contrast-enhanced spectral mammography: A potential exclusion diagnosis modality in dense breast patients.* Cancer Med, 2020.

47. Zubor, P., et al., *Why the Gold Standard Approach by Mammography Demands Extension by Multiomics? Application of Liquid Biopsy miRNA Profiles to Breast Cancer Disease Management.* Int J Mol Sci, 2019. **20**(12).

48. Yi, Z., et al., *Landscape of somatic mutations in different subtypes of advanced breast cancer with circulating tumor DNA analysis.* Sci Rep, 2017. **7**(1): p. 5995.

49. Sikkema-Raddatz, B., et al., *Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics.* Human mutation, 2013. **34**(7): p. 1035-1042.

50. Collins, F.S., et al., *New goals for the US human genome project: 1998-2003.* science, 1998. **282**(5389): p. 682-689.

51. Avila-Flores, R. and R.A. Medellin, *Ecological, taxonomic, and physiological correlates of cave use by mexican bats.* Journal of Mammalogy, 2004. **85**(4): p. 675-687.

52. Hindson, B.J., et al., *High-throughput droplet digital PCR system for absolute quantitation of DNA copy number.* Analytical chemistry, 2011. **83**(22): p. 8604-8610.

53. Chang, F. and M.M. Li, *Clinical application of amplicon-based next-generation sequencing in cancer.* Cancer genetics, 2013. **206**(12): p. 413-419.

54. Newman, A.M., et al., *An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage.* Nature medicine, 2014. **20**(5): p. 548.

55. Samorodnitsky, E., et al., *Evaluation of hybridization capture versus amplicon-based methods for whole-exome sequencing.* Human mutation, 2015. **36**(9): p. 903-914.

56. Streubel, A., et al., *Comparison of different semi-automated cfDNA extraction methods in combination with UMI-based targeted sequencing.* Oncotarget, 2019. **10**(55): p. 5690-5702.

57. Jamal-Hanjani, M., et al., *Tracking the Evolution of Non-Small-Cell Lung Cancer.* N Engl J Med, 2017. **376**(22): p. 2109-2121.

58. Chabon, J.J., et al., *Integrating genomic features for non-invasive early lung cancer detection.* Nature, 2020. **580**(7802): p. 245-251.

59. Halvorsen, A.R., et al., *TP53 Mutation Spectrum in Smokers and Never Smoking Lung Cancer Patients.* Front Genet, 2016. **7**: p. 85.

60. Govindan, R., et al., *Genomic landscape of non-small cell lung cancer in smokers and never-smokers.* Cell, 2012. **150**(6): p. 1121-34.

61. Rao, R.C. and Y. Dou, *Hijacked in cancer: the KMT2 (MLL) family of methyltransferases.* Nat Rev Cancer, 2015. **15**(6): p. 334-46.

62. Ardeshir-Larijani, F., et al., *KMT2D Mutation Is Associated With Poor Prognosis in Non-Small-Cell Lung Cancer.* Clin Lung Cancer, 2018. **19**(4): p. e489-e501.

63. Ojlert, A.K., et al., *The immune microenvironment in non-small cell lung cancer is predictive of prognosis after surgery.* Mol Oncol, 2019. **13**(5): p. 1166-1179.

64. Kelleher, F.C., et al., *Fibroblast growth factor receptors, developmental corruption and malignant disease.* Carcinogenesis, 2013. **34**(10): p. 2198-205.

65. Kim, H.R., et al., *Fibroblast growth factor receptor 1 gene amplification is associated with poor survival and cigarette smoking dosage in patients with resected squamous cell lung cancer.* J Clin Oncol, 2013. **31**(6): p. 731-7.

66. Imielinski, M., et al., *Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing.* Cell, 2012. **150**(6): p. 1107-20.

67. Chen, J., et al., *Genomic landscape of lung adenocarcinoma in East Asians.* Nat Genet, 2020. **52**(2): p. 177-186.

68. Ding, L., et al., *Somatic mutations affect key pathways in lung adenocarcinoma.* Nature, 2008. **455**(7216): p. 1069-75.

69. Zhang, B., et al., *Nidogen-1 expression is associated with overall survival and temozolomide sensitivity in low-grade glioma patients.* Aging (Albany NY), 2021. **13**(6): p. 9085-9107.

6. Annex

6.1. List of Tables

Table 1: Mutated genes in tumor patients and their amount

Gene	n
KMT2D	18
TP53	12
CREBBP	8
LTBP4	6
FANCA	5
KEAP1	5
NF1	5
NOS1	5
NOTCH1	5
POLQ	5
USH2A	5
ARID1A	4
DNMT3A	4
FOXP1	4
PTPRD	4
APC	3
BRCA2	3
ERCC2	3
HGF	3
MSH2	3
ADGRB3	2
AFF3	2
BRAF	2
CDK4	2
CNTN1	2
EGFR	2
FGFR1	2
FGFR3	2
MPPED1	2
NBN	2
NKX2-1	2

NLRP4	2
NRAS	2
PTEN	2
ROS1	2
SETD2	2
SMARCA4	2
STK11	2
TERT	2
TP63	2
ALK	1
ARFGEF2	1
ATM	1
BPIFB1	1
BRIP1	1
CTNNB1	1
FBXW7	1
FGFR2	1
HLA-A	1
KDM6A	1
KRAS	1
MAP2K1	1
MET	1
NRP2	1
PIK3CA	1
PPPRD	1
RASA1	1
RBM10	1
RET	1
SOX2	1

Table 2: Detailed Case Overview

ID	Age	Sex	Histology	Stage	pTNM	Resection	Type of operation	Neoadjuvant therapy	malignant past diseases	Type of previous malignant disease
BC 10	83	m	adeno	IA3	pT1c, pN0	R0	left UL-resection	No	No	
BC 101	77	m	adeno	IIB	pT1b, pN2	R0	left UL-resection	No	No	

BC 103	86	m	adeno	IA3	pT1c, pN0	R0	right UL-resection	No	Yes	c.a. invasive ductal breast cancer/ basal cell carcinoma/ squamous cell carcinoma shoulder left
BC 104	60	f	adeno	IIIC	ypT4, ypN2	R0	right LL-resection	Cx	No	
BC 107	65	f	adeno	IIIA	pT3, pN1	R0	right LL-resection	No	No	
BC 109	75	f	adeno	IA3	pT1c, pN0	R0	left LL-resection	No	No	
BC 11	59	f	adeno	IA2	pT1b, pN0	R0	left UL-resection	No	No	
BC 110	70	f	squamous	IA2	pT1b, Nx	R0	right UL-resection	No	No	
BC 111	83	m	squamous	IB	pT2, pN0	R0	left LL-resection	No	No	
BC 113	61	m	adeno	IIIA	pT1b, pN2	R0	right UL-resection	No	No	
BC 117	71	f	adeno	IIB	pT3, pN0	R0	left UL-resection	No	No	
BC 119	68	m	squamous	IIIA	pT4, pN0	R0	left LL-resection	No	No	
BC 121	62	f	adeno	IA2	pT1b, pN0	R0	right UL-resection	No	No	
BC 123	77	m	adeno	IB	pT2a, pN0	R0	right UL-resection	No	No	
BC 124	83	m	squamous	IB	pT2a, pN0	R0	right pneuectomy	No	Yes	plasmocytoma
BC 125	78	f	adeno	IIIA	pT4, pN0	R0	right LL-resection	No	No	
BC 126	61	f	NET	IB	pT2a, pNx	R0	right UL-resection	No	No	
BC 127	62	f	squamous	IA2	pT1b, pN0	R0	left LL-resection	No	No	
BC 128	62	m	adeno	IA2	pT1b, pN0	R0	right UL-resection	No	No	
BC 129	75	m	squamous	IIB	pT1c, pN1	R0	right UL-resection	No	No	
BC 130	75	m	adeno	IIIA	pT4, pN0	R0	left UL-Resektion	No	No	
BC 132	68	f	adeno	IA2	pT1b, pN0	R0	left LL-resection	No	No	
BC 134	65	f	adeno	IIIA	pT3, pN1	R0	left LL-resection	No	Yes	c.a. meningoim
BC 135	73	f	adeno	IB	pT2a, pN0	R0	left LL-resection	No	No	
BC 137	55	f	adeno	IIB	pT3, pN0	R0	right LL-resection	No	No	
BC 139	71	f	adeno	IA2	pT1b, pN0	R0	right enucleation UL	No	No	
BC 142	75	f	squamous	IA2	pT1c, pN0	R0	left UL-resection	No	Yes	c.a. benign mamma tumor left, Brenner tumor
BC 144	76	m	squamous	IIIA	pT3, pN1	R0	left UL-resection	No	No	
BC 15	59	m	adeno	IIB	pT1c pN2	R0	right UL-resection	No	No	
BC 151	73	m	squamous	IIIA	pT2a, pN2	R0	right upper bilobectomy	No	Yes	oral carcinoma c.a.
BC 152	57	f	adeno	IIB	pT1b, pN1	R0	right ML-resection	No	No	
BC 153	73	m	adeno	IA2	pT1b, pN0	R0	right LL-resection	No	No	

BC 154	60	m	squamous	IIB	pT2a, pN1	R0	right UL-resection	No	No	
BC 16	50	m	adeno	IA2	pT1b, pN0	R0	right UL-resection	No	No	
BC 162	61	f	adeno	IIIA	pT1c, pN2	R0	right ML-resection	No	No	
BC 18	79	f	adeno	IIB	pT2b, pN0	R0	left UL-resection	No	No	
BC 19	63	m	adeno	IA1	pT1a, pN0	R0	right UL-resection	No	Yes	suspicion of stomach carcinoma
BC 28	69	f	adeno	IB	pT2a, pN0	R0	right UL-resection	No	No	
BC 29	77	m	adeno	IIA	pT1b, pN1	R0	left LL-resection	No	No	
BC 31	64	m	adeno	IA2	pT1b, pN0	R0	left UL-resection	No	No	
BC 33	51	f	adeno	IA1	pT1a, pN0	R0	right UL-resection	No	No	
BC 41	65	m	adeno	IA1	ypT1a, ypN0	R0	pneuectomy left pneuectomy	Cx + Rx	No	
BC 44	62	f	adeno	IA1	pT1a, pN0	R0	right UL-resection	No	No	
BC 46	64	m	adeno	IIB	pT3, pN0	R0	right LL-resection	No	No	
BC 51	61	m	adeno	IIA	pT2b, pN0	R0	right UL-resection	No	No	
BC 52	77	f	adeno	IB	pT2a, pN0	R0	left LL-resection	No	Yes	c.a. hastrointerstinal strumor of the small gastric cavity
BC 53	74	f	adeno	IB	pT2a, pN0	R0	right UL-resection	No	No	
BC 54	67	f	adeno	IIIA	ypT4, ypN0	R0	right pneuectomy	Cx + Rx	No	
BC 55	71	m	adeno	IIIA	pT4, pN1	R0	right pneuectomy	No	No	
BC 56	71	m	adeno	IIIB	pT3, pN2	R0	left UL-resection	No	No	
BC 58	75	m	squamous	IIB	pT3, pN0	R0	right UL-resection	No	Yes	c.a. epithelial epi+oropharynx Ca (1990) oral floor PE Ca (2003) prostate Ca (2012)
BC 59	57	m	adeno	IIA	pT2b, pN0	R0	right UL-resection	No	No	
BC 60	59	m	squamous	IIIA3	pT2b, pN2	R0	left LL-resection	No	No	
BC 65	80	m	squamous	IA1	ypT1a, pN0	R0	pneuectomy left pneuectomy	Cx + Rx	Yes	c.a. prostate carcinoma
BC 66	70	m	squamous	IA2	pT1b, pN0	R0	right UL-resection	No	No	
BC 67	73	m	squamous	IA3	pT1c, pN0	R0	right lower bilobectomy	No	No	
BC 68	70	m	adeno	IIB	pT1b, pN2	R0	right upper bilobectomy	No	No	
BC 71	78	f	squamous	IIIB	pT3, pN2	R0	right lower bilobectomy	No	No	
BC 72	76	m	adeno	IA2	pT1b, pN0	R0	left UL-resection	No	Yes	c.a. muscle-invasive bladder cancer
BC 76	76	f	adeno	IA3	pT1c, pN0	R0	right UL-resection	No	No	
BC 77	76	f	adeno	IIIA	pT4, pN0	R0	right upper bilobectomy +	No	No	

							wedge resection S6 right			
BC 78	62	m	adeno	IIIB	pT4, pN2	R0	pneuectomy left pneuectomy	No	No	
BC 80	79	f	adeno	IIIA	pT3, pN1	R0	left LL-resection	No	No	
BC 86	67	m	adeno	IIIA	pT3, pN1	R0	right pneuectomy	No	No	
BC 88	64	f	squamous	IIIA	pT4, pN0	R0	left UL-resection	No	No	
BC 90	58	f	adeno	IA2	pT1b, pN0	R0	right LL-resection	No	No	
BC 91	78	f	adeno	IA3	pT1c, pN0	R0	left LL-resection	No	Yes	c.a. mamma tumor
BC 92	72	f	squamous	IIB	pT3, pN0	R0	left LL-resection	No	Yes	c.a. parotid carcinoma
BC 96	71	f	adeno	IIB	pT3, pN0	R0	right LL-resection	No	Yes	c.a. mamma tumor
BC 97	65	m	squamous	IIIA	pT4, pN0	R0	right UL-resection	No	No	
BC 98	77	m	adeno	IA1	pT1a, pN0	R0	right UL-resection	No	Yes	c.a. squamous cell carcinoma of the vocal folds
BC 84	60	f	squamous	IIA	pT1c, pN1	R0	pneuectomy left pneuectomy	No	No	
BC 148	74	m	squamous	IIB	pT3, pN0	R0	right lower bilobectomy	No	No	

Table 3: Mutations at T-stage

T1: 1.44 Mut/Pat

T2: 3.08 Mut/Pat

T3: 4.21 Mut/Pat

T1 (n= 34)		**T2 (n= 13)**		**T3 (n= 20)**	
Gene	**n**	**Gene**	**n**	**Gene**	**n**
KMT2D	8	HGF	3	KMT2D	7
LTBP4	3	KMT2D	3	TP53	7
TP53	3	PTPRD	3	FGFR1-3	5
POLQ	3	CREBBP	2	CREBBP	4
BRAF	2	DNMT3A	2	NOTCH1	3
CREBBP	2	KEAP	2	APC	3
FANCA	2	ROS1	2	ARID1A	3
NF1	2	TP53	2	KEAP1	3
NOS1	2				
USH2A	2				

Table 4: Mutations in the N-stage

N0: 2.59

N1: 2.6

N2: 1.8

N0 (n=49)		N1 (n=10)		N2 (n=8)	
Gene	**n**	**Gene**	**n**	**Gene**	**n**
KMT2D	14	CREBBP	4	KMT2D	2
TP53	9	NF1	2	LTBP4	2
NOTCH1	5	ROS1	2		
POLQ	5	TP53	2		
USH2A	5				

Table 5: Mutations in smokers

Smoker: 2,53 Mut/Pat

Exsmoker: 2,55 Mut/Pat

Non-Smoker: 2,13 Mut/Pat

Smoker (n=32)		Exsmoker (n=20)		Never (n=8)	
Gene	**n**	**Gene**	**n**	**Gene**	**n**
KMT2D	8	CREBBP	4	KMT2D	2
TP53	8	PTPRD	4	EGFR	2
LTBP4	5	TP53	4	FANCA	2
NF1	5	NOTCH1	3	NOTCH1	2
FGFR 1-3	4				

Table 6: Mutations in neoadjuvant treated patients

Gene	n
NID1	2
CNTN1	1
FGFR1	1
KEAP1	1
LTBP4	1
POLQ	1
ROS1	1
SETD2	1
SMARCA4	1
USH2A	1

Table 7: Mutations in patients with benign disease

Gene	n
KNT2D	2
USH2A	2
ADGRB3	1

ALK	1
ARID1A	1
BRCA2	1
EGFR	1
FLT3	1
MPPED1	1
NF1	1
NID1	1
NOS1	1
NOTCH1	1
PTPRD	1
SETD2	1
SMARCA4	1
TP63	1

7. Vorabveröffentlichungen von Ergebnissen

Vorträge:

1. 23. Chirurgischen Forschungstage vom 12.-14.09.2019: Detection of circulating tumor DNA in patients with lung cancer in non-metastatic stage as diagnostic tool
2. 199. Tagung der Kölner Chirurgenvereinigung am 04.12.2019: Liquid Biopsy bei Lungenkrebs – Detektion zirkulierender tumor DNA in Patienten mit Lungenkrebs im operablen Stadium

Veröffentlichungen:

1. Detection of circulating tumor DNA by digital droplet PCR in resectable lung cancer as a predictive tool for recurrence

   Asmae Gassa[ab1], Jana Fassunke[c1], Sarah Schueten[abd], Luca Kuhlmann[abd], Marie Scherer[bd], Jie Qien[b], Yue Zhao[b], Max Michel[e], Heike Loeser[c], Juergen Wolf[f], Reinhard Buettner[c], Fabian Doerr[a], Matthias Heldwein[a], Lars Hagmeyer[g], Konrad Frank[h], Sabine Merkelbach-Bruse[c], Alexander Quaas[c], Christiane Bruns[b], Khosro Hekmat, Jonathan Weiss, Thorsten Wahlers, HakanAlakus[b2]

   Lung Cancer - Available online 5 November 2020