Blood- and tumor-based analyses for improved prognostics in lung cancer

Isaksson, Sofi

2019

Document Version:
Publisher's PDF, also known as Version of record

Blood- and tumor-based analyses for improved prognostics in lung cancer

SOFI ISAKSSON
FACULTY OF MEDICINE | LUND UNIVERSITY

Printed by Media-Tryck, Lund 2019 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

Department of Clinical Sciences

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2019:52
ISBN 978-91-7619-781-3
ISSN 1652-8220

Blood- and tumor-based analyses for improved prognostics in lung cancer

Blood- and tumor-based analyses for improved prognostics in lung cancer

Sofi Isaksson

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended in B152, LUX, Helgonavägen 3, Lund, Sweden.
May 23, 2019 at 9.00 am.

Faculty opponent
Professor Tom Dønnem, M.D, Ph.D
Department of Clinical Medicine,
UiT The Arctic University of Norway
Department of Oncology,
University Hospital of North Norway
Tromsø, Norway

Organization LUND UNIVERSITY	Document name DOCTORAL DISSERTATION
	Date of issue May 23, 2019
Author(s) Sofi Isaksson	Sponsoring organization

Title and subtitle

Blood- and tumor-based analyses for improved prognostics in lung cancer

Abstract:

Lung cancer is a common cancer type associated with poor prognosis. Even in the group of patients with early-stage tumors, for which curative treatment is possible, lung cancer recurrence is frequent. The aim of this thesis was to investigate the potential prognostic role of mutations in tumor specimens, blood-based proteins and circulating tumor DNA.

In study I, we performed *EGFR* mutation detection by immunohistochemical (IHC) staining in 350 tumors (two cohorts; 298 tumors and 52 tumors, respectively). Mutation detection by IHC staining was not reliable, the antibodies used for *EGFR* mutation-specific staining was sometimes unspecific, and the optimal dilution of the antibodies was not clear. Thirteen of 17 mutations detected by IHC staining could be verified with PCR or Sanger sequencing in the larger cohort.

Mutational spectra was further investigated in study II, where mutational profiling was based on next generation sequencing (NGS) with a panel of 26 solid tumor-related genes. NGS was implemented as part of clinical routine in an autonomous health care region in southern Sweden (Region Skåne) in 2015 and we present the first 1.5-year of clinical NGS-based mutational profiling and its clinical impact. The most frequently mutated genes in this panel were *TP53* and *KRAS*. Mutations in these genes were not associated with progression-free survival (PFS) or overall survival (OS) in patients with advanced lung cancer treated with platinum-based chemotherapy. Targetable alterations in driver oncogenes (for example *EGFR*) were detected in 16% of all tumors and in 59% of tumors from never-smokers.

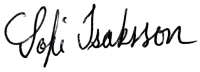
In study III and IV, we investigated tumor markers in serum and cell-free circulating tumor DNA (ctDNA) in plasma obtained pre-operatively from early stage lung cancer patients treated with curatively intended surgery and studied their relation to lung cancer recurrence. In study III, analysis of five tumor markers, implemented in the monitoring of other solid tumors, revealed a tendency of correlation between tumor markers and lung cancer recurrence in a cohort of 107 lung adenocarcinomas. The two most interesting markers were CA 19-9 and CA 125 which were further analyzed in study IV, comprising 58 lung adenocarcinomas, in which we also analyzed pre-operative plasma for ctDNA. With limited amount of plasma, we still detected seven cases with positive ctDNA in pre-operative plasma. Both ctDNA and the combination of ctDNA and tumor markers were associated with worse outcome but also more frequent in higher stage. Thus, larger studies are necessary to study the prognostic potential of these blood-based markers.

In summary, this thesis evaluates treatment-predictive molecular methods and investigates their potential prognostic impact. It presents the mutational spectra in a population-based lung cancer cohort with clinically well-characterized follow up, and suggests a possible role for blood-based prognostication in lung cancer.

Lung cancer, EGFR, NGS, family history, never-smoker, ctDNA, tumor marker

Classification system and/or index terms (if any)

Supplementary bibliographical information		Language English
ISSN and key title 1652-8220		ISBN 978-91-7619-781-3
Recipient's notes	Number of pages 111	Price
	Security classification	

Signature  Date 2019-04-15

Blood- and tumor-based analyses for improved prognostics in lung cancer

Sofi Isaksson
Faculty of Medicine
Department of Clinical Sciences, Lund
Division of Oncology and Pathology

LUND
UNIVERSITY

Supervisor: Associate Professor Maria Planck, M.D, Ph.D
Faculty of Medicine
Department of Clinical Sciences, Lund
Division of Oncology and Pathology

Co-supervisor: Associate Professor Johan Staaf, Ph.D
Faculty of Medicine
Department of Clinical Sciences, Lund
Division of Oncology and Pathology

Co-supervisor: Dr. Mats Jönsson, Ph.D
Faculty of Medicine
Department of Clinical Sciences, Lund
Division of Oncology and Pathology

Co-supervisor: Professor Åke Borg, Ph.D
Faculty of Medicine
Department of Clinical Sciences, Lund
Division of Oncology and Pathology

Coverphoto by Kennet Ruona

Faculty of Medicine
Department of Clinical Sciences Lund
Division of Oncology and Pathology

ISBN 978-91-7619-781-3
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2019

MADE IN SWEDEN

To my grandfather

Table of Contents

Thesis at a glance....11

Studies included in the thesis....12

Study contributions....13

Related studies not included....14

Abbreviations....15

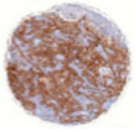
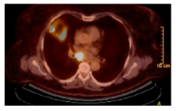
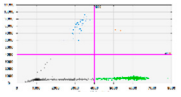
Populärvetenskaplig sammanfattning....17

Introduction and background....23

- Lung cancer research areas....23
- Epidemiology and etiology....24
- Molecular landscape....25
- Histology....30
  - Adenocarcinoma....33
  - Squamous cell carcinoma....34
  - Large cell carcinoma....34
  - Neuroendocrine lung tumors....34
- Diagnostics....35
  - Diagnostic work-up....35
  - Molecular diagnostics....38
- Screening....39
- Staging....39
- Prognostic factors....42
- Treatment....44
  - Surgery....45
  - Radiotherapy....47
  - Chemotherapy....49
  - Immunotherapy....50
  - Targeted therapy....52
  - Treatment options for lung cancer recurrence and progression....55
  - Treatments of patients in this thesis work....56

Aims 57
Patient material 59
Methods 65
Immunohistochemistry, mutated and total EGFR (Study I) 65
Mutation-specific PCR (Study I) 66
Fluorescence in situ hybridization (Study I) 66
Sanger sequencing (Study I) 66
Quantitative real-time PCR (Study I) 67
NanoString technology (Study II) 67
Next generation sequencing (Study II, III, IV) 67
Electrochemiluminescence immunoassay (Study III, IV) 69
Droplet digital PCR (Study IV) 69
Statistical methods 70
Limitations 73
Results and discussion 75
Study I and II: Mutational profiling and clinical outcome 75
Study III and IV: Blood-based markers in early stage lung cancer 86
Conclusions 93
Future perspectives 95
Acknowledgement 97
References 99

Thesis at a glance

Study		Question	Patients and material	Methods	Results and conclusion
I		Is *EGFR* mutation detection with IHC feasible? Are *EGFR* mutations, *EGFR* gene copy number and EGFR protein expression correlated?	• Retrospective consecutive patient cohorts: 350 surgically treated patients with early stage lung cancer • Tumor specimens	• IHC • PCR methods • DNA sequencing • FISH	IHC is not reliable for *EGFR* mutation detection. EGFR expression was common and associated with increased gene copy number.
II		How is the Region Skåne lung cancer mutation spectrum and does it correlate with clinical data?	• 599 patients, stage I-IV, in clinical routine analysis for mutational profiling using NGS • Tumor specimens/cytology	• NGS • NanoString technology	92 % of the tumors had at least one mutation. Mutations in *TP53* and *KRAS* were most frequent but did not display any clear correlation with chemotherapy response.
III		Do tumor markers in serum correlate with lung cancer relapse?	• 107 surgically treated patients with lung adenocarcinoma • Pre-operative serum, tumor specimens	• Electrochemiluminescence	Positive finding of at least two tumor markers seems to correlate with worse prognosis in patients with surgically treated early stage lung cancer.
IV		Do serum tumor markers and plasma cell-free circulating tumor DNA correlate with lung cancer relapse?	• 58 surgically treated patients with lung adenocarcinoma • Pre-operative plasma and serum, tumor specimens	• NGS (tumors) • Droplet digital PCR (plasma) • Electrochemiluminescence immunoassay (serum)	Circulating tumor DNA and/or tumor markers might correlate with worse prognosis in early stage patients but larger volume of plasma and serial monitoring are necessary in further studies.

Studies included in the thesis

I. Detecting EGFR alterations in clinical specimens - pitfalls and necessities. ***Isaksson S****, Bendahl PO, Salomonsson A, Jönsson M, Haglund M, Gaber A, Jirström K, Jönsson P, Borg A, Johansson L, Staaf J, Planck M.* Virchows Archiv. 2013 Dec;463(6):755-64.

II. Next generation sequencing based mutation analysis in lung cancer – first 1.5-year experience. ***Isaksson S****, Hazem B, Jönsson M, Reuterswärd C, Karlsson A, Griph H, Engleson J, Oskarsdottir G, Öhman R, Holm K, Rosengren F, Annersten K, Jönsson G, Borg Å, Edsjö A, Levéen P, Brunnström H, Lindquist KE, Staaf J, Planck M.* Manuscript 2019

III. CA 19-9 and CA 125 as potential predictors of disease recurrence in resectable lung adenocarcinoma. ***Isaksson S****, Jönsson P, Monsef N, Brunnström H, Bendahl PO, Jönsson M, Staaf J, Planck M*. PLoS One. 2017 Oct 19;12(10).

IV. Pre-operative plasma cell-free circulating tumor DNA and serum protein tumor markers as predictors of lung adenocarcinoma recurrence. ***Isaksson S****, George AM, Jönsson M, Cirenajwis H, Jönsson P, Bendahl PO*, *Brunnström H, Staaf J, Saal LH, Planck M.* Acta Oncologica 2019 (Accepted for publication)

Study contributions

My contributions to each study were as follows:

I. I performed PCR, direct DNA sequencing and all the immunohistochemical evaluations. I was responsible for the writing and revision of the manuscript.

II. I was involved in study design and application for ethical approval. I collected the clinical information, performed the statistical analyses and wrote the manuscript.

III. I collected the clinical information and performed the statistical analyses. I was responsible for the writing and revision of the manuscript.

IV. I participated in study design, performed DNA extraction from plasma, and participated in the laboratory work of droplet digital PCR analysis. I collected the clinical information, wrote the manuscript and participated in the revision of the manuscript.

Related studies not included

Relation between smoking history and gene expression profiles in lung adenocarcinomas. *Staaf J, Jönsson G, Jönsson M, Karlsson A,* ***Isaksson S****, Salomonsson A, Pettersson HM, Soller M, Ewers SB, Johansson L, Jönsson P, Planck M.* BMC Med Genomics. 2012 Jun 7; 5:22.

Landscape of somatic allelic imbalances and copy number alterations in human lung carcinoma. *Staaf J,* ***Isaksson S****, Karlsson A, Jönsson M, Johansson L, Jönsson P, Botling J, Micke P, Baldetorp B, Planck M.* Int J Cancer. 2013 May 1;132(9):2020-31.

Identification of transcriptional subgroups in EGFR-mutated and EGFR/KRAS wild-type lung adenocarcinoma reveals gene signatures associated with patient outcome. *Planck M,* ***Isaksson S****, Veerla S, Staaf J.* Clin Cancer Res. 2013 Sep 15;19(18):5116-26.

Histological specificity of alterations and expression of KIT and KITLG in non-small cell lung carcinoma. *Salomonsson A, Jönsson M,* ***Isaksson S****, Karlsson A, Jönsson P, Gaber A, Bendahl PO, Johansson L, Brunnström H, Jirström K, Borg Å, Staaf J, Planck M.* Genes Chromosomes Cancer. 2013 Nov;52(11).

Genomic and transcriptional alterations in lung adenocarcinoma in relation to EGFR and KRAS mutation status. *Planck M, Edlund K, Botling J, Micke P,* ***Isaksson S****, Staaf J.* PLoS One. 2013 Oct 24;8(10)

Abbreviations

AC	Adenocarcinoma
ALK	Anaplastic Lymphoma Kinase
AMP	Association for Molecular Pathology
ASCO	American Society of Clinical Oncology
ATP	Adenosine Triphosphate
BRAF	Rapidly Accelerated Fibrosarcoma kinase B
BRCA	Breast Cancer Early Onset Gene
CA 125	Cancer Antigen 125
CA 19-9	Carbohydrate Antigen 19-9
CAP	College of American Pathologists
CEA	Carcinoma Embryonic Antigen
CEP7	Centromere Probe of chromosome 7
CI	Confidence Interval
CNS	Central Nervous System
CRT	Chemoradiotherapy
CT	Computed Tomography
CTLA-4	Cytotoxic T-Lymphocyte-associated Antigen 4
cfDNA	Cell-free DNA
ctDNA	Circulating tumor DNA
ddPCR	Droplet Digital PCR
DNA	Deoxyribonucleic Acid
DFS	Disease-Free Survival
EBUS	Endobronchial Ultrasound
EGFR	Epidermal Growth Factor Receptor
EML4	Echinoderm Microtubule-Associated Protein-Like 4
FDA	Food and Drug Administration
FDG-PET	Fluorodeoxyglucose-Positron Emission Tomography
FFPE	Formalin Fixed Paraffin Embedded
FISH	Fluorescence in situ Hybridization
Gy	Gray
H&E	Hematoxylin and Eosin
HE4	Human Epididymis Protein 4

HPV	Human Papillomavirus
HR	Hazard Ratio
IASLC	International Association for the Study of Lung Cancer
IHC	Immunohistochemistry
irAE	Immune Related Adverse Event
LCC	Large Cell Carcinoma
LDCT	Low-Dose Computed Tomography
LCNEC	Large Cell Neuroendocrine Carcinoma
MET	Mesenchymal Epithelial Transition factor
mTOR	Mammalian Target of Rapamycin
NGS	Next Generation Sequencing
NOS	Not Otherwise Specified
NSCC	Non-Small Cell Cancer
NSCLC	Non-Small Cell Lung Cancer
NSE	Neuron-Specific Enolase
NTRK	Neurotrophic Receptor Tyrosine Kinase
OS	Overall Survival
PCR	Polymerase Chain Reaction
PD-1	Programmed Death 1
PD-L1	Programmed Death Ligand 1
PFS	Progression-Free Survival
PI3K	Phophoinositide 3-Kinase
PS	Performance Status
RCT	Randomized Controlled Trial
RECIST	Response Evaluation Criteria In Solid Tumors
RET	Rearranged during transfection
RFI	Recurrence-Free Interval
RNA	Ribonucleic Acid
ROS1	C-Ros oncogene 1
SBRT	Stereotactic Body Radiotherapy
SNP	Single Nucleotide Polymorphism
SqCC	Squamous Cell Carcinoma
TKI	Tyrosine Kinase Inhibitor
TMA	Tissue Microarray
TMB	Tumor Mutational Burden
TNM	Tumor Node Metastasis
VAF	Variant Allele Frequency
WHO	World Health Organization

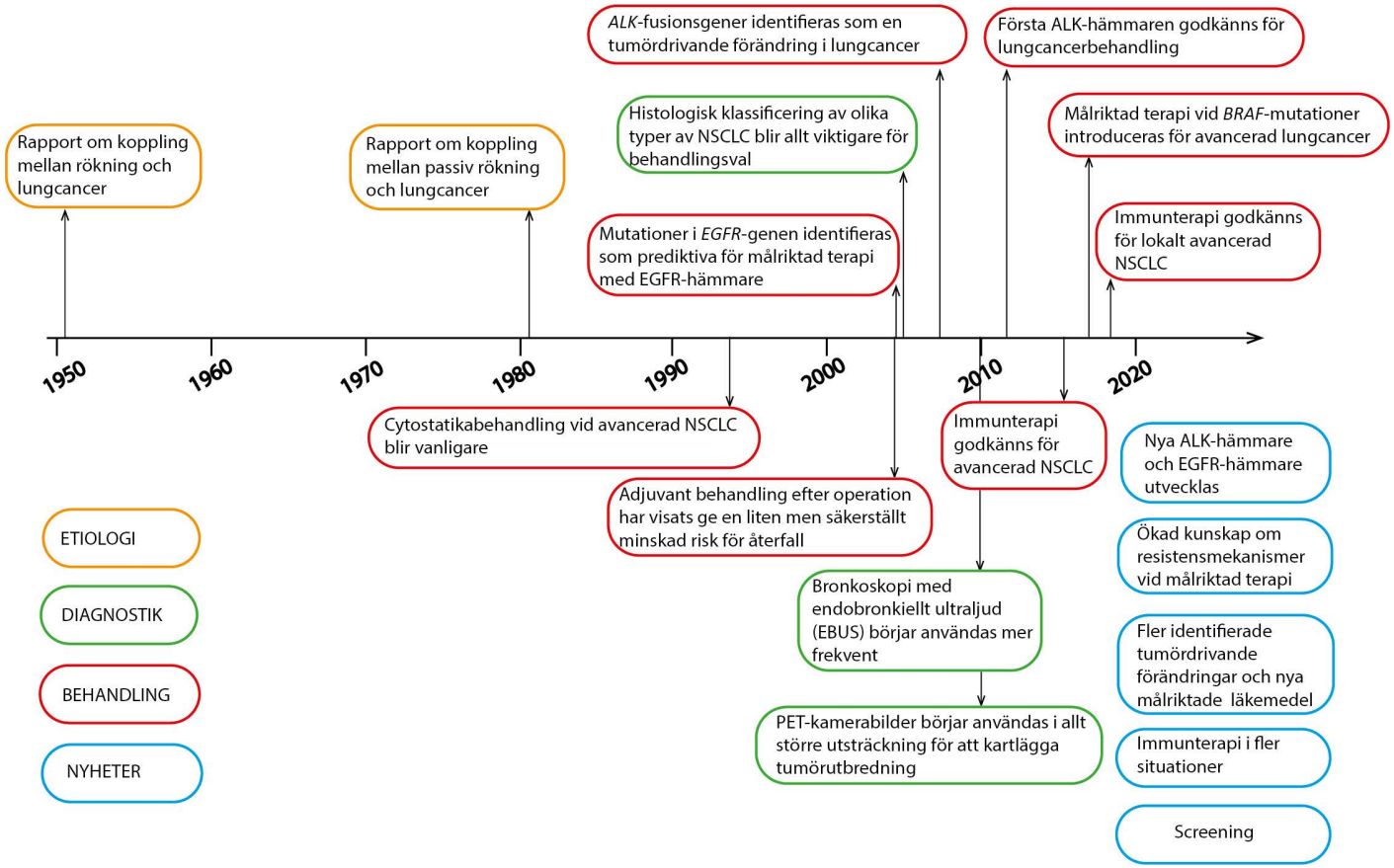
Populärvetenskaplig sammanfattning

Den brittiske kungen George VI drabbades av en elakartad lungtumör och genomgick därför en stor lungoperation 1951, en så kallad pulmektomi vilket innebär att en lunga opereras bort. Som många andra i Europa och USA efter andra världskriget var kungen aktiv rökare. Bara året innan kungen opererades hade den första rapporten om en koppling mellan rökning och lungcancer publicerats (Doll och Hill, British Medical Journal). Idag anses rökning vara en bakomliggande orsak till omkring 80% av alla lungcancerfall och även om antalet dagliga rökare i Sverige liksom i andra västländer har minskat sedan kung Georges insjuknande är en stor andel av lungcancerdrabbade tidigare rökare och prevalensen daglig-rökare i andra delar av världen hög. Årligen får omkring 1,8 miljoner människor globalt en lungcancerdiagnos och i Sverige är det ca 4000 människor som drabbas varje år. Femårsöverlevnaden är omkring 20% och den dåliga prognosen beror till stor del på att en betydande andel av alla lungcancerfall diagnosticeras i ett skede då cancern har spridit sig.

Lungcancer delas in i två huvudgrupper beroende på tumörernas egenskaper, icke-småcellig lungcancer (non-small cell lung cancer, NSCLC) och småcellig lungcancer (small cell lung cancer, SCLC) och behandlingarna skiljer sig åt. Icke-småcellig lungcancer utgör den större andelen av lungtumörer, ca 75%, och kan delas in ytterligare i undergrupper. Den här avhandlingen berör den stora gruppen icke-småcellig lungcancer, både opererade patienter och patienter med spridd lungcancer.

Kung Georges hälsa fortsatte att försämras och han avled ett par månader efter operationen. Idag ser både diagnostik och behandlingsmöjligheter naturligtvis annorlunda ut men länge var operation, om det var möjligt, den enda behandlingen som stod till buds för många som diagnosticerades med lungcancer. De största kliven framåt har framförallt skett under 2000-talet. En sammanfattning av viktiga framsteg inom lungcancerfältet med fokus på NSCLC är illustrerat i en tidslinje (se Figur 1).

Behandlingsmöjligheterna, representerade av de röda boxarna i tidslinjen, vid framförallt avancerad ej operabel lungcancer som än så länge utgör den största andelen av nydiagnostiserade lungcancerfall, har utökats mycket de senaste 15 åren. Vid spridd lungcancer rör det sig i nuläget om palliativa behandlingar som syftar till symtomlindring, bromsad tumörtillväxt och förlängd överlevnad. Ett steg framåt



Figur 1. Viktiga milstolpar inom lungcancerområdet.

togs på 90-talet då studier påvisade en förbättrad överlevnad för patienter med spridd lungcancer som fick cytostatika jämfört med de som enbart fick symtomlindring. Den största utvecklingen inom behandling av spridd lungcancer har dock skett genom så kallad målriktad terapi och senare immunterapi. Målriktad terapi (targeted therapy) hämmar specifika tumörprotein som ofta är starkt tumördrivande. Sådana defekta proteiner kan bildas genom mutationer, d.v.s. förändringar i tumörernas DNA-sekvens (arvsmassa), eller genom fusionsgener, vilket innebär att två gener i tumörens DNA smält samman. Med dagens kunskap är det hos en minoritet av lungcancerpatienterna vi finner sådana förändringar och kan ge målriktad terapi. Två viktiga milstolpar, *EGFR*-mutationer och *ALK*-fusionsgener ses på tidslinjen och de senaste åren har man identifierat fler drivande tumörförändringar (ex. i *BRAF*, *ROS1*, *RET*) mot vilka det finns läkemedel eller läkemedel under utveckling. Det senaste tillägget till behandling av spridd lungcancer är immunterapi som påverkar kroppens immunförsvar till att bekämpa cancercellerna. Riktigt vilka patienter som svarar bäst på immunterapi vet man inte. Som ses på tidslinjen har indikationen nyligen breddats och kommer säkert att ändras ytterligare kommande år.

Botande behandling av icke-småcellig lungcancer handlar fortfarande främst om kirurgi men även strålbehandling av små tumörer och kombination av strålning och cytostatika. Hade kung George opererats idag hade han utretts inför operationen med skiktröntgen och säkert även med PET-kamera och med ultraljudsledd provtagning av lymfkörtlar i thorax, vilket ofta görs inför operation för att kartlägga tumörspridning.

Utvecklingen inom lungcancer, framförallt vad gäller behandlingsmöjligheterna vid spridd lungcancer, har således gått fort de senaste åren och studier pågår kring nya behandlingskombinationer och behandlingsmöjligheter.

Detta avhandlingsarbete utgörs av fyra delarbeten med det övergripande målet att utvärdera möjliga prognostiska faktorer vid lungcancer. Ökad kunskap om prognostiska faktorer kan på sikt leda till förbättrad stratifiering av lungtumörer med möjlighet till mer individualiserad behandling och uppföljning.

Delarbete I och II berör mutationsdiagnostik, EGFR-förändringar och kliniska korrelationer. I delarbete I undersöks *EGFR*-mutationer, ökat antal genkopior av *EGFR* och uttryck av EGFR-proteinet i opererade lungtumörer i en retrospektiv kohort bestående av opererade lungcancerpatienter. I detta projekt testades mutationsdiagnostik av de två vanligaste mutationstyperna i *EGFR*-genen genom immunhistokemisk färgning och vidare analyserades ett bredare spektrum av *EGFR*-mutationer genom andra metoder och uttryck av EGFR-proteinet och ökat

kopietal av *EGFR*-genen. Immunhistokemisk färgning är rutin inom patologin för diagnostik av bland annat vilken vävnadstyp lungtumörer har men visade sig i vårt arbete inte vara pålitligt för EGFR-mutationsdiagnostik.

Delarbete II handlar liksom delarbete I om mutationstestning. I januari 2015 infördes mutationstestning med så kallad next generation sequencing (NGS), vilket innebär att många olika gener i tumörer från flera patienter kan undersökas samtidigt. Under den första tiden sekvenserades delar av 26 gener, däribland *EGFR* som är huvudsyftet med mutationstestningen. Under första 1.5 året som NGS användes i klinisk rutin i Region Skåne för mutationstestning av lungtumörer sekvenserades tumörer från över 600 lungcancerpatienter. Sammanlagt hade 599 patienter med lungcancer, exkluderat SCLC, ett konklusivt mutationsresultat och i delarbete II samlades kliniska data in för de 599 patienterna och sammanställdes tillsammans med sekvenserings-resultaten. I drygt 90% av tumörerna kunde minst en mutation i någon av de undersökta generna hittas. Mutationsfynden såg olika ut i olika undergrupper av lungtumörer och det var främst i gruppen lungcancer av adenokarcinomtyp som mutationer detekterades. De vanligaste mutationerna fanns i generna *TP53* och *KRAS* men någon tydlig koppling mellan mutationer i dessa gener och behandlingssvar eller progressionsfri överlevnad (tid under och efter behandlingen då tumören inte växer) hos cytostatikabehandlade patienter med avancerad lungcancer kunde inte ses. Vidare visar den kliniska sammanställningen att eventuell exponering och hereditet ofta inte efterfrågas. Däremot var uppgifter om rökning oftast väl dokumenterat och vi såg att hos lungcancerpatienter som aldrig rökt var mutationer som det finns målriktad behandling mot, vanligt. Exempelvis hittades *EGFR*-mutation i 44% av dessa tumörer.

Delarbete III och IV berör den mindre gruppen av opererade lungcancerpatienter och deras risk för lungcanceråterfall. Trots tilläggsbehandling med cytostatika efter operation så är risken för opererade patienter att drabbas av återfall hög. I delarbete III undersöks fem tumörmarkörer i serum tillvarataget innan operation från drygt 100 opererade patienter. Serumanalyserna utfördes vid Klinisk Kemi, Skånes Universitetssjukhus, där de undersökta tumörmarkörerna sedan tidigare används som del i den kliniska handläggningen av andra tumörformer än lungcancer. Det sågs en tendens till koppling mellan recidiv och tumörmarkörer, i synnerhet för två av markörerna, CA 19-9 och CA 125, men kohorten är liten och resultaten ska därför tolkas försiktigt.

Delarbete IV bygger vidare på delarbete III, och de 58 inkluderade patienterna är delvis överlappande med patienterna i delarbete III, med målet att undersöka blodbaserade markörer och deras koppling till lungcanceråterfall: Tumör-markörerna CA 125 och CA 19-9 samt fritt cirkulerande, cell-fritt tumör-DNA undersöktes i preoperativa blodprover från patienterna. Fritt cirkulerande tumör-DNA (ctDNA) utgör en liten andel av allt fritt cirkulerande DNA i blodet och kräver

känsliga analysmetoder för att detekteras. Trots att det här var tidigt diagnostiserade tumörer kunde vi detektera ctDNA i blodprover innan operation i sju fall varav sex senare drabbades av lungcancerrecidiv. Vi kunde också, i likhet med resultaten i delarbete III, se en trend gentemot sämre prognos om CA 19-9 eller CA 125 var förhöjt preoperativt. För att verkligen utröna om dessa blodbaserade markörer kan nyttjas i kliniken för att identifiera opererade patienter med ökad risk att drabbas av recidiv krävs större studier.

Sammanfattningsvis utgörs denna avhandling av fyra delarbeten med frågeställningar direkt kopplade till kliniska utmaningar och med det övergripande målet att individualisera och förbättra omhändertagandet av lungcancerpatienter baserat på resultat från tumör- och blodprover.

Introduction and background

Lung cancer research areas

Lung cancer is a heterogenous disease and lung cancer research is a rapidly evolving field. Some of the important, major lung cancer research areas and unmet challenges, in which increased knowledge are highly necessary, are listed below.

- *Etiology*: Increased knowledge about etiology is highly relevant to prevent lung cancer. An end to tobacco use is the most urgent and important goal. However, almost a fifth of lung cancer patients are never smokers and more research is highly needed in order to improve lung cancer prevention.
- *Diagnosis/early detection*: Prognosis of lung cancer is dependent on early diagnostics. A screening program might be introduced shortly and easily accessible markers e.g. in blood or sputum could potentially play an important complementary role to radiology and clinical examination in early diagnosis.
- *Prognostics*: Despite adjuvant therapy or very low tumor burden many surgically treated lung cancer patients are diagnosed with recurrence, which make stratification beyond stage necessary to further individualize treatment and follow-up.
- *Treatment prediction*: Through recent years, many new drugs have been approved for lung cancer and, with new targets being identified, an increasing number of approved drugs are assumed. Individualized treatment requires tumor- and/or patient-related factors that predict response or non-response. For immunotherapy and chemotherapy, treatment predictive variables are insufficient or lacking.
- *Resistance mechanisms*: The rapid evolution of new drugs for lung cancer requires more knowledge about mechanisms of resistance behind disease progression during therapy.

The studies in this thesis focus on the potential of blood- and tumor-based markers to identify patients at high risk of recurrence after curative surgery and in prognostication of advanced disease, respectively. Tumors studied are primarily non-small cell lung cancer (NSCLC), in particular adenocarcinoma (AC) and squamous cell carcinoma (SqCC). Other types of lung cancer are therefore only briefly mentioned in the introduction and background. Study I and II focus on mutational profiling of tumors, and on the mutations' impact on patient outcome. The first study includes early-stage surgically treated lung cancer patients and study II involves all stages. Blood-based markers and their potential association with patient outcome are investigated in study III and IV. Both studies include surgically treated early-stage lung cancer patients.

Epidemiology and etiology

Globally is lung cancer not only the most common cancer but also the deadliest form of cancer with an estimated 2.1 million new cases and 1.8 million deaths in 2018[1]. Cigarette smoking, initially described as associated with lung cancer in the 1950s[2], contributes to approximately 80% or more of lung cancers in the Western population[1].

A more recent chapter in the long history of tobacco consumption is electronic cigarettes (e-cigarettes) with vaporized tobacco, originally produced as a smoking cessation device. E-cigarettes are considered less harmful than traditional cigarettes. However, long term effects and other aspects, i.e. potential harm from flavoring chemicals, remain to be investigated[3,4].

Several other lung cancer risk factors, apart from tobacco smoke have been identified. Studies have shown evidence of an increased risk of lung cancer also by environmental cigarette smoke, radon through alpha particles emitted from radon decay, air pollution, arsenic and asbestos. [5-8]. Chronic obstructive pulmonary disease (COPD) is smoking-related, like lung cancer, and an association between COPD and lung cancer is supposed to be largely attributed to the confounding of smoking[9]. Many studies have investigated a possible role for oncogenic infections in lung cancer, for example certain types of human papillomavirus (HPV), which are known carcinogens and highly relevant in cervical cancer and oropharyngeal cancer. The possible role of HPV infections in lung carcinogenesis are not fully known and studies have shown divergent prevalence of the HPV in lung tumors [10,11].

While lung cancer is generally associated with tobacco smoking and other environmental factors, heredity might play a prominent role in a small proportion of lung cancer cases, possibly associated with germline mutations interfering with

DNA repair, such as mutations in *BRCA* or *TP53* [12-14]. A germline mutation in exon 20 of the epidermal growth factor receptor gene (*EGFR*), resulting in the point mutation T790M, has also been shown to predispose to lung cancer, often with a second activating mutation in *EGFR* [15,16].

Molecular landscape

The transition from normal cells to cancer cells is a multistep procedure in which the tumor cells stepwise alter. Principles that characterize tumorigenesis can be organized in the hallmarks of cancer, described by Weinberg and Hanahan[17]. These principles, initially six but extended in the up-dated publication from 2011, are briefly summarized below.

1) **Sustaining proliferative signaling**. Mutations in *EGFR* or *BRAF* can lead to constitutively active growth factor receptors triggering pathways promoting cell growth.
2) **Evading growth suppressors.** TP53 normally functions as a suppressor that inhibits cell growth or promotes apoptosis when receiving signals of abundant stress or DNA damage. Alterations of TP53 might enable the tumor to evade this normal function.
3) **Resisting cell death.** Tumors can escape cell death by loss of TP53 function or by upregulate expression of antiapoptotic regulators.
4) **Enabling replicative immortality.** In normal conditions a cell's ability to pass cell- growth- and division cycles is limited. The chromosomes are flanked by telomeres, a repeated nucleotide sequence, that shortens for each division, and the chromosome becomes vulnerable to end-to-end fusions and ultimately leading to apoptosis for the cell. The enzyme telomerase might be upregulated in tumors and as it extends the telomeres the cells become resistant to senescence and apoptosis.
5) **Inducing angiogenesis.** Tumors need vessels to grow and can stimulate endothelial cells to format new vessels by expressing proangiogenic signals through ligands such as vascular endothelial growth factor (VEGF) or fibroblast growth factor (FGF).
6) **Activating invasion and metastasis.** By alterations in shape, attachment to other cells and the extracellular matrix, for example by loss of E-cadherin, tumor cells may exhibit invasion and metastasis.
7) **Genome instability and mutation.** Genomic instability, e.g. by alterations in genes involved in DNA repair, results in genetic alterations that promotes tumor progression.

8) **Tumor-promoting inflammation.** Inflammation can benefit several hallmark capabilities by supplying factors, for example growth factors in the tumor microenvironment leading to sustained proliferative signaling.
9) **Reprogramming of energy metabolism.** Tumors might reprogram their cellular energy metabolism, skipping the mitochondrial oxidative phosphorylation, yielding less efficient ATP production but compensating with an increased glucose uptake. The rationale for this is not fully understood but one hypothesis depicts that glycolytic intermediates might be used in synthesis of nucleosides and amino acids, crucial for developing new cells.
10) **Evading immune destruction.** Tumors are presumed to avoid being destructed by the immune system in various ways, including secreting immunosuppressive factors.

There are widespread differences of genetic alterations between subtypes of NSCLC and between smokers and non-smokers[18]. Cigarette smoke contains thousands of compounds, of which >60 classifies as carcinogens, and leads to DNA damage by radical oxygen species or DNA adducts. The transversion type of mutations, in for example the *TP53* or kirsten rat sarcoma viral oncogene homolog (*KRAS*) genes, are enriched in tumors of smokers[19,20]. Furthermore, the mutagenic effects from smoking contribute to the very high rate of mutations per megabase in especially squamous cell carcinoma (SqCC), which is strongly associated with smoking. While all lung cancer types are associated with smoking, adenocarcinomas (AC) are the most frequent histology among never-smokers diagnosed with lung cancer[18]. Mutational profiling of AC and SqCC reveal frequent *KRAS* mutations in AC (approximately 30% in Caucasian populations and 5-10% in Asian populations) but rarely in squamous cell carcinomas. Also associated with AC are targetable *EGFR* mutations which are rarely detected in SqCC. Tumor suppressor gene *TP53* mutations are frequent in both AC and SqCC but very prominent in SqCC. Other mutated genes in SqCC are *PIK3CA*, *PTEN*, *KEAP1* and *RB1*. *CDKN2A* is a tumor suppressor gene inactivated in a majority of SqCC cases, through mutations. Other genetic features prominent in SqCC are gene copy number alterations, for example gain or amplification of chromosome 8p (involving *CDKN2A*) or deletion of chromosome 9p (involving *CDKN2A*)[18,21,22].

Oncogenic driver alterations

Some of the above-mentioned mutations are oncogenic driver alterations, which refer to genetic alterations that are essential for tumor-cell survival by initiating and maintaining tumorigenesis - vital functions for the tumors that are dependent on these oncogenic drivers, a phenomenon called oncogene addiction. Oncogenic driver alterations are typically mutually exclusive and involving genes relevant for proliferation and survival. Many of them are primarily found in tumors of AC histology and patients with sparse or no smoking[18,23].

In this section, some important driver mutations and gene rearrangements are described. Some of these altered proteins are actionable, with drugs either approved or under investigation, which is described in the Treatment section (Targeted therapies).

EGFR mutations, copy number alterations and protein expression of EGFR were the focus of study I, whereas in study II, multiple mutations, including variants in *EGFR*, *KRAS*, the B-rapidly accelerated fibrosarcoma gene (*BRAF*), mesenchymal epithelial transition factor (*MET*), are described together with clinical aspects such as prognosis. For study IV, surgically removed early stage tumors with mutations in *EGFR*, *BRAF* or *KRAS* were included. However, the mutations themselves were not the main focus, but rather used as a footprint of the tumor when analyzing plasma for cell-free circulating mutant DNA originating from the tumor cells.

EGFR is a transmembrane receptor encoded belonging to a tyrosine kinase family of four receptors; EGFR (ERBB1, HER1), ERBB2 (HER2), ERBB3 (HER3) and ERBB4 (HER4). Ligand binding to a receptor leads to homo-or hetero-dimerization of receptors and subsequent phosphorylation of the tyrosine kinase domain located intracellularly with phosphate derived from ATP bound within the tyrosine kinase domain. This activation triggers a signaling cascade through several pathways, including the RAS/RAF/MEK pathway and phosphoinositide 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR) pathway, ultimately resulting in cell proliferation, cell survival, cell motility and cell invasion[23,24].

EGFR mutations in exon 18-21, encoding part of the tyrosine kinase domain, are presumed to occur early in lung cancer development and lead to constitutive EGFR activation. The EGFR signaling pathways are illustrated in Figure 2. Many of these mutations, in particular those in exon 18, 19 and 21, are targetable with tyrosine kinase inhibitors (TKI) as described in the section Targeted therapy. Mutations in exon 20 typically confer resistance to EGFR TKI. The prevalence of *EGFR* mutations in NSCLC vary with populations studied, approximately 10-15% in Europe and about 40% in Asia, and generally higher prevalence among never-smokers, females and NSCLC with AC histology[22,24-26].

Activating mutations in *ERBB2* (*HER2*) have been identified in a smaller subset of AC, with clinical features similar to those seen in *EGFR*-mutated lung cancer[18,22].

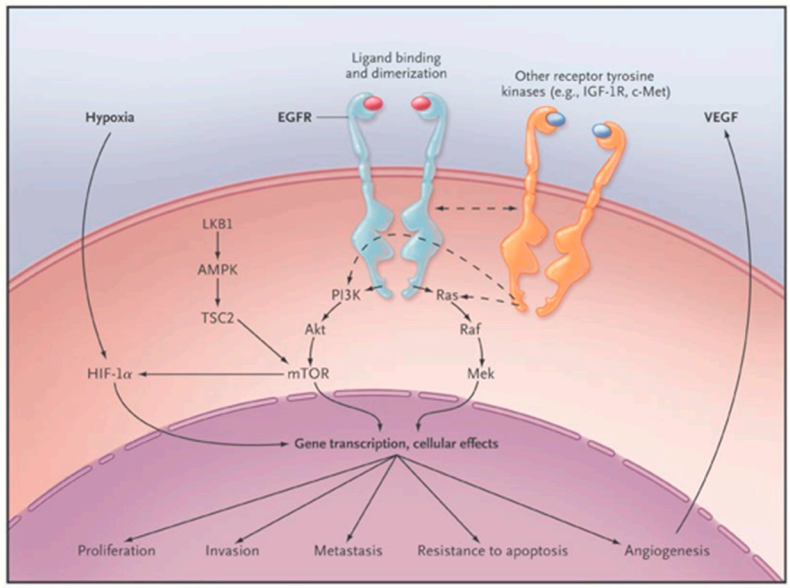


Figure 2. EGFR activation and cell signaling pathways. Activated EGFR acts through several pathways, including Ras/Raf/Mek and PI3K/Akt/mTOR and the intracellular signaling regulating several tumorigenic processes. Reproduced with permission from[24] Copyright Massachusetts Medical Society.

KRAS encodes the protein KRAS, downstream of EGFR in the RAS/RAF/MEK pathway (Figure 2). Mutations in *KRAS* are typically associated with AC histology, mutually exclusive from *EGFR* mutations, and often smoking-related based on G to T transversions affecting exon 12 or 13[22,24]. Mutations in *NRAS*, encoding a GTPase related to KRAS, are much less frequent than *KRAS* mutations but also suggested to define a small subset of oncogene addicted lung cancer that seems to correlate with smoking[22,27].

Downstream of KRAS in the MAPK signaling pathway is another interesting drug target in NSCLC, BRAF. Activating mutations of the *BRAF* gene, encoding the kinase BRAF downstream of RAS in the RAS/RAF/MEK/ERK signaling pathway regulating cell growth, is present in 2-10% of AC. There is no strong correlation to absence of smoking history, in contrast to *EGFR* mutations. The *BRAF* V600E mutation is of specific interest since it is targetable by a combination of BRAF and MEK inhibiton[18,28,29].

Activation of RAS and RAF follows by activation of mitogen-activated protein kinase 1 (MAP2K1/MEK1). Mutations in *MAP2K1* have been identified as oncogene drivers in AC[22].

In the pathway consisting of PI3K, AKT and mTOR, involved in proliferation and cell survival, activating alterations in *PIK3CA*, which are associated with SqCC, are suggested to act as oncogenic drivers[21,24].

MET encodes the hepatocyte growth factor (HGF) receptor, a receptor tyrosine kinase involved in regulating development and cell growth activated through ligand binding and subsequent homodimerization and phosphorylation of intracellular tyrosine residues. The MET pathway might be dysregulated in a variety of mechanisms in lung cancer. *MET* exon 14-skipping mutations define a subset of oncogene addicted lung cancers and have been identified in 2-4% of NSCLC. In the regular process, introns flanking exon 14 are spliced out and mRNA containing exon 14 is subsequently translated into a MET receptor. A wide variety of *MET* exon 14 skipping alterations have been identified, including mutations that disrupt the splice sites flanking exon 14 leading to aberrant splicing and mRNA lacking exon 14 (Figure 3). Exon 14 encodes a binding site of ubiquitin ligase and the outcome of the aberrant splicing is a MET protein with increased stability and decreased degradation[30-33].

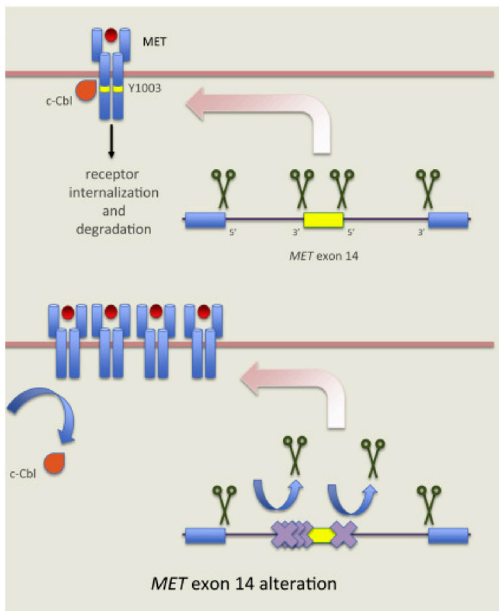


Figure 3. Aberrant splicing of *MET* resulting in decreased degradation of MET. Top: Normal splicing with exon 14, holding the binding site for ubiquitin ligase, complete. Bottom: Mutations disrupting the splice sites flanking exon 14 result in decreased degradation of MET and increased signalling.
Reprinted with permission from Elsevier[32].

Several tumorigenic, and targetable, gene rearrangements/translocations in NSCLC, primarily AC, have been identified. The anaplastic lymphoma kinase (*ALK*) gene on chromosome 2p23 encodes a receptor tyrosine kinase acting in, e.g., RAS-MAPK activated pathways. *ALK* translocations are derived from fusion of a part of the *ALK* gene comprising the tyrosine kinase domain with a partner gene, which results in constitutive tyrosine kinase activation with subsequent uncontrolled proliferation and survival of the tumor cells. The breakpoint within *ALK* is highly conserved but multiple fusion partners are identified, with the major partner being the echinoderm microtubule-associated protein-like 4 (*EML4*) gene on chromosome 2p21. *ALK* rearrangements are detected in approximately 5% of NSCLC[34]. Other fusion genes in NSCLC with similar fusion patterns (a proto-oncogene with a tyrosine kinase domain) include for example c-ros oncogene 1 (*ROS1*), rearranged during transfection (*RET*) and neurotrophic receptor tyrosine kinase (*NTRK*)[18].

Histology

Non-small cell lung cancer (NSCLC) accounts for the majority (approximately 75%) of the lung cancer diagnoses, further subtyped into the main histological types adenocarcinoma (AC), squamous cell carcinoma (SqCC) and large cell carcinoma (LCC)[35]. Large cell neuroendocrine cancer (LCNEC) was previously defined as a variant of LCC but rather share features with the other neuroendocrine tumor types, i.e. small cell lung cancer (SCLC), which is the most frequently occurring neuroendocrine lung cancer, and the less frequent, rarely aggressive, carcinoids.

The current histological classification of lung cancer refers to the 2015 WHO classification. Immunohistochemistry (IHC) has a more crucial role in this classification compared to the prior 2004 WHO classification, in which morphology with hematoxylin and eosin (H&E) staining alone was the common diagnostic procedure [18]. Histological subtype became more important in the renewed classification due to chemotherapy regimens with different effect on AC and SqCC, respectively, and due to treatment predictive mutations being more associated with AC. Since the majority of lung cancer patients are not operable, biopsies and cytology are important in lung cancer diagnosis, but the prior 2004 WHO classification addressed resected specimens. A publication from 2011 which was initially settled to update AC classification, also included a proposed classification for non-resection specimen including both AC and other histological types to meet the need of refined diagnostics[36].

A suggested work-up of small biopsies or cytology, i.e. non-resectable tumors, in the 2015 WHO classification starts with morphology. If morphology reveals a clear AC or SqCC further diagnostic work is generally not necessary. Immunohistochemical staining and mucin staining can aid in the further diagnostic of NSCLC (in WHO classification termed non-small cell cancer; NSCC) not otherwise specified (NOS). If no clear staining for AC or SqCC markers the tumor is classified as NOS but if positive AC markers and negative SqCC markers the tumor is suggested to be AC while the opposite result of staining suggests SqCC. A specimen that displays features and/or IHC of both AC and SqCC respectively, might be an adenosquamous tumor which should then be stated although the definitive diagnosis can only be established in a resected tumor in which two distinct cell populations with ≥ 10% of each component can be confirmed. Morphology, sometimes in combination with IHC, can also reveal SCLC or suspected LCNEC[18]. Swedish guidelines suggest that at least TTF-1 (or napsin A) are performed even in cases when AC morphology is clear as a positive result strongly suggests lung origin. In the case of a poorly differentiated NSCLC, a minimal panel consisting of p40 (or CK5) and TTF-1 (or napsin A) and possibly mucin staining is suggested. If neuroendocrine morphology is present, neuroendocrine marker are added[37]. The histological pictures of the main non-endocrine subtypes AC, SqCC and LCC with typical corresponding features are summarized in Figure 4.

The tumors studied in this thesis work are NSCLC, except for a few cases of SCLC in study I with surgically treated patients of varied histological subtypes. In study I, the tumors are classified according to 2004 WHO classification[38]. Study II includes all NGS-tested patients in Region Skåne for 1.5 year, diagnosed with primary lung cancer or lung cancer recurrence/progression 2015-2016, and includes mainly NSCLC but also a few LCNEC whereas three SCLC were excluded. The proposed histological types for mandatory mutational profiling have been changing over time. The majority of the patients in study II had a tumor of AC histology, which reflects the mutational profiling being performed mainly in AC and non-squamous histology on a routine basis, although SqCC were also present. Many lung cancer cases are subject to NGS test early during investigation, with the analyses thus performed on small tumor samples sometimes insufficient for a definitive histological diagnosis insufficient for certain classifications like e.g. that of LCNEC. Study III and IV include surgically treated early-stage lung cancers of AC histology and, in addition, the predominant type of AC (explained below) was considered in sub-analyses.

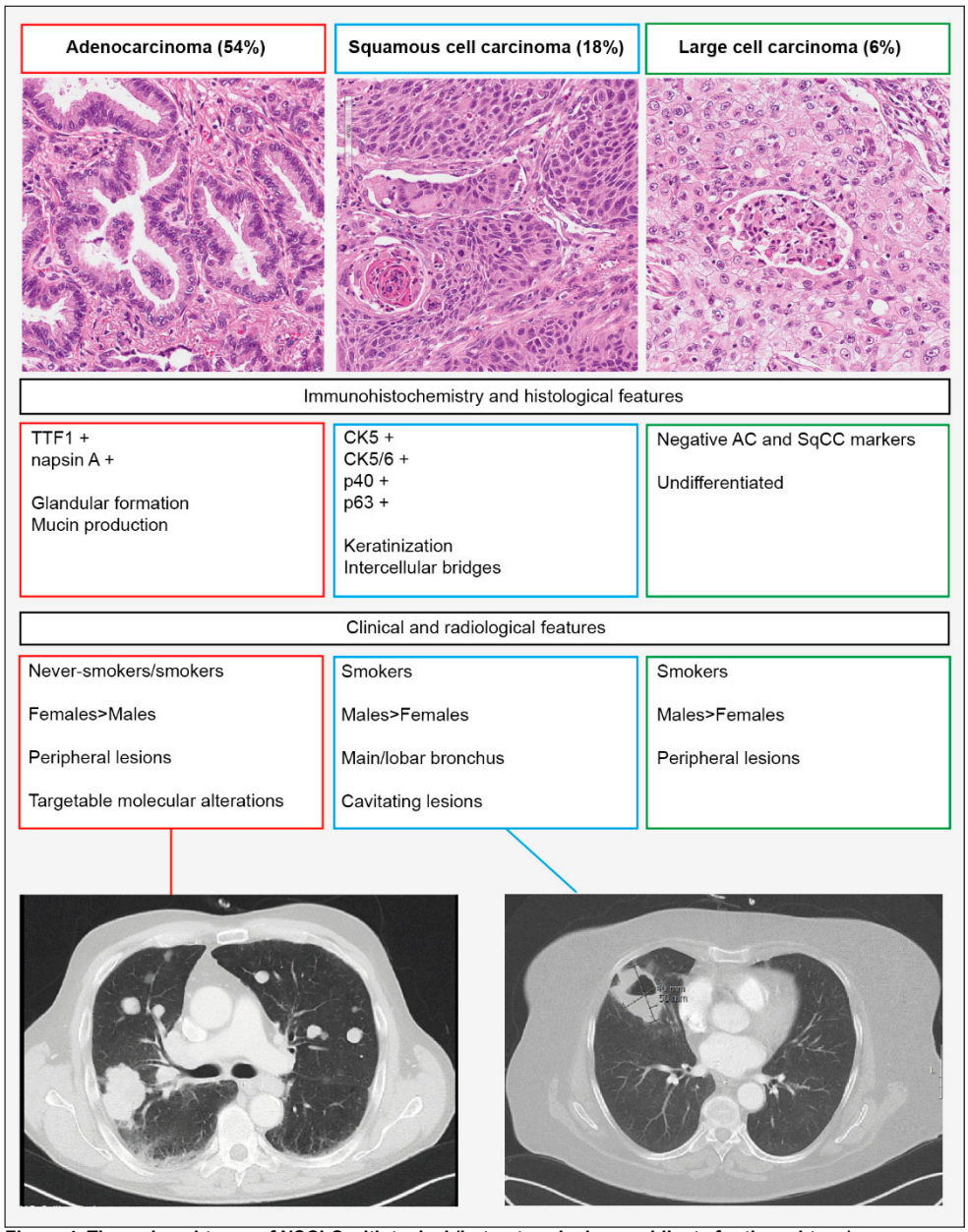


Figure 4. The main subtypes of NSCLC with typical (but not exclusive or obligate for the subtype) characteristics. Pictures of H&E stains and computed tomography (CT) scans.
Photos kindly provided by Dr. Hans Brunnström and Dr. Håkan Griph.

Adenocarcinoma

Decades ago SqCC was the most frequent lung cancer histology but this has shifted and AC is now the most commonly diagnosed lung cancer type[18]. In Sweden, AC accounted for more than half of the diagnosed lung cancers in the period 2015-2017[35].

Although smoking is associated with all lung cancer types, lung adenocarcinoma is the predominant histology among never-smokers with lung cancer. Driver gene alterations in for example *EGFR*, *KRAS*, *ALK*, *ROS1*, *BRAF*, *RET* and *ERBB2* are almost exclusively found in lung tumors of AC histology and most of them preferentially in patients with no or sparse smoking. Lung adenocarcinomas are generally located in the periphery of the lungs[18].

The tumors are characterized by glandular differentiation, mucin production and positive IHC of TTF1 and napsin A. Not all features have to be present for a diagnosis of adenocarcinoma. Likewise the IHC markers are not present in all lung AC and can also be positive staining in other histologies[18].

The main invasive AC subtypes are described in Table 1. Invasive AC usually comprise a mixture of subtypes and according to WHO classification, a predominant subtype should be noted, and present subtypes evaluated in 5% increments. In small biopsies/cytology, identifiable present patterns may also be described. Beside the invasive variants presented in Table 1, there are pre-invasive, minimally invasive AC types and other variants of invasive AC not presented here.

Table 1. Subtypes of invasive adenocarcinomas

Adenocarcinoma subtype	Features
Lepidic	• Tumor cells grow along the surface of alveolar walls • Invasive component in at least one focus of >5 mm in greatest dimension
Acinar	• Glands • Mucin may be present in neoplastic cells and glandular spaces
Papillary	• Growth of glandular cells along central fibrovascular cores
Micropapillary	• Small and cuboidal tumor cells • Psammoma bodies • Growth in papillary tufts, no fibrovascular cores
Solid	• Polygonal tumor cells • Tumor cells forming sheets

Squamous cell carcinoma

SqCC is the second most common lung cancer type, comprising 17.6% of lung cancer diagnoses in Sweden 2015-2017[35].

SqCC is more strongly associated with smoking than AC and observed changes in global incidence are closely related to altered smoking behavior. Tumors of SqCC histology usually arise in a main or lobar bronchus and the tumors tend to be locally aggressive and may invade local structures. Hypercalcemia, a paraneoplastic syndrome, might affect lung cancer patients, especially those with tumors of SqCC histology. In comparison to AC, treatment predictive mutations are rare. SqCC typically comprise a very high rate of mutations per megabase and gene copy numbers alterations are common. Frequently mutated genes include for example *TP53*, *PTEN* and *PIK3CA*[18].

The morphological features are more prominent in better differentiated tumors than poorly differentiated tumors, but characteristics of SqCC include keratinization, intercellular bridges and expression of a squamous cell carcinoma marker; CK5, CK5/6, p40 or p63[18].

Large cell carcinoma

LCC are undifferentiated NSCLC without histological, cytological or immunohistochemical resemblance of other histological types. Due to the more prominent role of IHC in the 2015 WHO classification compared to previous version, many tumors previously classified as LCC would now be classified as AC or SqCC. Since LCC is established much by ruling out other diagnoses, the diagnosis require resected tumor specimens[18]. In Sweden, 5.6% of the lung cancers 2015-2017 were classified as LCC[35].

Clinical characteristics of these tumors include male gender and most of the patients have a history of smoking. The tumors are typically peripherally located in the lungs[18].

Neuroendocrine lung tumors

Neuroendocrine tumors share morphological, immunohistochemical and molecular features. Nearly all patients diagnosed with SCLC are smokers, among the major lung cancer subtypes SCLC has the strongest association with smoking. Histological characteristics of SCLC include densely packed small cells with sparse cytoplasm and a high mitotic count. Tumors of SCLC histology can express positive staining of neuroendocrine markers[18]. SCLC accounted for 11.7% of lung cancer diagnoses

in Sweden 2015-2017[35]. LCNEC was previously classified as a variant of LCC. Similar to SCLC, these tumors are strongly associated with smoking. Both SCLC and LCNEC tend to be aggressive tumors. LCNEC is characterized histologically by rosettes and peripheral palisading and is by definition positive for at least one neuroendocrine marker[18].

Diagnostics

The diagnostic work-up aims to confirm the lung cancer diagnosis and to determine the disease stage, histology, and molecular pathology, which are all crucial for treatment decision in combination with an evaluation of the patient's comorbidities and general performance status.

Diagnostic work-up

Symptoms like cough, hemoptysis or pain might have led to the first radiological examination subsequently leading to diagnosis but when symptoms are present, the lung cancer has usually metastasized which is the situation in >50% of newly diagnosed lung cancers[35]. Typically, an early diagnosis might have been set due to radiology performed by other reasons leading to an accidental finding of a suspicious nodule in the lung. With x-ray of the lungs, lung cancer suspicious alterations can usually be detected, but further examination with contrast-enhanced computed tomography (CT) of thorax and upper abdomen, for metastasis screening in for example liver and adrenal glands, fluorodeoxyglucose-positron emission tomography (FDG-PET) with or without CT is required. PET is indicated foremost in patients with potentially curative disease to exclude dissemination and as a part of evaluation of thoracic lymph nodes which is also pivotal in a potential curative situation[39].

Flexible bronchoscopy is routinely used in lung cancer diagnostics, enabling tumor assessment from bronchial/transbronchial biopsies and bronchial brushing and plays a central role in both potentially curative and palliative situations[39-41]. When the tumor lesion is located peripherally, transthoracic percutaneous image (CT or ultrasound)-guided fine needle aspiration or core biopsy can be performed[39,40].

In potentially operable, early-stage tumors, bronchoscopy can be combined with transbronchial needle aspiration (TBNA) of lymph nodes by endobronchial ultrasound (EBUS). Lymph nodes that can be assessed by EBUS include position 2, 4, 7, 10 and 11 and sometimes position 1 (Figure 5). Position 5, 6, 8 and 9 cannot be assessed by EBUS[39]. The previous standard method for lymph node assessment,

mediastinoscopy, is more invasive and EBUS is the preferred method today. During optimal conditions there is a cytologist present during the bronchoscopy and EBUS to immediately process the obtained cytological material and evaluate if the samples are representative. Lymph node assessment should be performed pre-operatively in the situations of abnormal lymph nodes visualized with PET, enlarged lymph nodes (≥15 mm), centrally located tumor or a tumor with low uptake of FDG. When the tumor is small, PET-positive and located in the peripheral third of the lung and lymph nodes are <15 mm and without uptake on PET, lymph node examination is not considered obligate. PET has a high negative predictive value and high sensitivity but a positive predictive value around 50%. A positive uptake on PET can also be caused by infections while it can be necessary to evaluate lesions detected by PET with biopsy or cytology and furthermore, small lesions (<1cm) or slowly growing tumors, including some adenocarcinomas, does not always present with uptake on PET[39].

If metastatic lesions are suspected due to symptoms, directed diagnostic examinations can be performed for diagnostic purpose and to plan specific metastatic therapy, including for example bone scan, PET or magnetic resonance imaging (MRI). Metastatic sites might also be accessible for biopsy/cytology for verification of lung cancer diagnosis as an alternative to bronchoscopy or for tumor material needed for additional histopathology or molecular analyses. Pleural effusion is common, and thoracentesis can be used both for diagnostic and therapeutic purposes[39,40]. There is no overall consensus whether radiology of the brain should be performed in asymptomatic patients[41], but in Sweden the national guidelines recommend brain metastasis screening of locally advanced NSCLC when curative treatment is planned[39].

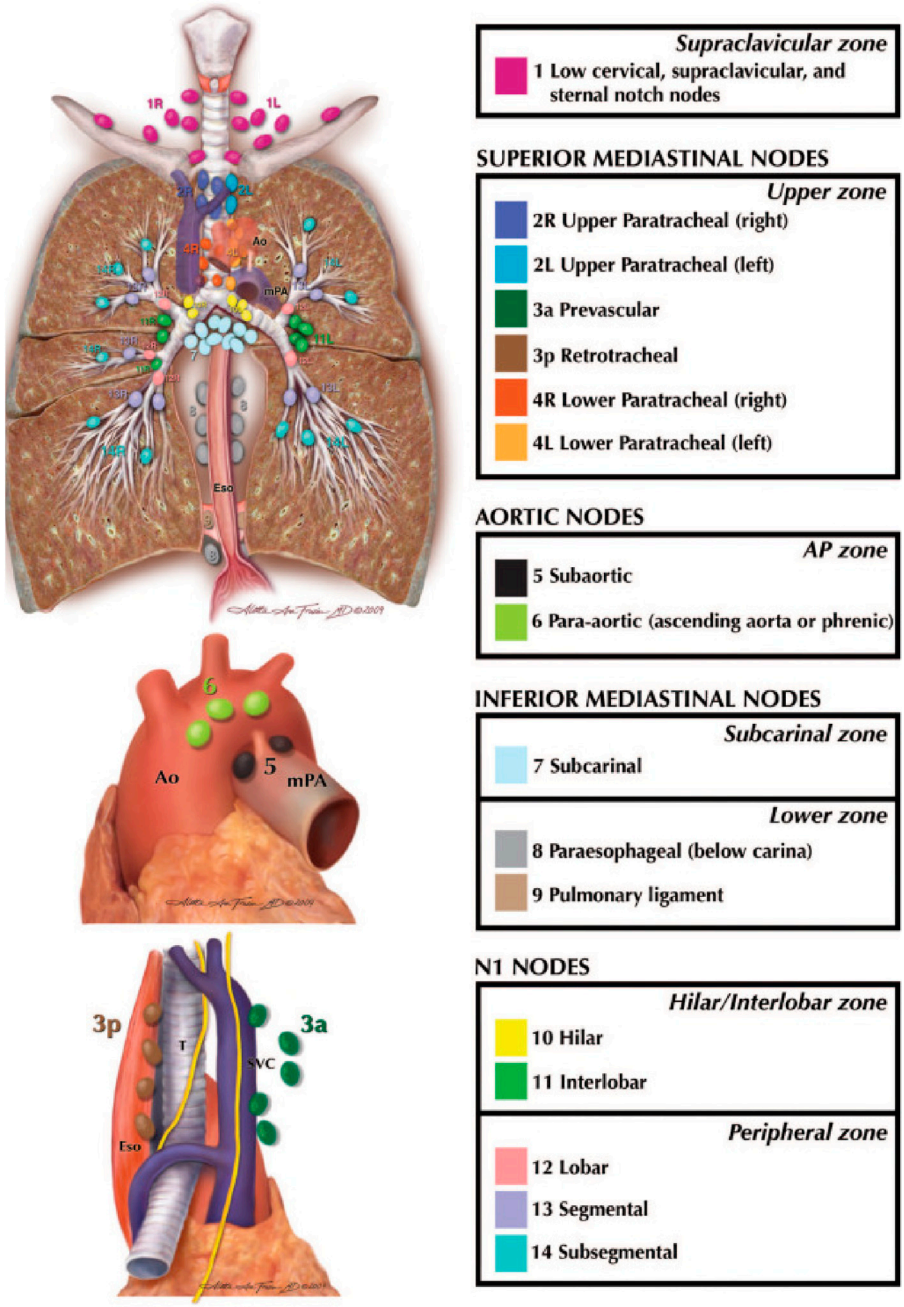


Figure 5. The International Association for the Study of Lung Cancer (IASLC) lymph node map with stations and zones.
Reprinted with permission from Elsevier[42].

Molecular diagnostics

The tumor specimens obtained should be sufficient both for diagnosis, either by morphology alone or with IHC, and for molecular pathology. The importance of molecular diagnostics is growing and constantly evolving. This is noticed also in various guidelines within the field, trying to keep pace with new facts. In their latest guideline for molecular testing, the College of American Pathologists (CAP), the International Association for the Study of Lung Cancer (IASLC) and the Association for Molecular Pathology (AMP) highlight that new findings (regarding e.g. *BRAF*-mutated lung cancers and use of immunotherapies) were published between literature review and guideline drafting[43]. In their guidelines, the major statements include that a minimum set of genes for test of advanced lung adenocarcinomas are *EGFR*, *ALK* and *ROS1*, followed by a recommendation of an additional set of genes (*BRAF*, *MET*, *RET*, *ERBB2* (*HER2*) and *KRAS*) to be included either in the first test in an expanded panel or in sequence after *EGFR*/*ALK*/*ROS1*. Although targeted therapy for *BRAF*-mutated lung cancer was FDA-approved at the time of publish of the guidelines, the authors stated that the published evidence for *BRAF* testing as a routine stand-alone assay were not sufficient at the time of review to warrant an international recommendation but acknowledge that the next guideline will include a stronger recommendation of *BRAF*-testing[43]. These guidelines were recognized by the American Society of Clinical Oncology (ASCO). The ASCO Expert panel endorsed minor modifications, including recommending *BRAF* test together with *EGFR/ALK/ROS1*[44]. In the guidelines by ASCO and CAP/IASLC/AMP, molecular testing beyond adenocarcinoma is discussed and recommended for e.g. selected patients with certain clinical features[43,44]. In the end, guidelines are intended to ease but not being final for clinical decision. Worldwide, lack of resources often means difficulties in implementing these guidelines.

In Region Skåne, molecular testing of lung cancer by next generation sequencing (NGS, or massive parallel sequencing, MPS) was set up in a clinical framework in January 2015[45]. The clinical aspects of the first 1.5 year of molecular profiling of lung cancer by NGS in Region Skåne is described in study II. The first NGS-panel included targeted sequencing of specific exons in 26 genes including *EGFR*, *KRAS*, *BRAF*, *ERBB2* and *MET* but with major clinical focus on *EGFR* and *KRAS*. *ALK* translocations were at the time analyzed with IHC and/or FISH and *ROS1* rearrangements were only analyzed in a minor fraction of the patients, in these cases with FISH and/or IHC.

Today (March 2019), a larger set of genes and fusion genes are routinely analyzed by NGS in Region Skåne. Mutation analysis and fusion gene detection is performed with NGS on DNA and RNA, respectively, reflecting a change towards more extensive molecular analysis and simultaneous detection of mutations and fusion genes. Moreover, blood-based analysis of selected mutations in singleplexed assays are about to be implemented in the clinic (May 2019).

Screening

Due to generally poor prognosis related to diagnosis in a late disease stage and the possibility to cure the disease if diagnosed early, screening to detect lung cancer before dissemination could be very beneficial. Implementing screening requires defined criteria for the population subjected to screening and a program for follow-up. Based on large studies there are evidence of benefit of lung cancer screening and an implementation in Europe is probably on the edge. In The National Lung Cancer Screening Trial (NLST) screening by low-dose computed tomography (LDCT) compared to chest x-ray was compared in a cohort comprised of more than 53000 American current or former smokers. Results in NLST showed a 20% relative reduction in lung cancer mortality rate by LDCT screening compared to screening by chest x-ray[46].

Recently, a large Dutch-Belgian lung cancer screening trial, NELSON, revealed the results from a 10-year follow-up of current and former smokers randomized into screening with LDCT or to no screening. After 10 years, the lung cancer mortality reduction was 26% among the males in the screening arm compared to the control arm whereas an even larger (39%, however not reaching statistical significance) benefit was shown in the small subset of females in the screening arm compared to the control arm[47].

Staging

The patients in this thesis work are staged according to the 7th edition of tumor, node and metastasis (TNM) classification published in 2009. The component T refers to the size, localization, and certain other characteristics of the primary tumor, the N component describes the lymph node involvement and M refers to metastasis. Combinations of T, N and M are grouped into stages I-IV in which I-III are subdivided into A and B [48]. The specific characteristics for T, N and M within the 7th edition of the classification, are described in Table 2 and stage in Table 3.

Table 2. T, N and M descriptors in the 7th edition of TNM classification of lung cancer[48].

T/N/M Description in the 7th edition of lung cancer classification	
T: Primary tumor	**Tumor size and description**
TX	Primary tumor cannot be assessed, or tumor proven by the presence of malignant cells in sputum or bronchial washings but not visualized by radiology or bronchoscopy
T0	No evidence of primary tumor
Tis	Carcinoma in situ
T1	Tumor ≤ 3 cm in greatest dimension, surrounded by lung or visceral pleura, no bronchoscopic evidence of invasion more proximal than the lobar bronchus
T1a	Tumor ≤ 2 cm in greatest dimension
T1b	Tumor > 2 cm but ≤ 3 cm in the greatest dimension
T2	Tumor > 3 cm but ≤ 7 cm or any of the following features: • Invades visceral pleura • Involves main bronchus ≥ 2 cm distal to the carina • Associated with atelectasis/obstructive pneumonitis extending to the hilar region but not involving the entire lung
T2a	Tumor > 3 cm but ≤ 5 cm in the greatest dimension
T2b	Tumor > 5 cm but ≤ 7 cm in the greatest dimension
T3	Tumor > 7 cm or any of the following: • Direct invasion of the chest wall, diaphragm, phrenic nerve, mediastinal pleura or parietal pericardium • Tumor in the main bronchus < 2 cm distal to the carina without involvement of the carina • Associated atelectasis/obstructive pneumonitis of the entire lung or separate tumor nodule(s) in the same lobe as the primary tumor
T4	Tumor of any size that invades any of the following: mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, vertebral body, carina or separate tumor nodule(s) in an ipsilateral lobe to that of the primary tumor
N: Regional lymph node involvement	**Description**
Nx	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes, including involvement by direct extension
N2	Metastasis in ipsilateral mediastinal and/or subcarinal lymph nodes
N3	Metastasis in contralateral hilar, contralateral mediastinal, ipsilateral or contralateral scalene or supraclavicular lymph node(s)
M: Distant metastasis	**Description**
M0	No distant metastasis
M1	Distant metastasis
M1a	Separate tumor nodule(s) in a contralateral lobe; tumor with pleural nodules or malignant pleural or pericardial effusion
M1b	Distant metastasis in extrathoracic organ(s)

Table 3. Stage according to 7th TNM classification for lung cancer[48]

T/M	N0	N1	N2	N3
T1a	IA	IIA	IIIA	IIIB
T1b	IA	IIA	IIIA	IIIB
T2a	IB	IIA	IIIA	IIIB
T2b	IIA	IIB	IIIA	IIIB
T3	IIB	IIIA	IIIA	IIIB
T4	IIIA	IIIA	IIIB	IIIB
M1a/b (Any T)	IV	IV	IV	IV

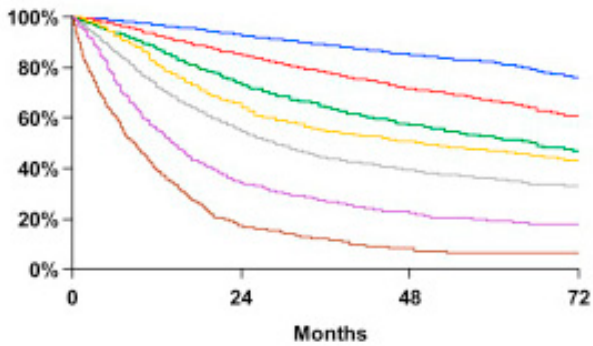
The patients in study I were all diagnosed before 2009 which also accounts for some of the patients in study III and IV. Stage in study I is based on clinical and pathological information available at the time. The tumor stage in study III and IV have been reclassified to TNM 7th edition by a thoracic pathologist. Patients in study II were diagnosed 2015-2016 and the tumors were staged according to the TNM 7th edition. The current classification system, the 8th edition of lung cancer classification, became worldwide standard in January 2017[50]. In Sweden it was implemented in the beginning of 2018.

Changes from the 7th edition in the 8th edition include further subdivision of the T component and M component, whereas the N component remained unchanged. The T category was subdivided in 1-cm increments up to 5 cm and T1 was extended from T1a and b to T1a-c (T1c including tumors of size >2 cm-≤3cm). Tumors between >5 cm - ≤ 7 cm are categorized as T3 instead of T2 as in the 7th edition and T4, depending on descriptors other than size in the 7th edition, include tumors > 7 cm. Furthermore, qualitative descriptors of T are changed; central tumors involving a main bronchus are classified as T2a regardless of distance to the carina, tumors causing obstructive atelectasis are classified as T2 regardless of partial or complete atelectasis of the lung, tumors involving the diaphragm are classified as T4 and the former T3 descriptor mediastinal pleural involvement has been omitted. The M component has been subdivided into M1a-c in which M1a is unchanged from the 7th edition whereas M1b represents tumors with a single distant (extrathoracic) metastasis and M1c represents tumors with multiple metastases in a single organ or multiple metastases in several organs[49,50].

Prognostic factors

Prognostic factors provide information on the natural course of the disease unrelated to treatments, whereas a predictive factor predicts response of a certain treatment. In some cases, a factor can be both prognostic and predictive.

Lung cancer survival is poor, with a five-year survival rate of only around 20% and an even worse prognosis for SCLC[51]. There is a clear variation through stage according to TNM (Figure 6), which is a well-established prognostic factor in lung cancer. Also, the degree of tumor burden within stage IV has a prognostic impact and in the 8th edition of TNM, stage IVA includes M1a and M1b (i.e. intrathoracic or single distant metastasis, respectively) while stage IVB includes M1c (i.e. multiple metastases) and has a worse prognosis than stage IVA[49].



7th Ed.	Events / N	MST	24 Month	60 Month
IA	1119 / 6303	NR	93%	82%
IB	768 / 2492	NR	85%	66%
IIA	424 / 1008	66.0	74%	52%
IIB	382 / 824	49.0	64%	47%
IIIA	2139 / 3344	29.0	55%	36%
IIIB	2101 / 2624	14.1	34%	19%
IV	664 / 882	8.8	17%	6%

Figure 6. Overall survival by clinical stage according to the 7th edition of the TNM classification for lung cancer.
Reprinted with permission from Elsevier[49].

Performance status (PS), a subjective grading of the patient's level of functioning in daily life, has been widely studied as a prognostic factor and there is now much evidence of PS being linked to outcome, i.e. a better PS predicts a more favorable outcome[52]. For PS in the studies within this thesis, we refer to a frequently used grading of PS from the Eastern Cooperative Oncology Group (ECOG) [53] (Table 4). Another patient-related prognostic factor, besides PS, is weight loss, where less weight loss at baseline is associated with better outcome[52].

Table 4.Performance status (ECOG)

Grade	Explanation of activity
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out any work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all selfcare but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited selfcare, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any selfcare. Totally confined to bed or chair
5	Dead

Women tend to have better prognosis than men, also when adjusting for stage, histology or smoking respectively[52,54]. Smoking is a well-known, established risk factor for lung cancer and several studies have also concluded that smoking is a negative prognostic factor[52,55,56]. Suggested explanations for this negative prognostic effect include not only smoking-related comorbidities but also an interference with the immune system leading to reduced capacity to clear cancer cells [55].

Adenocarcinomas are associated with better prognosis than other histological types of NSCLCs[57]. However, the classification of adenocarcinoma published in 2011[36] links subtypes of AC to outcome independent of stage, with predominant lepidic tumors associated with better outcome and micropapillary tumors being one of the subtypes associated with worse outcome[58,59]. Further suggested stratification of adenocarcinomas for prognosis include different epigenetic subgroups based on methylation pattern[60] and a subgroup of adenocarcinomas with shared gene expression pattern showing less aggressiveness[61].

In clinical settings, prognostic factors would aid in stratifying lung tumors into risk groups, in order to choose the best-suited treatment or monitoring of the patient. Adjuvant treatment is one example, described in the treatment section, where the benefit is significant but the absolute effect on survival is moderate. Although stage is a very important issue in the adjuvant treatment decision-making, it is insufficient, and additional risk stratification tools are needed. Another example is in the

palliative setting, where additional prognostic factors could be helpful when deciding the follow-up intensity or the choice of more or less aggressive treatments. Furthermore, studies of prognostic factors might be a step towards further stratification of lung tumors, generating hypotheses for functional studies of these factors and for subsequent treatment-related studies.

Many projects investigate prognostic factors, but few have been implemented. With the above-mentioned examples of important, strong and in some cases suggested prognostic factors, the value of investigations of a single potential prognostic marker might be of limited value. Identified single prognostic markers should be further investigated in relation to other variables, considering the complexity of lung cancer.

Treatment

Treatment opportunities for NSCLC have been evolving rapidly in recent years much due to an increasing number of targeted therapy drugs and immunotherapy as a new cornerstone in the treatment arsenal. Further treatment options, both new drugs and new combination of drugs, are in pipeline and the treatment overview for newly diagnosed NSCLC, illustrated in Figure 7, will probably be changing in a near future. Additionally, recommended first hand choice in curative situations or first line treatments illustrated in Figure 7, and described in the following sections, cannot be applied for all patients due to comorbidities, tolerability issues and patients' preferences. Furthermore, despite efforts, staging, molecular pathology, and accurate diagnosis can be challenging. In order to discuss treatment strategies, multidisciplinary conferences with radiologists, pathologists, pulmonologists and oncologists are generally performed at all Swedish hospitals responsible for management of lung cancer patients.

A central part in choice of treatment concerns tumor-related factors, most importantly disease stage and mutational or fusion gene profile. The staging helps to decide if curative intention is possible, whereas the mutational and fusion gene analysis is to determine whether targeted therapy can be given in a palliative situation. Unfortunately, the majority of lung cancer patients are diagnosed in a late stage (approximately 50% with metastatic disease and 70% with either locally advanced or metastatic disease)[51], when treatment generally has a palliative intent, although, in selected cases with single solid metastases curative intention is sometimes possible. A description of lung cancer patients treated in a single Swedish health care region is presented in study IV.

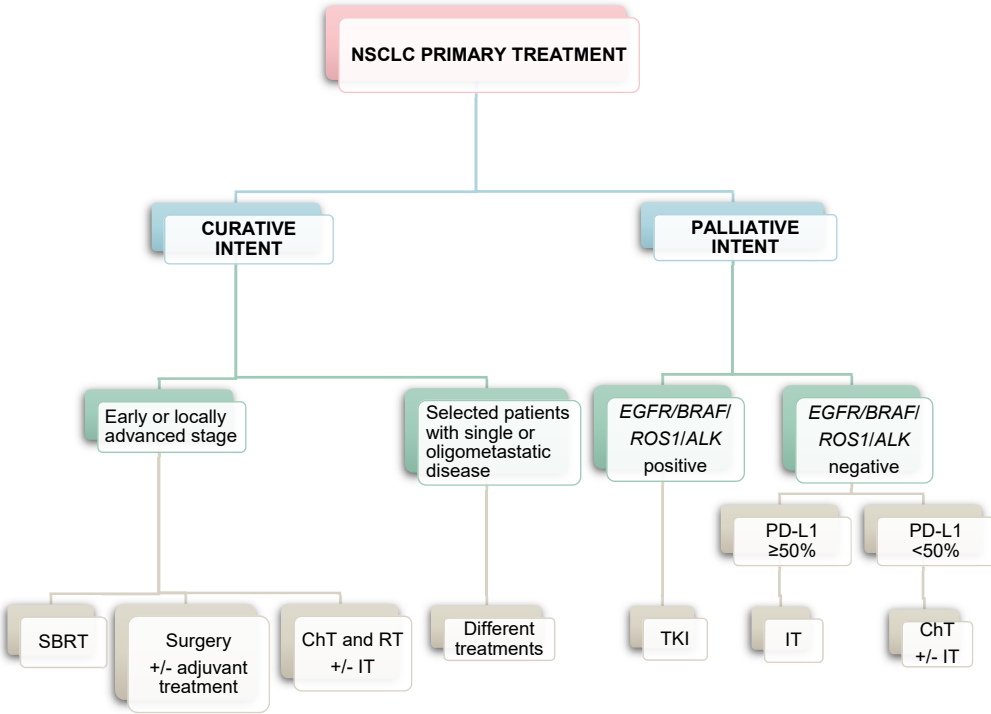


Figure 7. Treatment overview in newly diagnosed NSCLC. In a majority of newly diagnosed lung cancers, curative treatment is not possible. In these situations, mutational profile and result from PD-L1 test guide treatment choice. Only first line treatment options are illustrated and size of patient groups in each treatment alternative is not taken into account. Abbreviations: ChT=chemotherapy, IT=Immunotherapy, TKI=tyrosine kinase inhibitor, RT=radiotherapy

Surgery

Surgically treated patients with early stage tumors comprise all included patients in study I, III and IV, and part of the patients in study II, in this thesis work. Surgery is the prominent option of curative treatment and should be considered in early stages and for selected patients with locally advanced stage. Pre-surgical examination should include a proper staging and lymph node assessment and a general risk evaluation considering co-morbidities and specifically the patient's cardiopulmonary function and predicted post-surgical status. For patients with single station N2 disease, surgery is not excluded, but in case of multiple N2 disease or N3, surgery is not recommended[41].

Lobectomies/bilobectomies, in which one or two lobes, respectively, are removed are the most common surgical procedures in lung cancer[39]. In study II, III and IV the proportion of lobectomies/bilobectomies varied between 76-86%. An alternative

to lobectomy is sublobar resections (segmentectomies that follow anatomy and the non-anatomical wedge resections), which can be the preferred choice when patient's cardiovascular function is not sufficient for a lobectomy. It could possibly also be an alternative if the tumor is small and without lymph node involvement, but lobectomy is still the recommended procedure. In the mid-nineties, a randomized controlled trial (RCT) was published with the findings of a higher rate of local recurrence after a sublobar resection compared to lobectomy for stage I NSCLC[62]. So far, no newer RCT have been published although staging and surgery have developed. Thus, results from older studies might not be entirely applicable today and patients who can be treated with sublobar resection without increased risk of recurrence in comparison with a lobectomy still needs to be defined. Although some studies have supported the conclusion of lobectomy being superior to sublobar resections in stage I other studies have shown that lobectomies and sublobar resections are equal and that several variables such as histology, tumor size, age and choice of sublobar resection are important to consider [63-65]. Results from new randomized trials are awaited to shed new light on the question of lobar versus sublobar resections[66-68].

The first successful pneumonectomy/pulmectomy of a lung cancer patient, a procedure in which an entire lung is removed, was performed in 1933 by the surgeon Dr Evarts A. Graham who later died of lung cancer himself, survived by the patient and friend who he had operated 25 years earlier[69]. Pulmectomies (approximately 20% of the lung cancer surgical procedures) are necessary in cases of centrally located tumors or tumors with engagement of all ipsilateral lobes. This procedure is associated with a higher risk of complications[39].

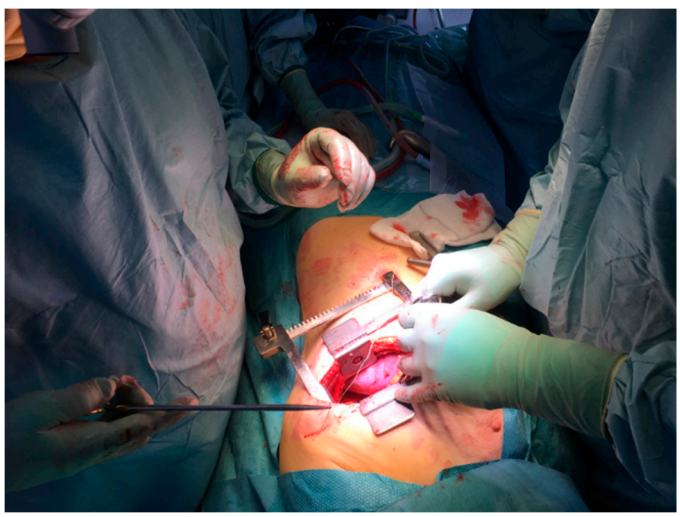


Figure 8. Thoracotomy (Department of Cardiothoracic surgery, Skåne University Hospital, Lund)
Photo kindly provided by Dr. Sandra Lindstedt Ingemansson.

Lung cancer surgery is performed either with open thoracotomy (Figure 8) or by video-assisted thoracoscopic surgery. VATS is associated with improved post-operative recovery than thoracotomy[70,71]. The next development to have a breakthrough in lung cancer surgery might be robotic-assisted lobectomy, i.e. computer-aided surgery. There are no RCTs or large multicenter retrospective studies, but existing studies have suggested equal outcome or possibly even better outcome of robotic surgery compared to other surgical methods[72,73].

Radiotherapy

Radiotherapy (Figure 9) has for long been, and still is, a central therapy in lung cancer treatment both in curative and palliative situations. Since adverse effects from radiotherapy are dose-dependent, the dose-planning depends on the treatment intent. Some of the important side effects include inflammatory pneumonitis, lung fibrosis, esophagitis and toxicity of the heart[39].

Figure 9. Linear accelerator at the Oncology department, Skåne University Hospital, Lund.
Photo kindly provided by Dr. Jens Engleson.

Curative intent

Radiotherapy can be given with curative intent as the only oncological treatment for patients with small tumors or in combination with chemotherapy for patients with locally advanced tumors. Furthermore, radiotherapy plays a role in the neoadjuvant setting and in post-operative cases with incomplete resections.

Medically inoperable patients might be offered curatively intended stereotactic body radiotherapy (SBRT), also called stereotactic ablative radiotherapy (SABR). Small, preferably peripheral, tumors are best-suited for this treatment. In stereotactic radiotherapy the radiation beams are delivered with very high precision and the treatment is usually well-tolerated. In Sweden, SBRT is usually given in three fractions of 15 Gy[39]. RCTs investigating the outcome of SBRT versus surgery are ongoing, where patient selection and type of surgery for comparison are essential. Some former RCTs attempting to compare effectiveness between these modalities have closed early, facing difficulties to recruit patients[39,41,74-76].

In locally advanced NSCLC, radiotherapy might be given in combination with chemotherapy (CRT) as a curatively intended therapy which can be sequential, starting with chemotherapy and subsequent radiotherapy, or concomitant. Concomitant treatment has been shown to be superior to sequential treatment in terms of survival[77]. A platinum-based combination is the recommended choice of chemotherapy regimen. There is a dose-response relation in effect of radiotherapy, but higher doses generally come with increased risks of toxicity. Suggested optimal doses vary between 60 and 68 Gy[39,41].

Other situations in which radiotherapy is given in curative situations are pre- and post-surgery. Trimodality therapy, i.e. the combination of CRT followed by surgery, is applied for selected locally advanced cases[40]. Post-operative radiotherapy should be considered if surgery has not been radical[39,41].

Palliative intent

Radiotherapy in the palliative setting may serve both as local tumor control and symptom relief. If the patient has a good performance status up to 36 Gy against the primary lung tumor can be given, otherwise approximately 20 Gy is an option. An important role of radiotherapy is also local control and symptomatic treatment of metastases such as to the bone or CNS[39,40].

Chemotherapy

Chemotherapy is used in curatively intended treatments, either in combination with radiotherapy in cases of unresectable NSCLC as described above, or as adjuvant therapy after surgery. In metastatic disease, chemotherapy might be given alone or in combination with immunotherapy in first-line. Adverse effects of chemotherapy partly depends of regimen but generally include neutropenia, anemia, nausea, gastro-intestinal side effects and fatigue.[39]

Curative intent

Adjuvant chemotherapy after radical surgery has been shown to be beneficial in tumors with lymph node engagement or tumors larger than 4 cm, i.e. patients with stage IA tumors are not considered for adjuvant chemotherapy. The benefit of adjuvant chemotherapy is an absolute increase in survival of 4-5% at five years post-surgery, thus, a modest but significant increase in survival, based on evidence from several randomized trials and meta-analyses. Adjuvant chemotherapy is usually given in four cycles and the regimen recommended is a cisplatinum-based combination, preferably vinorelbine for which most robust evidence exist[78,79].

Palliative intent

Chemotherapy was proven to be more beneficial than best supportive care in patients with metastatic lung cancer in the 1990's and has since then been a cornerstone in the treatment of metastatic NSCLC[80]. The proportion of patients treated with chemotherapy alone in first line is expected to decrease following the pending implementation of a recommended combination of chemotherapy and immunotherapy. Standard chemotherapy regimen in first line is four cycles of a platinum-based combination with a taxane, gemcitabine, vinorelbine or pemetrexed to patients with PS 0-2. Monotherapy is an option for patients who do not tolerate doublet chemotherapy[39,40]. Histology guides the choice of chemotherapy combination, where platinum combined with pemetrexed have been shown to be effective particularly in patients with non-squamous histology[81,82]. In patients with a good performance status, tumor with non-squamous histology and no progression after four cycles of platinum-based chemotherapy, maintenance therapy, that is a continuous treatment, can be considered and for which pemetrexed has the indication[39,40].

Since 2015, immunotherapy has an increasing role in patients with negative predictive tests for targeted therapies. So far, immunotherapy in first-line treatment has been restricted to monotherapy and in patients selected through positivity of an IHC-evaluated marker, programmed cell ligand 1 (PD-L1), whereas the remaining patients were treated with chemotherapy. However, during 2018, studies demonstrating immunotherapy and chemotherapy combinations being superior to

chemotherapy alone in first-line treatment of metastatic NSCLC were published[83-85]. The studies concluded effect also regardless of PD-L1 expression and first line treatment of metastatic NSCLC is therefore now about to include immunotherapy for all patients without contraindications. Thus, the treatment of patients with negative predictive tests for TKI will consist of two main directions as illustrated in Figure 7, chemotherapy and immunotherapy or immunotherapy as monotherapy.

Immunotherapy

Immunotherapy is the latest added modality in the treatment arsenal of lung cancer. As seen in Figure 7, immunotherapy now has a role in both curative and palliative treatments of newly diagnosed NSCLC. However, hitherto, the majority of patients receiving immunotherapy has been in later (second or beyond) lines of palliative treatment.

The immune system recognizes differences between normal and malignant cells but gained abilities in the tumor or modulations of the microenvironment can disturb this normal function, making the tumor escape the immune system. Tumors might avoid elimination by the immune system by loss of antigenicity, loss of immunogenicity or establishing an immunosuppressive environment. Examples of these mechanisms in tumors include defects in antigen presentation or upregulation of the immunoinhibitory membranous molecule PD-L1. By acting on activating signals in the immune system or by inactivating inhibitory signals, immunotherapy alters the interactions between the immune system and the tumor steering it towards tumor regression[86].

The immunotherapies used in lung cancer are so called immune checkpoint inhibitors that target either the ligand PD-L1 (atezolizumab, durvalumab) or its receptor programmed death 1 (PD-1; nivolumab, pembrolizumab). Immunotherapy is overall well-tolerated but adverse effects such as immune related adverse events (irAEs), e.g. pneumonitis, thyroiditis, and colitis, occur. Many irAEs can be managed with corticosteroids[83-85,87].

Membranous IHC of PD-L1 expression on tumor cells is a marker used for selecting patients to immunotherapy in first line. However, patients without PD-L1 expression can benefit from immunotherapy and vice versa. Another potential marker for immunotherapy is tumor mutational burden (TMB), the number of nonsynonymous mutations in tumor DNA. A suggested definition of high TMB is $\geq$ 10 mutations/megabase of genome examined.

Other potential predictive markers for response to immunotherapy under investigation include e.g. quantity of tumor infiltrating lymphocytes and their PD-L1 expression, gene-signature expression, and the intestinal microbiome[88,89]. Mismatch repair defects have been associated with immune checkpoint inhibitor efficacy in several other tumor types but seem to be rare in lung cancer[90,91].

Curative intent

A recently published RCT for unresectable stage III NSCLC demonstrated a survival advantage of durvalumab versus placebo when administered after standard concurrent CRT to patients without progression during the CRT. It is suggested that chemoradiation up-regulates PD-L1 expression in tumor cells and, by administering a PD-L1 inhibitor, this suppression of the immune system would be reversed[87].

Palliative intent

The start of immunotherapy in NSCLC treatment was nivolumab in second line, requiring positivity in PD-L1 for non-squamous histology but not for SqCC where nivolumab had shown to be superior to docetaxel in second line in RCTs[92,93]. A first line indication of immunotherapy was reached in February 2017 with pembrolizumab[39] and included patients with advanced NSCLC and PD-L1 expression on ≥ 50% of tumor cells[94].

More recently, studies concluded that addition of pembrolizumab to platinum-based chemotherapy results in longer overall survival (OS) and improved progression-free survival (PFS) than platinum-based chemotherapy alone regardless of PD-L1 positivity or not. However, if patients with at least 50% positive PD-L1 expression on tumor cells benefit from immunotherapy combined with chemotherapy or if immunotherapy as monotherapy is equally beneficial remains to be investigated[84,85]. Furthermore, the addition of atezolizumab to the angiogenesis inhibitor bevazicumab plus chemotherapy has shown to significantly improve OS and PFS[83] and a combination of nivolumab and ipilimumab (an antibody that is directed against cytotoxic T-lymphocyte-associated antigen 4, CTLA-4, and thereby obstructs T-cell-inhibiting signals), have shown to be effective in patients with high TMB[95].

In summary, in first-line treatment of advanced NSCLC, immunotherapy as monotherapy is considered for patients without targetable mutations/gene fusions and with a PD-L1 expression ≥ 50% and for patients with <50% PD-L1 expression a swift from chemotherapy to chemotherapy plus immunotherapy is being implemented in Sweden during 2019.

Targeted therapy

The identification of treatment predictive mutations in *EGFR* in 2004 marks the beginning of an era of rapidly evolving personalized medicine in lung cancer. Today, multiple targeted therapies are available for NSCLC, offering patients treatments that are generally more tolerable and more effective than chemotherapy. However, these treatment-predictive genetic alterations are only present in a subset of patients and there is a tough battle against resistance mechanisms. In addition to EGFR inhibitors, approved targeted therapies for NSCLC are currently available for NSCLC harboring *ALK* fusions, *ROS1* fusions, or *BRAF* mutations. In addition, targetable genes where approval currently awaits results from clinical trials include e.g. *RET*, *MET*, and *NTRK*.

EGFR

In 2004, a correlation between activating mutations in the kinase domain of *EGFR* and response to the reversible EGFR TKI gefitinib was discovered. It had been noticed that response to EGFR TKI was associated with specific features, such as female gender, being a never- or light-smoker, East Asian origin, and adenocarcinoma histology[96,97]. At the time of this milestone discovery, EGFR TKI (gefitinib or erlotinib) was indicated as monotherapy for chemotherapy refractory patients with NSCLC[98,99].

The *EGFR* exons 18-21 all code parts of the tyrosine kinase domain of the receptor but, with a few exceptions, all TKI-sensitizing mutations are found in exon 18, 19, or 21, whereas mutations in exon 20 predict resistance. The majority (approximately 85%) of the sensitizing mutations is comprised of exon 19 in-frame deletions and a substitution mutation in exon 21, L858R. The prevalence of *EGFR* mutations in NSCLC vary with populations studied, approximately 10-15% in Europe and about 40% in Asia[25,26,100,101].

The discovery of *EGFR* mutations being predictive of EGFR TKI-response led to a large number of subsequent RCTs comparing first-generation (erlotinib or gefitinib, both with reversible blocking of EGFR signaling) or second-generation (afatinib, an irreversible blocker of EGFR signaling) EGFR TKI to chemotherapy in first-line for patients harboring *EGFR* mutations concluded a better response and PFS with EGFR TKI in this subset of patients which led to the first FDA approval of EGFR TKI as a first-line options in patients with EGFR mutations 2013 (erlotinib). EGFR TKIs are generally well-tolerated, side effects include for example rash and gastrointestinal disorders[102-105].

By unambiguously increasing both the survival time and the quality of life for patients with disseminated disease, the introduction of EGFR inhibitors marked a major turning point in the treatment of NSCLC. However, nearly all patients at some

point progress due to acquired resistance in the tumor. A particular mutation in exon 20, T790M, is responsible for about 50-60% of the resistant cases. Some other resistance mechanisms include *MET* or *HER2* amplification leading to a by-pass of the inhibited EGFR by upregulation of alternative signaling pathways, downstream activation through mutations in *PIK3CA,* epigenetic alterations, and histological transformation to SCLC. Different resistance mechanisms might be present at the same time[106,107].

Osimertinib is a third-generation EGFR TKI that targets both T790M and sensitizing *EGFR* mutations. Accordingly, studies have revealed effect both in the situation of resistance developed during treatment with first- or second-generation EGFR TKI and in TKI-naïve patients. Improved PFS with osimertinib compared to first-generation EGFR TKI in first line[40,108] has made osimertinib a first-line treatment option for patients with *EGFR*-mutated tumors[40]. More effective first-line treatments push the second-line treatment considerations further ahead but the problems with acquired resistance remains. Mechanisms of resistance to osimertinib as a first-line treatment is not yet fully investigated. However, some identified resistance mechanisms to osimertinib in the situation of pre-treated T790M-positive NSCLC are similar to those for first- and second-generation EGFR TKI, i.e. activation of alternative pathways and histologic transformation to SCLC but also loss of T790M-mutant clones and *EGFR* mutation C797S[109,110].

ALK

In 2007, the first reports of anaplastic lymphoma kinase (*ALK*) fusions in lung cancer were published[111,112]. *ALK* rearrangements are present in 3-7% of NSCLC, with varying frequencies in different geographic areas, and, similar to EGFR-positive lung cancer, these tumors are associated with younger age and adenocarcinoma histology and are more frequently found in patients who have never smoked or who smoked sparsely[113,114].

The first ALK TKI, crizotinib, was approved in Sweden 2014 for ALK-positive lung cancer patients previously treated with chemotherapy[115] and was subsequently demonstrated to achieve significantly better PFS in first-line compared to chemotherapy in ALK-positive NSCLC[116]. Since then, additional ALK inhibitors have been introduced and approved for use in first line; Ceritinib and alectinib are ALK inhibitors which have shown better intracranial activity than crizotinib[117 118] and alectinib is currently the preferred first line option in most Swedish lung cancer units. For second-line treatment or beyond, yet another ALK inhibitor, brigatinib, was recently approved based on a trial reporting promising PFS in patients previously treated with crizotinib, and is currently also compared to crizotinib in ALK-positive patients with no prior ALK inhibition, revealing a superior PFS in the first interim analysis[119].

Similar to anti-EGFR therapies, the ALK inhibitors lose their antitumoral effect when the tumor sooner or later develop resistance. Resistance mechanisms to ALK inhibition are complex and include e.g. several mutations within *ALK* itself, with differences partly varying depending on the ALK inhibitor used[120]. Also the adverse effects from ALK inhibition vary between the different TKIs but may include e.g. nausea, gastrointestinal side effects and elevated aminotransferases[40].

ROS1

ROS1 fusions, present in approximately 1% of NSCLC, define yet another small subset of NSCLC with an actionable target, associated with little or no smoking and adenocarcinoma histology. Crizotinib is currently the first-line choice for patients with ROS1-rearrangements[39,40,121].

BRAF

Dual BRAF- and MEK-inhibition by dabrafenib and trametinib is approved for treatment of *BRAF* V600-mutated lung cancer based on two non-randomized multicenter trials demonstrating efficacy in advanced NSCLC treated as first-line and second-line therapy, respectively. The targetable V600E mutation in *BRAF* is detected in approximately 1-2% of NSCLC[28,40,122].

Other

Other oncogenic drivers in NSCLC that are actionable with targeted therapies but not yet established in clinical routine include for example *RET* fusions, alterations in *MET*, and *NTRK* fusions.

RET fusions are discovered in 1-2% of NSCLC. There are no RET-selective inhibitors approved yet, and multikinase inhibitors have hitherto unfortunately not reached the same results as seen for other TKIs[40,123].

Among *MET* alterations, splice site mutations have yielded the largest interest in lung cancer. These mutations are found in 3-4% of NSCLC, usually in adenocarcinoma or sarcomatoid carcinoma, and lead to aberrant splicing and *MET* exon 14 skipping. Crizotinib has shown effect as a multi-kinase-inhibitor in this patient group and, furthermore, MET-directed TKIs are under development[31,40].

NTRK fusions are very rare (<1%) in NSCLC, but studies of targeted therapies, e.g. larotrectinib, have shown encouraging results[40,124].

The monoclonal antibody bevacizumab is directed against vascular endothelial growth factor (VEGF) to inhibit tumor angiogenesis, and can be added to chemotherapy (most evidence in carboplatin/paclitaxel) in non-squamous advanced NSCLC or to erlotinib in *EGFR*-mutated NSCLC[39,40,125,126].

Furthermore, the TKI nintendanib, targeting the VEGF receptor (VEGFR), can be used in combination with docetaxel as a second-line treatment for adenocarcinoma[127]. Currently there are no predictive markers for response to angiogenesis inhibitors[128].

Treatment options for lung cancer recurrence and progression

Despite being the best chance of cure from what we know today, only about 30 to 70% of surgically treated patients (i.e. stage I to III) are free from lung cancer recurrence within five years from diagnosis[129]. Only selected cases of strictly local recurrences are available for salvage with radiotherapy or, more seldom, surgery. Beside comorbidities and patient performance status, the treatment options in the situation of lung cancer recurrence therefore most often depends on mutation status and PD-L1 expression and are thus equal to first-line options in primary disseminated disease; chemotherapy, possibly in combination with immunotherapy, immunotherapy as monotherapy, or (in case of targetable mutations or gene fusions) TKI.

Patients in stage III treated with CRT are at higher risk of recurrence[41]. The five-year survival rate in this group have been about 15-30%, but the recently implemented treatment strategy with addition of durvalumab therapy after concurrent CRT without progression have shown a survival advantage compared to placebo. Thus, for a subgroup of patients with unresectable stage III disease an increased long-term survival might crystallize[87].Treatment of progressive disease after CRT is generally dependent on previous chemotherapy given, duration of response after CRT, molecular profiling of the tumor and the patient's comorbidities.

Traditionally, chemotherapy has been given to patients with advanced disease and no treatment-predictive alterations, yielding a moderate increase in absolute survival. However, studies leading to implementations of maintenance therapy showed a median overall survival of almost 14 months with pemetrexed after four cycles of cisplatin and pemetrexed[40]. In second line, for patients treated with chemotherapy in first line, immunotherapy should be considered, and molecular profiling performed if not achieved before. Chemotherapy, e.g. docetaxel or other regimens, comprised the alternative before immunotherapy was introduced and might still be the preferred choice for patients with contraindications for immunotherapy. Chemotherapy as a single modality in first line will be given to fewer patients due to immunotherapy and the newly introduced combination of chemotherapy and immunotherapy, however, its role in second line is indisputable.

Patients harboring treatment-predictive *EGFR* mutations and treated with EGFR TKI (erlotinib, gefitinib) have been shown to have a median progression-free survival of 11 months versus 5.6 months with chemotherapy in first line. Osimertinib, previously given to patients with T790M resistance mutation detected at progression on first- or second-generation EGFR TKI, is now a first-line option with significantly longer median progression-free survival than with standard EGFR TKI[108]. A standard-option of second-line in these patients is not established and there are probably local variations. Chemotherapy or inclusion in studies might be alternatives.

Treatments of patients in this thesis work

In study I, III and IV the patients are surgically treated and received no pre-operative therapy. In study I, it can be presumed that the patients were not subject to any post-operative adjuvant treatment due to the time of surgery (the majority in the 1980's and 1990's). In study III and IV, comprising surgically treated early stage patients during 2005-2014, adjuvant treatment was given to selected patients with stage IB and in general to all patients with stage II or III, unless there were contraindications or patients' preferences were against it.

The patients in study II were diagnosed either with a primary lung cancer or with a lung cancer recurrence/progress in 2015-2016. Clinical follow-up ended in May 2018 and treatments reflect the recent rapid development of the therapy arsenal in lung cancer. Thus, immunotherapy was mainly a choice in second line palliative treatment and, when it comes to targeted therapies, crizotinib was the first hand choice for *ALK*-positive cases, followed by ceritinib at progression, and EGFR TKIs in first line included erlotinib, gefitinib and afatinib, whereas third generation EGFR TKI (osimertinib) did not have an indication of first line treatment at the time. *ALK*-rearrangements were tested for in a high proportion of the patients but *ROS1*-rearrangements only in a small subset of the patients. PD-L1 test increased during the study period but only a few patients received immunotherapy in first line. *BRAF* V600E was included in the NGS panel but was in general not included in the statement from the pathology department to the clinician since BRAF/MEK inhibition was not an approved lung cancer treatment at the time.

Aims

Overall aim

In this thesis work, I used a combined clinical and molecular approach to characterize lung cancers. With the overall and long-term goal to individualize and improve patient management, I aimed to investigate and correlate potential prognostic information from blood-based markers, tumor mutations, and clinical data.

Specific aims

Study I and II

The aims of study I and II were to investigate the frequency and variety of mutations, in particular *EGFR* mutations but also a wider spectrum, and furthermore to investigate the potential association between mutations and clinical features, including patient outcome. More specifically, the usefulness of mutation-specific IHC staining, the frequencies of mutations across 26 genes, and the prognostic impact of the most commonly occurring mutations, were examined.

Study III and IV

The aims of study III and IV were to investigate blood-based markers and their potential prognostic role in surgically treated lung cancer. In study III, five potentially lung cancer-related tumor markers in serum and their relation to lung cancer relapse were studied, followed by combined analyses of two of these serum markers and plasma cell-free circulating tumor DNA in study IV.

Patient material

The four studies of this thesis work are based on tumor specimens/cytology, blood samples and clinical variables from four different patient cohorts summarized in Table 5.

Table 5. Patient cohorts in the thesis.

	Study I	Study II	Study III	Study IV
Biobank	Two retrospective cohorts	Clinically NGS-tested patients	Southern Swedish Lung Cancer Study	Southern Swedish Lung Cancer Study
Patients (n)	Cohort I: 298 (35% F, 65% M) Cohort II: 52 (44% F and 56% M)	599 (51% F, 49% M)	n=107 (68% F, 32% M)	n=58 (66% F, 34% M)
Median age at lung cancer diagnosis	Cohort I: 66 Cohort II: 70	70	68	69
Smoking	Not available	Current/former: 87% Never: 12% Unknown: 1%	Current/former: 85% Never-smoker: 15%	Current: 57% Former: 28% Never: 16%
Period of inclusion for patients in the study	Cohort I: 1981-84 and 1995-97 Cohort II:1993-03	January 2015 to June 2016	2005-2011	2005-2014
Histology	Cohort I: 42% SqCC, 36% AC, 18% LCC, 3% SCLC Cohort II: 77% AC, 23% SqCC	70% AC 16% SqCC 14% Other	AC	AC
Disease stage	Surgically treated early stage patients	All stages	Surgically treated early stage patients	Surgically treated early stage patients
Tumor material	• Resection ○ FFPE ○ Fresh frozen	• Cytology • Biopsy/resection ○ FFPE	• Resection ○ FFPE ○ Fresh frozen	• Resection ○ FFPE ○ Fresh frozen
Blood samples			Pre-operative serum	Pre-operative serum and plasma

Abbreviations: F=females, M=males , FFPE=formalin-fixed paraffin embedded

Study I

Study I is based on two retrospective cohorts, cohort I and II, consisting of lung cancer patients in early stage treated surgically with curative intent and without pre-surgical treatment. No patients were overlapping in the two cohorts.

Cohort I consisted of 298 patients surgically treated for clinical stage I-IIIA primary lung cancer at the Lund University Hospital (renamed into Skåne University Hospital in 2010). The patients were treated either in 1981-1984 (48%) or 1994-1997 (52%). Cohort II comprised 52 patients with early stage tumors (N0, M0) surgically treated at the Lund University Hospital in 1993-2003, non-overlapping with cohort I and without pre-surgical treatment. Tumor material analyzed consisted of formalin-fixed paraffin embedded (FFPE) tumor tissue and/or fresh frozen tumor tissue.

Studies on these retrospective cohorts were approved by the Regional Ethical Review Board in Lund, Sweden (Registration no. 762/2004).

Study II

Starting in January 2015, lung cancer patients in the Southern Health Care Region in Sweden go through predictive mutational testing by NGS. Study II includes all 599 lung cancer patients in Region Skåne whose tumor(s) were mutationally profiled with NGS the first 1.5 year (January 2015 until June 2016), as described in Figure 10. Depending on the clinical situation, cytology or FFPE tumor specimens, either from biopsies or from resections, were used.

Data on multiple pre-defined variables were obtained from patient charts. These variables included baseline patient characteristics, lung cancer treatments and treatment outcomes. More specifically, the following variables were assembled: date of diagnosis, age, performance status, smoking history, family history of cancer, lung cancer-related treatments and outcome of treatments, tumor extension at time of diagnosis and tumor dissemination during follow-up including time to progression or recurrence and subsequent treatment lines. Date of diagnosis was defined as date of histological/cytological proof of lung cancer, or cancer if more specific diagnosis could not be reached for which the summarized evaluation of clinical, radiological and histological information led to lung cancer diagnosis. Smoking history was categorized as never-smoker, former smoker and current smoker. Smoking cessation within a year from lung cancer diagnosis was categorized as current smoker. A frequently suggested definition of never-smoker is <100 cigarettes, but in 599 patient records this definition was never used. This definition is thus probably not realistic to use in a retrospective study based on patient records. We included also patients with a very short period of irregular

smoking as never-smokers. Date of death and primary malignancies were obtained from the Southern Swedish Cancer Registry.

Chemotherapy response in a subgroup of patients within study II (101/297 patients with advanced or metastatic disease) were evaluated according to Response Evaluation Criteria for Solid Tumors (RECIST) 1.1 criteria[130] in order to study possible treatment-predictive associations. Patients included were treated with at least two cycles of platinum-based chemotherapy combination. Patients with discontinuation after less than four cycles for other reasons than progression was not included. Furthermore, in order to be able to evaluate chemotherapy response, radiotherapy against the lung tumor or mediastinum prior to or concomitant with chemotherapy was an exclusion criterion for this part of the study. Due to the retrospective nature of the study and available radiological examinations, a few modifications of the criteria were necessary. In one case, progression was confirmed purely by a clinical finding and in one case a progression was confirmed by a bronchoscopy statement compared to a previous bronchoscopy statement.

The study was approved by the Regional Ethical Review board in Lund, Sweden (registration no. 2014/32, 2015/575 and 2017/620).

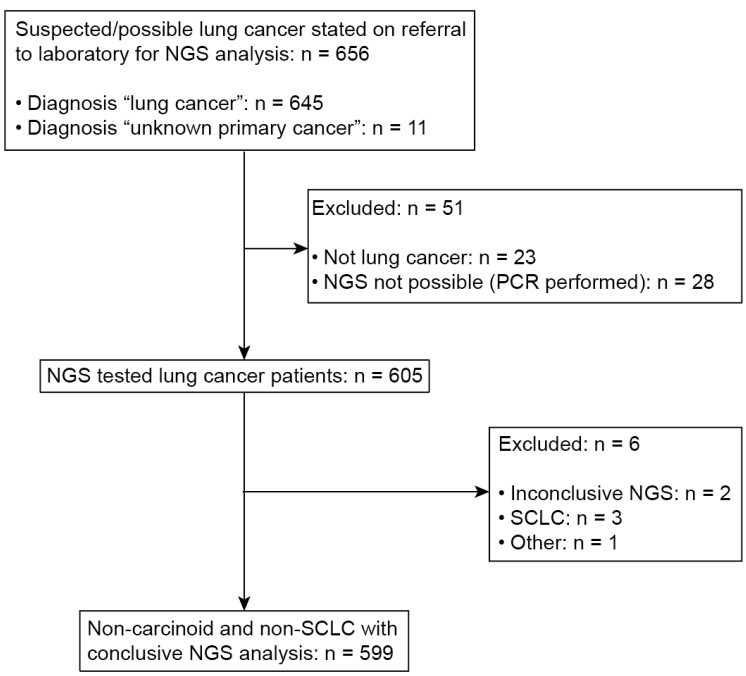


Figure 10. Inclusion of patients in study II.

Study III and IV

Study III and IV include patients from our internal biobanking, the Southern Lung Cancer Study. This consecutively collected biobank contains blood samples pre-and post-surgery and tumor specimens from surgically treated lung cancer patients in the Southern Swedish Health Care Region (including the counties Skåne, parts of Halland, Småland and Blekinge) between 2005-2014.Tumor specimens, blood, serum and plasma were stored at -80°Celsius (C). Patients signed a written consent and the study was approved by the Regional Ethical Review Board in Lund, Sweden (Registration no. 762/2004).

Pre-operative serum and tumor specimens and samples from 107 patients with lung adenocarcinoma stage I-IIIA surgically treated between 2005 until September 2011 were analyzed in study III.

Study IV included tumor specimens and pre-operative serum and plasma from 58 patients fulfilling the criteria of tumor in stage I-IIIA with mutation in either *BRAF*/*KRAS*/*EGFR* and adenocarcinoma histology. Figure 11 displays the selection process.

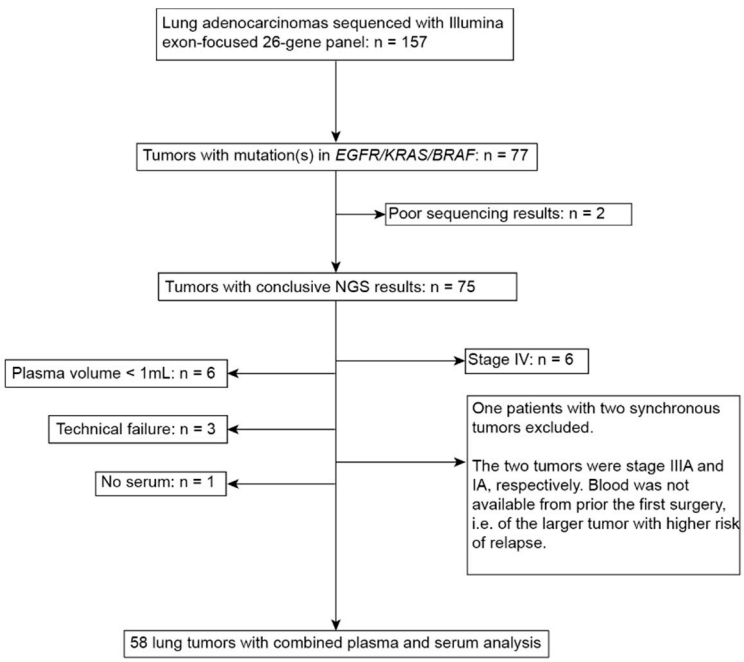


Figure 11. Study scheme, outlining inclusion of patients in study IV.

Baseline characteristics including age, date of surgery, smoking history, adjuvant treatment and lung cancer follow-up was obtained through patient records. Smoking cessation within a year from lung cancer diagnosis was categorized as current smoker. None of the patients in study III and IV received neoadjuvant treatment. Stage and histology were reviewed by a pathologist according to 7th edition of TNM and guidelines in WHO[18,48].

Methods

The analyses performed in this thesis include detection of mutations, protein expression, gene amplifications, gene fusions and blood-based markers. The methods used in each study are summarized in Table 6.

Table 6. Methods used in the four studies.

Method	Analysis	Study I	Study II	Study III	Study IV
IHC, mutated EGFR	Mutation status	X			
***EGFR* mutation-specific PCR**	Mutation status	X			
IHC, total EGFR	Protein expression	X			
FISH	Gene copy number	X			
Sanger sequencing	Mutation status	X			
Quantitative real-time PCR	Gene copy number	X			
NanoString technology	Fusion genes		X		
NGS	Mutation status		X	X	X
Electrochemiluminescence immunoassay	Tumor markers			X	X
Droplet digital PCR	Cell-free circulating tumor DNA				X

Immunohistochemistry, mutated and total EGFR (Study I)

In study I, *EGFR* mutation analysis by IHC and total EGFR expression were evaluated. Exon 21 point mutation p.L858R and exon 19 deletion E746_A750 in tumor specimens were analyzed by IHC staining with two rabbit monoclonal antibodies. An automated immunostainer was used. Tumors were evaluated on a

tissue microarray (TMA). With few exceptions triplicate 1.0-mm cores from each case on a tissue microarray (TMA) were evaluated. Positive cases regarding mutation-specific antibodies were repeatedly stained on whole tumor sections to evaluate tumor heterogeneity.

Mutation-specific PCR (Study I)

Tumors in cohort I that were mutation-positive according to IHC were further analyzed regarding *EGFR* mutations with a real-time PCR-based kit. This kit detects 29 *EGFR* mutations including 19 exon 19 deletions, T790M, L858R, L861Q, G719 point mutations, S768I and three exon 20 insertions. It relies on the amplification-refractory mutation system, that DNA polymerase effectively distinguish a match and a mismatch at the 3´end of a PCR primer. The kit also uses Scorpion primer/probe for detection.

Fluorescence in situ hybridization (Study I)

Gene copy number of *EGFR* was determined by fluorescence in situ hybridization (FISH) in cohort I. A dual-color *EGFR* probe, labeled with spectrum orange for *EGFR* (7p12) and spectrum green for centromere localization (CEP7) was used for hybridization. If less than 3 copies of the *EGFR* gene or chromosome was detected the case was considered negative. Tumors with amplification (EGFR/CEP7 ratio ≥2) or polysomy (≥3 copies of both EGFR gene and chromosome 7) were considered positive.

Sanger sequencing (Study I)

EGFR mutation status, including exon 18-21, of fresh frozen tumor specimens in cohort II were evaluated by Sanger dideoxy chain termination DNA sequencing. After PCR amplification of the DNA template, a sequencing reaction containing fluorescently labelled dideoxynucleotides mixed with normal nucleotides were assembled. When the DNA polymerase incorporates a fluorescently labeled dideoxynucleotide the DNA elongation process terminates which leads to DNA strands of different lengths with a fluorophore molecule in the end. The DNA strands of different length are separated by size through capillary electrophoresis and detected, followed by sequence analysis.

Quantitative real-time PCR (Study I)

In cohort II, *EGFR* gene copy number was evaluated by quantitative real-time PCR (qPCR). A SYBR Green-based assay was used. When the SYBR Green binds double-stranded DNA a fluorescent signal is emitted. For each PCR cycle the fluorescence increases and then reaches a threshold, C_T, of measurability. Thus, the more DNA product, the less PCR cycles need to be completed before the fluorescence reaches the threshold. Relative *EGFR* gene copy numbers were calculated by comparing C_T values of *EGFR* with an endogenous reference gene, in this case the genes of albumin or glucokinase respectively, in the target DNA with a reference DNA (commercial human diploid DNA). A ratio ≥ 1.5 was defined as amplification. In each run, serial dilutions of a reference DNA were used to generate a standard curve from which the amplification efficiency was determined. To ensure lack of DNA contamination a no template control with all essential components of the amplification reaction except the template was also included.

NanoString technology (Study II)

The focus in study II is on clinical data in relation to mutational status. However, in addition to NGS, a subset of tumors was analyzed for fusion genes by NanoString technology. The NanoString assay used herein is based on probes hybridizing to RNA purified from tumors. It has the advantages that no amplification of the sample is performed, and it tolerates samples of low quality such as FFPE derived RNA. After hybridization, the sample is loaded onto a cartridge. After purification, hybridized molecules are stretched out on an optical surface where they are counted and identified via the target-specific molecular identifier. For fusion gene analysis we used probes targeting exons on both sides of the junction, so called imbalance probes. An expressed fusion gene usually means a higher expression of the part involved in the fusion, for example the kinase domain of ALK. Specific probes spanning the junction were used to identify the exact fusion.

Next generation sequencing (Study II, III, IV)

Next generation sequencing (NGS), or massive parallel sequencing, is a high-throughput sequencing method enabling parallel analysis of multiple samples covering large part of the genome, as opposed to Sanger sequencing described above, which is restricted to smaller DNA sequences. In study II, III and IV NGS-

based mutation analysis was performed using the Illumina Trusight Tumor (TST) exon-focused 26 gene panel, including specific exons across 26 solid tumor-related genes, as previously described[45]. The assay is optimized for FFPE derived DNA. It uses two pools with gene specific oligonucleotides and each sample is sequenced on both strands of DNA. Samples are indexed and amplified using PCR and then pooled prior to sequencing. During sequencing, strands are attached to the surface of a flow cell where clusters are generated from one original DNA fragment. The clusters are sequenced by synthesis in a parallel manner. Fluorescently labeled nucleotides are incorporated base by base and the fluorescent emission is recorded. Since both strands of the original DNA are sequenced, formalin-induced artefacts misinterpreted as mutations can be avoided, as only alterations present in both strands are considered as true mutations. The tumors in this work were sequenced on a MiSeq instrument, the data were aligned to the Human UCSC hg19 reference genome and mutation detection was performed in the software Variant Studio, supplied by Illumina.

NGS-based mutation analysis is the main focus in study II, investigating mutation spectrum and clinical aspects. In study III, NGS-results for *KRAS* and *EGFR* were considered in sub-analyses and in study IV, mutations in either *KRAS*, *EGFR* or *BRAF*, were used as an inclusion criterion and the mutation detected in the tumor were subsequently analyzed in plasma (see description of droplet digital PCR below). In study II, we did a filtering of presumed single nucleotide polymorphisms (SNP) by identifying variants with a frequency of ≥ 1% in the general population as reported in Variant Studio by Illumina. Furthermore, we took into account that mutational load in lung cancer is high and although the gene-panel for NGS through the study period covers exons where hotspot alterations occur in different cancers, alterations detected have a wide variety of significance when it comes to tumorigenic effect. We distinguished some more potentially prognostic/predictive driver oncogenes (defined in[31,131-133]). The following were thus noted as driver alterations. In *KRAS:* mutations in codon 12,13 and 61, in *EGFR*: exon 19 deletions, exon 20 insertions, T790M, codon 719 and exon 18 deletion E709_T710delinsD, p.S768I (exon 20), exon 19 insertions, C797S (exon 20), codon 851 and 861 (exon 21), in *BRAF*: variants in codon 600, in *PIK3CA:* codon 542, 545 and 1047 variants, in *NRAS:* codon 12,13 and 61, in *MAP2K1:* codon 56 and 57 variants, in *ERBB2:* exon 20 insertions and in *MET:* variants involving the donor splice site position at the exon 14 intron-exon junction, leading to a *MET* exon 14 skipping variant.

Electrochemiluminescence immunoassay (Study III, IV)

Five tumor markers in pre-operative serum from early-stage lung cancer patients were analyzed in study III to investigate the relation between these markers and lung cancer recurrence in patients treated with curatively intended surgery. In study IV, the prognostic role for two of these tumor markers in serum in surgically treated early-stage lung cancer patients were further studied in combination with ctDNA in plasma. Tumor markers studied include carcinoembryonic antigen (CEA), cancer antigen 125 (CA 125), human epididymis protein 4 (HE4), neuron-specific enolase (NSE) and carbohydrate antigen 19-9 (CA 19-9). These blood-based protein or antigens are not routinely used in lung cancer monitoring but are implemented to varying degrees in the monitoring of other cancers. CEA is frequently used as a tumor marker in colon carcinoma[134], CA 125 is used in diagnostic procedure and management of ovarian cancer[135,136], HE4 is used as a marker in ovarian cancer[136], CA 19-9 is a marker in pancreatic cancer[137] and NSE is associated with tumors of neuroectodermal origin[138].

Tumor marker levels were analyzed with electrochemiluminescence immunoassay at the Division of Clinical Chemistry and Pharmacology, Department of laboratory Medicine, University Hospital, Lund, Sweden where these analyses are performed in clinical routine. In electrochemiluminescence immunoassay, the marker binds to specific antibodies and the complex is detected by luminescence produced during electrochemical reactions. The cutoff values for normal reference interval used were also applied in these studies; CEA < 5µg/L, CA 19-9 < 35 kE/L, CA 125 < 35 kE/L, NSE < 17 µg/L, HE4 premenopausal women < 92 pmol/L and postmenopausal women < 121 pmol/L. We did not have any information regarding menopause and presumed all women > 50 years to be postmenopausal. No cutoff values for men existed for HE4 and we chose to use the same reference interval as for postmenopausal women.

Droplet digital PCR (Study IV)

Droplet digital PCR (ddPCR) was used in study IV to detect cell-free circulating tumor derived DNA (ctDNA) in pre-operatively collected plasma (median 1500 microliters, range 1000-1600) from early-stage lung cancer patients. Prior to ctDNA analysis, tumors with a mutation in either of the genes *EGFR*/*BRAF*/*KRAS* had been selected and the specific mutation were analyzed in previously collected plasma from the respective patient. Twenty-three assays were designed based on the tumor NGS results and the performance of the assays was verified using mutated and normal human DNA as positive and negative controls respectively.

Digital PCR (dPCR) differs from PCR by compartmentalization of the sample, ultimately enabling a quantification of the target sequence. Instead of one reaction per sample, PCR-amplification occurs in each compartment, ideally containing one single molecule, in dPCR. In study IV, we used droplet digital PCR (ddPCR) in which the PCR reaction mix are randomly partitioned into thousands of water-in-oil droplets. Amplification of the target cleaves a fluorescent molecule from the target-specific probe, and each droplet is read as either positive or negative depending on its fluorescent intensity. At least two mutant-positive droplets were considered a positive finding of ctDNA. The specific ddPCR method used in this project was IBSAFE ddPCR, an improved method based upon ddPCR with improved limit of detection. This is achieved by sequential combination of linear amplification to increase copies of true target sequence followed by limited exponential amplification for signal generation, thereby increasing sensitivity and specificity by drastically minimizing the consequence of polymerase base-incorporation errors.

Statistical methods

Descriptive statistics were used in all four studies, with numbers and percentage presented for categorical data and median and range for continuous variables. Statistical analyses were performed with Stata 12.1 in study I, SPSS version 22 in study III and R version 3.5.2 in study II and IV.

In general, a significant probability value (p-value) for rejecting the null hypothesis was predefined as <0.05 - a commonly seen predefined value and approach. However, the p-value is a continuous measurement of evidence against the null hypothesis and strongly affected by sample size. Reasoning about the results might be a better approach than merely present p-values. Thus, in study II and IV (comprising the last projects of this thesis in chronological order) no p-value was defined as significant.

The main statistical tests used in the different studies of this thesis are listed in Table 7. Specifically, categorical variables were compared using Chi-square test or Fisher's exact test in situations when expected counts in at least one cell was <5. In case of two ordinal variables, linear by linear analysis was performed. Comparison of the distribution of a continuous variable over categorical variables were performed with the non-parametric (due to small data sets) Mann-Whitney's test and corresponding Jonkheree-Terpstra test if the categorical variable was ordinal. Differences in overall survival (OS) or recurrence-free time were evaluated with Kaplan-Meier plots and log-rank test to compare survival curves.

Cox regression was used to calculate an effect measure, a hazard ratio (HR) with confidence interval (CI). Cox regressions were either univariable or multivariable to adjust for confounding factors. Different clinical endpoints were used in the different studies as explained below.

Table 7. Statistical methods used in this thesis work.

Statistical method	Purpose	Study I	Study II	Study III	Study IV
Fisher's exact test / Chi-square test	Association between categorical variables	X	X	X	
Log-rank test / log-rank trend test	Comparison of survival curves	X	X	X	X
Cox regression	Estimate effects on survival curves with hazard ratio (HR)	X		X	X
Linear by linear test	Comparison of two ordinal variables			X	
Mann-Whitney test	Compare distribution of continuous variable over categorical variable			X	
Jonkheere-Terpstra test	Compare distribution of continuous variable over ordered groups			X	

[1]Fischer's exact test when expected counts in at least one cell was < 5.

Study I

Fisher's exact test was used for associations between the EGFR variables; mutations, total protein expression and gene copy number as well as comparisons between histopathological subtypes. Differences in overall survival (OS) according to the EGFR-related variables and other factors including gender and histology were evaluated with Cox regression analysis. Start of time period was date of surgery and endpoint was death of any reason. Follow-up time for patients alive was censored at the end of the study. Kaplan-Meier plots and log-rank test was used to compare differences in OS. Regression analyses were performed with and without adjustment for gender.

Study II

Associations of clinical variables with mutation findings were analyzed with the Chi-square test when comparing categorical variables; mutation status (*TP53*, *KRAS* and *EGFR* mutation or wildtype) and other primary malignancy than lung cancer (binary variable). Kaplan-Meier plots and log-rank test was used to compare progression-free survival (PFS) and OS, respectively, with mutation status (*KRAS*,

TP53 and the combination of these two alterations). The PFS endpoint was measured as the date of diagnosis until progression or death. Censoring of time period for patients without event was done at the date of their latest appointment at the lung department. The endpoint OS was defined as the date of diagnosis until death and time period for patients alive was censored at the end of the study. Survival analyses were performed using Kaplan-Meier plots and the log-rank-test.

Study III

In study III, tumor marker levels were analyzed in relation to tumor stage and *EGFR* and *KRAS* status, respectively, and in relation to recurrence. In the first part, tumor markers levels (continuous variables) were compared through stages with the Jonkheree-Terpstra test, while the number of elevated tumor markers (0-5 or 0-4 when NSE was excluded) of the markers through tumor stages were analyzed with linear by linear analysis. Relation between *EGFR* or *KRAS* status (binary variables) and discrete categorization of tumor marker levels were analyzed with the Chi-square test or Fisher's exact test. Tumor marker levels in tumors grouped according to *EGFR* and *KRAS* status respectively, were analyzed with the Mann-Whitney's test. The endpoint disease-free survival (DFS) was defined as the time period from surgery to recurrence. Time period for patients without recurrence was censored at time of death, diagnosis of a second primary lung cancer, diagnosis of a cancer of suspected origin other than lungs, or at the end of the study if alive (last check in patient files in February 2016). Survival analyses were performed using Kaplan-Meier plots and the log-rank test. Cox regression analyses were performed to adjust for stage and adjuvant treatment.

Study IV

Kaplan-Meier plots were constructed to illustrate differences in recurrence-free interval (RFI) in patients according to the binary variables ctDNA, tumor markers (CA 125 and/or CA 19-9) and combination of ctDNA and tumor markers (ctDNA and/or at least one tumor marker). RFI was defined as the period from surgery to recurrence. Time periods of patients without an event were censored at last lung cancer-related medical follow-up. Survival analyses were performed using Kaplan-Meier plots and the log-rank test. Cox regression analysis was performed to adjust for stage and adjuvant treatment.

Limitations

The studies in this thesis have several limitations, some related to the long duration of time for my thesis work (2011-2019 in parallel with clinical duty) but others associated with more study-related limiting factors. Through this time period, diagnostic procedures and treatments have advanced, thus potentially challenging the relevance of the work performed meanwhile.

The populations studied are generally small (in particular when there is need for relevant subgrouping by for example stage), retrospective and sometimes relying on old tumor material. When analyzing prognostic impact of different variables in lung cancer, stage is of course relevant to take into account but also other factors such as extent of lymph node involvement, mutational status and pathological factors which might be difficult due to the limited size of the cohorts. Study I includes patients treated in the early 1980's and 1990's. Staging of these tumors might be incorrect. Due to the diagnostic process and surgery performed at the time it would have been difficult, or probably not possible, to do it more accurately. The study also contains a variety of methods, merely because the second cohort was added to a pre-existing project and examined dependent on in-house methods at the time. Another aspect of the patient cohorts in this thesis is that, in contrast to study II (which includes all NGS-tested lung cancer patients in Region Skåne through 1.5 year), studies I, III and IV are not population-based, which would have provided a more accurate picture of for example frequency of mutations.

Regarding statistical analyses, a lot could probably have been improved with additional tests/considerations. Across the studies, different endpoints are used, which can contribute to making this thesis more difficult to read. Different endpoints were used partly due to necessity (e.g. PFS in the group of patients with advanced disease in comparison to RFI when studying recurrence in surgically treated patients) but also because we gained more insight into the different applications during the study progress. Furthermore, to conclude whether a patient had a lung cancer recurrence or not, or to distinguish a second primary lung cancer from a recurrence, can be challenging, in particular based on the retrospective design and available information in patient files. However, we have used all available information, i.e. radiology, molecular pathology, notes from multidisciplinary conferences and the clinical decisions, to define the events as accurate as possible.

Results and discussion

A rapid evolvement within the molecular lung cancer field accompanied this thesis work, with methodology expanding during the time period from single gene testing to need for multiplexed and, most recently, blood-based testing. The studies I-IV focus both on molecular and clinical profiling, investigating both blood- and tumor-based associations with patient outcome. Some results, with potential implications for future individualized and improved treatment and follow up for lung cancer, are discussed below.

Study I and II: Mutational profiling and clinical outcome

Study I and II focus on mutation analysis in lung cancer. *EGFR* mutations and other EGFR alterations are investigated in study I, whereas the first 1.5 year of clinical lung cancer mutational profiling by NGS is presented in study II.

In study I, the two most common EGFR mutations L858R and exon 19 deletion p.E746_A750 were analyzed with mutation-specific IHC. IHC is a well-established method in pathology. In lung cancer pathology specifically, IHC has subsequently become more central due to the need of subclassification for treatment choice[18,82,97]. Furthermore, gene copy number of *EGFR* and EGFR expression were evaluated.

Predictive testing of primarily *EGFR* mutations was the basis of the clinical inclusion in study II. Since NGS test for mutational profiling of lung cancer was implemented in the clinic in 2015 (Region Skåne, Sweden), additional genetic alterations in lung tumors have gained clinical interest and relevance and we present multiple mutational findings from NGS results in this clinically defined Southern Swedish cohort with focus on prognosis. In a well-characterized sub-cohort comprising patients with stage IIIB/IV and treated with platinum-doublet chemotherapy we studied PFS and OS in relation to *KRAS*/*TP53* mutation/wildtype.

The different techniques used in these two studies reflect the changes towards multiplexed mutational approaches in line with advances in methods and the increasing knowledge of targetable alterations in lung cancer.

Study I

The main aim of study I was to evaluate EGFR detection by mutation-specific IHC and, in addition to *EGFR* mutations, also investigate *EGFR* gene copy number alterations and EGFR protein expression and the association between these variables.

Evaluation of mutation-specific antibodies by IHC revealed 17 mutation-positive cases in cohort I, of which nine were positive for exon 19 deletion and eight for L858R. With DNA-based methods only 13/17 cases could be verified. In two of the discrepant cases, the repeated staining with the more diluted antibodies, 1:100 on whole section, did not display any positive staining, thus a titer-dependent false positive staining could be suspected. However, three of the PCR/sequencing-verified cases with positive mutation-specific IHC using the concentration 1:10 of the antibodies stained negative when using the titer 1:100.

The two remaining discrepant cases showed heterogenous staining within invasive areas on whole section with the titer 1:100. We investigated invasive areas with varying staining in a tumor with a verified mutation using PCR and pyrosequencing. The mutation was detected by both methods in all areas, including areas negative with mutation-specific antibody staining. The two discrepant cases could thus represent unspecific staining or, less likely, truly mutated cells but too few for detection with PCR.

In all, 13 tumors in cohort I had verified mutations, eight with exon 19 deletions and five with L858R. There was a distinct difference in mutation frequency across histological types ($p=0.006$, Fisher's exact test) with 11/13 of the mutation-positive tumors in tumors of AC and 2/13 in the group of SqCC resulting in totally 10% mutation-positive AC and 1.6% mutation-positive SqCC in cohort I. All mutations were detected in females.

This study demonstrated that these mutation-specific antibodies were not reliable due to the discrepancies explained. Furthermore, *EGFR* mutation analysis (covering exons 18-21) of cohort II by Sanger sequencing revealed five mutations, of which none were any of the two that should be recognized by the antibodies. One of the five mutations was an 18-base pair (bp) exon 19 deletion and was actually detected by mutation-specific IHC staining but only with 1:10 dilution and not 1:100. The tumors were resected in different time periods, altogether ranging from 1981-2003, and paraffin-embedded tissue might have been prepared in different ways during these wide time period and impact on IHC performance due to different handling of the material cannot be excluded. The TMA construction also displayed some pitfalls, the cores near the edges were sometimes clearly not optimally stained. In a few cases, a triplet of cores could not be evaluated. Despite our concerns about unspecific staining, unclear dilution and unsure sensitivity and specificity an

obvious reason for not using these antibodies is of course the technical advances in molecular diagnostics. Molecular testing by IHC are not recommended in clinical use in the latest guideline of molecular testing since these antibodies are suboptimal and because of much better alternatives even in small tumor samples making these antibodies unnecessary[43], thus, entirely in line with our findings and the direction in the clinic where NGS was implemented in 2015 which is further described in study II.

By FISH, 294/298 tumors in cohort II could be evaluated of which 123 (42%) showed increased gene copy number (by amplification in 22 tumors and polysomy in 101 tumors). Increased *EGFR* gene copy number in cohort II, detected by qPCR, was found in 6/52 (12%) of the cases. The different frequencies of increased gene copy number in the cohorts might be derived from the different methods, FISH and qPCR. The former includes a subjective evaluation and in qPCR the cell content can have different amounts of normal cells that affect the result.

Immunostaining for EGFR expression was evaluable in all but one case in cohort I and 186 tumors (63%) were positive. In cohort II, all except one of the 52 tumors could be evaluated with IHC for total EGFR and 30 (59%) displayed positivity. In contrast to *EGFR* mutations which clustered in tumors of AC histology, total EGFR expression was associated with SqCC in both cohorts (statistically significant in the larger cohort I, $p<0.001$, Fisher's exact test). Expression of EGFR is frequent in NSCLC, in particular SqCC where it has been reported in 70%, and it has also been suggested as a negative prognostic factor[139]. When it comes to clinical applications, EGFR expression stands in the shadow of the successful prediction of TKI-response with *EGFR* mutations, but actually has a small role in lung cancer treatment. Necitumumab is a monoclonal EGFR antibody that competes with natural ligands and through binding to EGFR inhibits downstream signaling. In advanced SqCC the addition of necitumumab to cisplatin and gemcitabine has shown a very moderate benefit in particular for the patients expressing EGFR[140]. This treatment combination has not been adopted as a standard treatment in Europe[40].

Analysis of the relation between EGFR variables displayed an association between FISH positivity and total EGFR positivity by IHC ($p<0.001$, Fisher's exact test) in cohort I and in cohort II similar tendency between the corresponding EGFR gene copy number by qPCR and total EGFR IHC in cohort II ($p=0.04$, Fisher's exact test). No trend of association between total EGFR and mutations were detected in any of the cohorts. Mutations and FISH positivity in cohort I were associated but in the smaller cohort II with five *EGFR*-mutations and six tumors with increased gene copy number by qPCR no association was detected.

OS analyses were performed in cohort I. The follow-up time of the survivors was 16 years in median (range 8-30). Female gender was the strongest prognostic factor in this material, in line with several other studies[54]. Females had a significantly

better survival than males (p=0.001, log-rank test; HR=0.66, 95% CI=0.51-0.85). Log-rank tests and univariable analyses of the EGFR variables and histology (AC vs. SqCC) displayed a significantly worse outcome for patients with positive total EGFR compared to those with negative immunostaining of total EGFR (p=0.004, log-rank test; HR=1.4, 95% CI 1.1-1.9). A trend of worse outcome for patients with SqCC compared to AC, mutation-positive compared to mutation-negative and also for FISH-positive compared to FISH-negative but none of these analyses reached significance. The worse outcome in the group of total EGFR positive cases remained when adjusting for gender. Furthermore, the worse outcome for males became more obvious when excluding the 13 cases with *EGFR* mutation. It is difficult to interpret the potential prognostic impact of *EGFR* mutations in this material with only 13 detected mutations, furthermore, except for the 13 PCR/Sanger verified mutations, cohort I is only screened for mutations by the mutation-specific IHC staining.

Study II

Background

In 2015, mutational profiling of lung cancer by NGS was implemented in Region Skåne, meeting the need for a multiplexed testing instead of methods covering only few mutations like e.g. the mutation-specific antibodies (never established in clinical settings) or real-time PCR (until recently the standard method for predictive testing in lung cancer). Study II summarizes the mutational findings from the first 1.5-year of clinical NGS test of lung tumors in Region Skåne and describes the clinical lung cancer patient cohort subjected to NGS. Moreover, associations between frequently occurring mutations and clinical outcome in the group of patients with advanced lung cancer treated with platinum-doublet chemotherapy were analyzed.

Enrolled patients (n=599) were grouped according to the situation of the NGS test. Since synchronous and metachronous tumors were sometimes selected for NGS test, the total amount of tumors was 611. In the majority of the patients (n=519 patients, with 530 tumors), mutational profiling by NGS was performed as part of the diagnostic procedure, i.e. prior start of treatment. In this sub-cohort, referred to as De novo-diagnosed cohort, histology and clinical baseline data were similar to the entire cohort. A comparison with the total number of lung cancer diagnoses (carcinoids and SCLC excluded) in Region Skåne through the same time (to estimate the coverage of NGS test among newly diagnosed lung cancer cases) revealed a coverage of 68%. The reasons for exclusion from NGS test of the remaining approximate third of newly diagnosed lung cancers can be derived from four main reasons; i) surgically treated patients not routinely selected for molecular profiling, ii) no mandatory routine testing of SqCC, iii) tumor specimens not

sufficient for molecular pathology (as stated by a pathologist) and iv) tumors selected for NGS test but tissue specimen turned out to be insufficient for NGS and PCR-based molecular testing was performed instead. Nonetheless, this cohort gives a comprehensive picture of the mutational spectrum and clinical course among lung cancer patients in a single Swedish health care region. It can also be concluded that NGS as the primary test for treatment-predictive mutation test of lung cancer patients could be effectively implemented.

Mutation spectrum and clinical baseline information are presented for the entire cohort (599 patients), while treatments and survival were investigated among the 519 patients with the NGS test as part of the primary diagnostic process (i.e. prior to lung cancer treatment), called De novo-diagnosed cohort. Furthermore, within the De novo-diagnosed cohort we looked deeper into the groups treated with TKI or with standard platinum-doublet chemotherapy in first line, respectively. The study scheme is outlined in Figure 12.

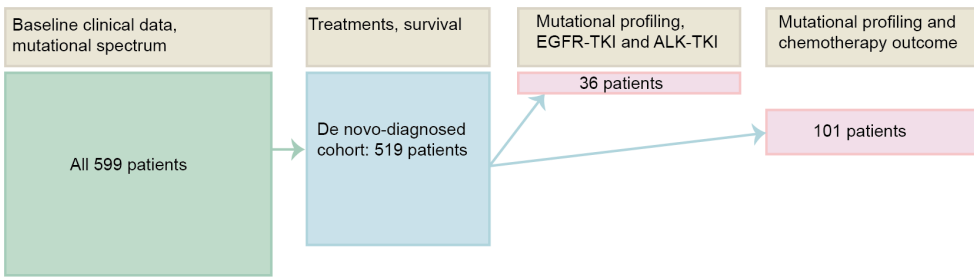


Figure 12. The scheme of study II in the top boxes and patients included in each theme in the lower boxes. The entire cohort comprises 599 patients for which clinical baseline data and mutational spectrum is described. In a majority of the cases (n=519 patients), NGS test was performed as part of the diagnostic procedure of the primary lung cancer, i.e. prior treatment. These patients comprise the De novo-diagnosed cohort for which treatment lines and results from NGS was put in relation to TKI-treatments and chemotherapy.

Description of the cohort

In the entire cohort of 599 patients we were able to group 99% into smokers, former smokers and never-smokers while pack-years could be estimated only in 78% of the cohort. Never-smokers comprised 12% (71 patients) of the cohort and at the other end of the spectrum were the 11% of the patients who had smoked 50 pack-years or more. A vast majority of the never-smokers (93%) had tumors of AC histology and 63% were women. Occupation was not as well documented as smoking but was reported in 85% of the cases. Specific information regarding exposure could be found only in 15% of the patient files, asbestos comprising the most frequently reported exposure (40 patients). By categorizing occupations into low-risk (e.g. health care, education or office work) it could be estimated that 48% of the never-smokers and 36% of the ever-smokers had low-risk occupation and no other reported exposure, however, the information was deficient.

Other variables in the baseline data for all 599 patients constituted other primary malignancies (other lung tumors and non-melanoma skin cancers excluded) and heredity of cancer. Information of other primary malignancies was obtained for all patients through the Swedish Cancer Registry and patient files. Twenty-one percent (n=126) had a history of another primary malignancy either before, in parallel with, or after lung cancer diagnosis/treatment. Most commonly occurring primary tumors were breast cancer and prostate cancer, in line with the incidence or tumors in Sweden. The frequency of other primary tumors was similar in never-smokers and ever-smokers, respectively. Neither did we detect any association between another primary malignancy (of any kind) and mutations in the three most frequently mutated genes; *TP53*, *KRAS* and *EGFR*.

Cancer in family was reported only in 185 (31%) of the patient files. Occurrence of any cancer in a first-degree relative (n=109) was not associated with mutations in any of the genes *TP53*, *KRAS* or *EGFR*. A family member diagnosed with lung cancer, specifically, was revealed in 39 cases, of which 35 had at least one first-degree relative with lung cancer. Hence, 5.8% of the patients reported a first-degree family member with lung cancer, which thus might be considered as the minimum proportion of lung cancer heredity in this cohort with information about heredity missing in approximately two thirds of the patient files. As data on family history come from patient-reported information (in patient files) and not the Cancer Registry, data should be carefully interpreted. Nonetheless, questions about heredity should probably be included in the management of newly diagnosed lung cancer patients.

Mutation spectrum

Mutational analysis by NGS in the 611 tumors revealed at least one variant (i.e. one of the 26 genes with at least one mutation) in 92%. *TP53* was the most frequently mutated gene in this 26-gene panel, followed by *KRAS* and *EGFR*, all three displaying variations between AC and SqCC. An overview of the mutations, both variants defined as drivers and other variants, in relation to histology, stage, gender and smoking status are presented in Figure 13.

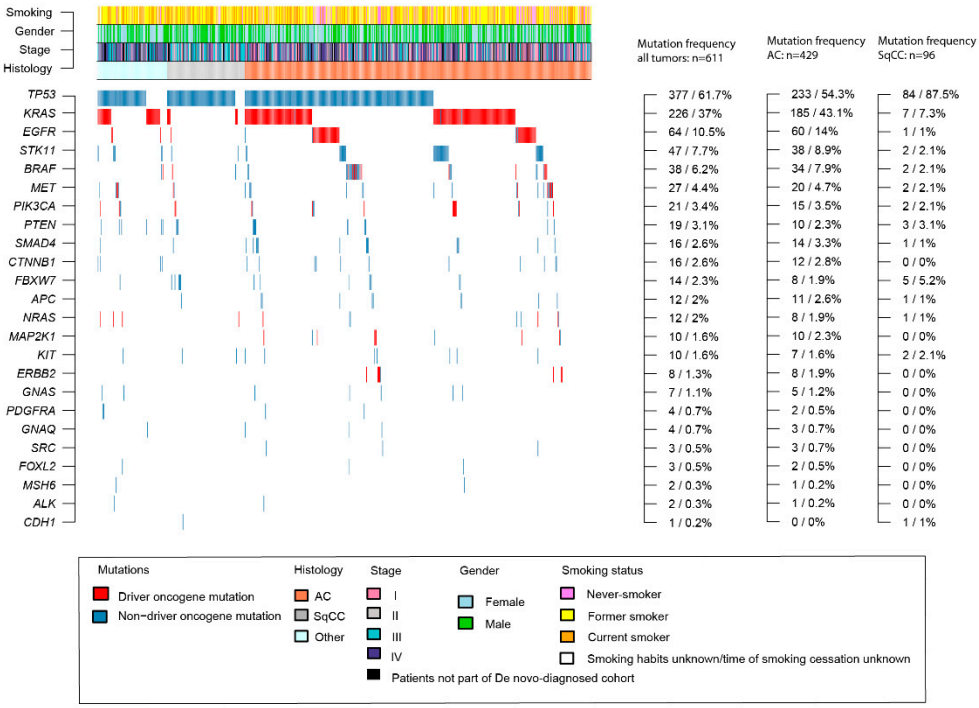


Figure 13. Mutations in the 611 tumors. Frequency of mutations in each gene for the entire cohort, all AC separately and SqCC separately are listed to the right.

The mutation spectrum largely reflects the European lung cancer population[141,142]. Among the oncogene drivers, *KRAS* mutations were most common, with 36% of the tumors harboring an oncogene driver mutation in *KRAS* followed by oncogene driver mutations in *EGFR* in 9.8% of the tumors. As expected, *KRAS* mutations were associated with smoking. Approximately a third (35%) had smoked 11-30 pack-years and another third (34%) over 30 pack-years. *EGFR* mutations were associated with no smoking or light smoking; 53% were never-smokers and additionally 10% had smoked no more than 10 pack-years.

Among the 60 *EGFR* mutations classified as driver mutations, half were exon 19 deletions and a third were L858R, four of the tumors with these common *EGFR* mutations also had a T790M. The remaining ten tumors with *EGFR* mutations displayed two exon 20 insertions, three L861X, one G719X, one exon 18 deletion and three with compound mutations. The three tumors with compound mutations comprised the following combinations; *EGFR*-mutations G719S and L861Q, *EGFR* mutations G719C and S768I and *EGFR* mutations L861Q and T725M.

Driver alterations in the other genes with predefined driver mutations (*BRAF*, *PIK3CA*, *MAP2K1*, *MET*, *ERBB2* and *NRAS*) were less common, with 7-14 tumors displaying oncogene driver mutations in each gene, respectively. The oncogene driver mutations were frequently mutually exclusive, but this was not obligate. Of the 14 tumors with driver mutations in *PIK3CA*, another driver alteration was present in nine of the tumors (of which one harbored three co-occurring driver mutations). In total, 14 tumors had two co-occurring drivers and one tumor had three, in most cases these driver alterations displayed different variant allele frequencies (VAF), suggesting different tumor clones. The clinical relevance of co-occurring driver mutations needs to be further studied. Notable, although only eight tumors with *MET* exon 14 splice site mutations were detected, no other driver mutations were present in these tumors. Similar to driver mutations in *EGFR,* and also in *ERBB2*, these *MET* mutations were associated with a non-smoking history. None of the patients *NRAS* or *PIK3CA* driver mutations, in which pack-years could be estimated, were never-smokers whereas never-smokers in patients harboring *BRAF* V600 or *MAP2K1* driver variants comprised 7% and 22%, respectively.

Some of the driver alterations are targetable with drugs; *BRAF* V600, *MET* exon 14 skipping mutations and *EGFR* mutations. In addition, targetable *ALK* and *ROS1*-rearrangements were analyzed in a subset of the patients. In the entire cohort one targetable alterations was detected in 100 tumors (16% of the tumors). However, due to incomplete testing of *ALK* fusions and *ROS1* fusions (combined results from NanoString technology and IHC/FISH performed in the clinic), *ALK* status was present in 90.2% and *ROS1* status only in 33%. In the group of never-smokers, comprising 12% of the 599 patients, a targetable alteration was detected in 59%. Among never-smokers *ALK* status was unknown in 5.6% and *ROS1* status in 73%. These findings highlight the need of an extensive molecular testing in lung cancer, especially since fusion genes were incompletely analyzed in this cohort and the NGS gene-panel could only detect one type of *MET* exon 14 splice site mutations and other potentially targetable fusion genes such as *RET* and *NTRK* fusions [123,124] and since the NGS gene-panel could only detect one type of *MET* exon 14 splice site mutations.

De novo-diagnosed cohort

In De novo-diagnosed cohort, i.e. the 519 patients with NGS test as part of the primary diagnostic procedure, we further evaluated the clinical implications of the mutational profiling. Survival curves for the entire De novo-diagnosed cohort displayed clear differences across stages, as expected (Figure 14).

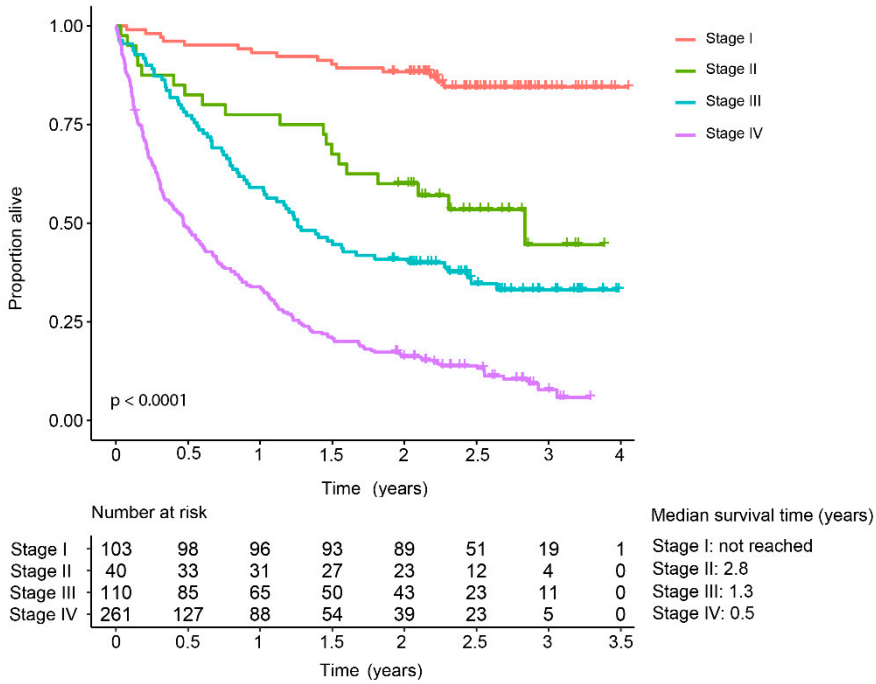


Figure 14. Survival plot, log-rank test and median survival time for De novo-diagnosed cohort. Five patients were excluded because stage could not be concluded.

TKI therapy and mutational findings in De novo-diagnosed cohort

Twenty-seven patients were treated with EGFR TKI in first-line and ALK inhibitor in nine patients. All nine patients with ALK inhibitor were given crizotinib, EGFR TKI varied but osimertinib did not have first-line indication at the time and was exclusively given to patients with progression on first-line EGFR TKI. These 27 and nine patients, respectively, are presented with some characteristics in Figure 15.

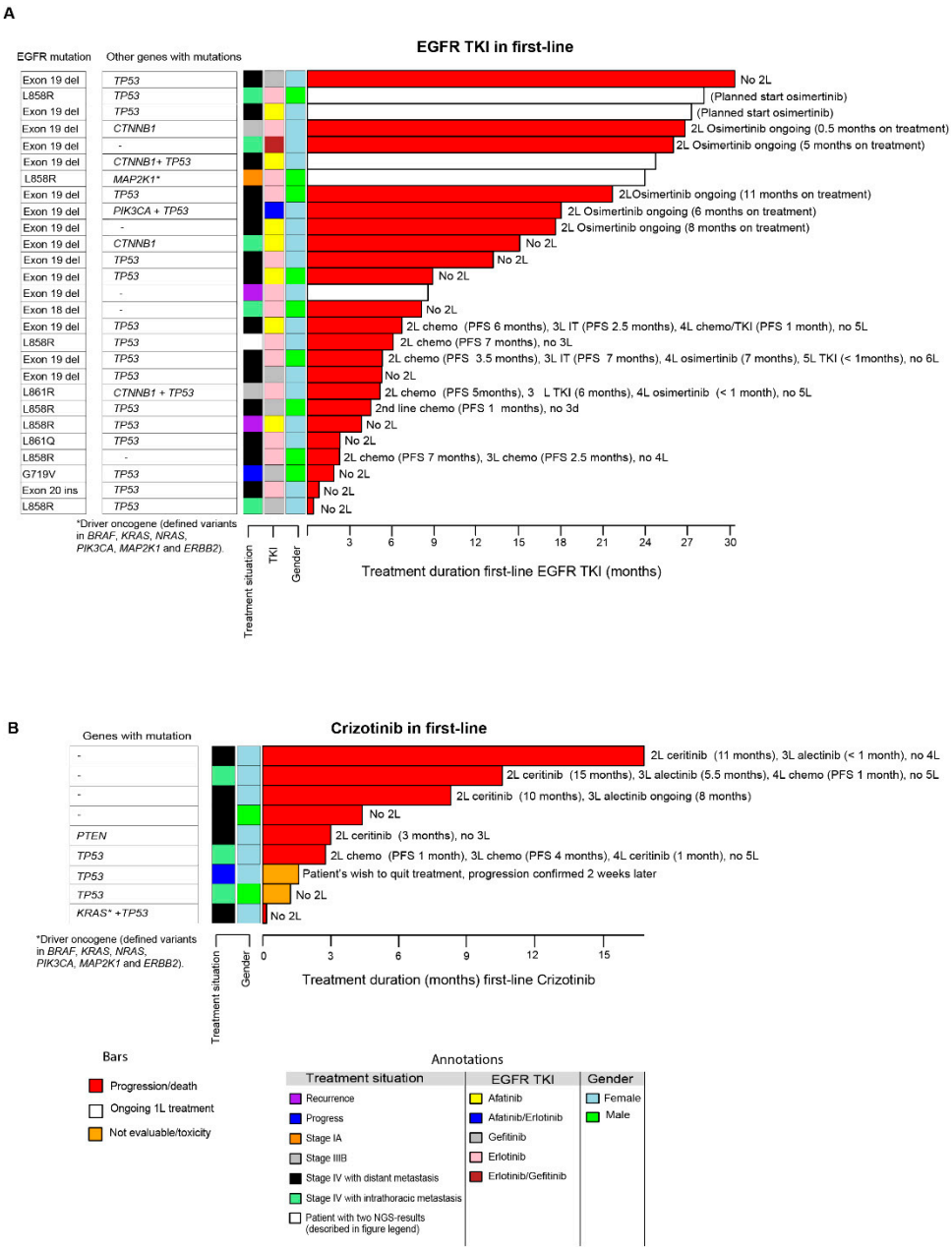


Figure 15. TKI as first-line treatment among patients in De novo-diagnosed cohort. A) Twenty-seven patients with EGFR TKI. For one patient (white box in treatment situation), stage could not be concluded. Given two NGS tests with very different results, one at time of diagnosis showing an *EGFR* mutation and no *KRAS* mutation and one after anti-EGFR therapy displaying a *KRAS* mutation and no *EGFR* mutation. Besides, ALK tests performed at the same time as NGS was negative and subsequently positive. Retrospectively, it cannot be concluded that this was two synchronous tumors from the start and not a intrathoracic metastatic disease. B) Nine patients with ALK inhibitor. Abbreviations: L=treatment line,

With so few patients it is difficult to distinguish any potential impact on mutational findings. An expected tendency of better response to TKI for patients with exon 19 del can be seen also in this small sub-cohort. One of the patients treated with EGFR TKI had a tumor of SqCC with an G719V mutation, where results from NGS also revealed a TP53 mutation. The patient was a current smoker at time of diagnosis and with a smoking history of 50 pack-years. Curative CRT resulted in partial remission but four months later treatment with EGFR TKI was initiated due to progression in the lungs. Treatment was ended two months later when progression, again in the lungs, was discovered. A new biopsy from the lungs concluded the diagnosis of SqCC, which had not been entirely clear in initial biopsies from start, and, in addition, the same treatment-predictive *EGFR* mutation was confirmed in NGS of the new biopsy. Due to the rarely seen *EGFR* mutation positive SqCC this case was reviewed with extra accuracy by our thoracic pathologists who agreed on the SqCC diagnosis. Although the tumor morphology was poorly differentiated, features of SqCC was recognized and the IHC staining (p40+ and TTF1- in both biopsies) supported SqCC. Although this patient did not respond on EGFR TKI, this case is a reminder of the rare but existing *EGFR* mutation-positive results in tumors of SqCC. Another case among the 27 patients on EGFR TKI illustrated in Figure 15 worth noticing was the patient with an exon 18 deletion (E709_T710delinsD). This patient was diagnosed in stage IV with tumor dissemination in the lungs and was on erlotinib for about eight months before progression. Less frequent *EGFR* mutations than exon 19 deletions and L858R have indeed been associated with shorter duration of response and might also respond differently depending on the type of EGFR TKI. For example, tumors with exon 18 mutations have been suggested to respond best to second generation EGFR TKI[133,143].

Platinum-based chemotherapy in stage IIIB/IV in De novo-diagnosed cohort

A group of 101 patients in stage IIIB/IV from De novo-diagnosed cohort was treated with platinum-doublet chemotherapy (inclusion and exclusion criteria for this group is described in section Patient material). Response to platinum-doublet chemotherapy was evaluated as partial remission in 43 patients, stable disease in 30 patients and progressive disease in 28 patients. We did not detect any correlation between mutational status in *KRAS*, *TP53* or a combination of *KRAS* and *TP53* or a *TP53* and *KRAS* combined. Neither OS nor PFS were significantly different between the patients grouped into *KRAS* or *TP53* mutation, respectively, versus wildtype or when divided into four groups according to mutation status; *KRAS* and *TP53* wildtype, *KRAS* mutation/*TP53* wildtype, *TP53* mutation/*KRAS* wildtype or *KRAS* and *TP53* mutation. A weak trend of worse outcome for patients with *TP53* mutations could be observed. In the literature, there has not been a clear consensus of the possible prognostic impact of *KRAS* mutations.

These variations between studies might depend on co-occurring mutations and on the cohort studied[144,145]. A meta-analysis revealed a non-significant worse PFS for patients on chemotherapy if *KRAS*-mutated tumor in comparison to patients with *KRAS* and *EGFR* wildtype tumors[146].

Summary

In summary, we examined the mutational spectrum in a well-characterized and population-based cohort and detected a large proportion of targetable alterations (59%) among the 12% never-smokers. Of the detected *EGFR* mutations, some were unusual variants and we found one case of *EGFR*-mutated SqCC. These findings highlight the importance of an extensive mutational and fusion gene panel and to not exclude non-adenocarcinomas from testing. Apart from the clear and intended predictive information from the mutational testing, we could not detect any additional impact on patient outcome. Specifically, we evaluated patients with advanced or disseminated lung cancer who were treated with chemotherapy. Chemotherapy is still the cornerstone in treatment within the palliative setting, either as a single modality or in combination with immunotherapy, but, today, no predictive markers for chemotherapy response exist and this is thus a topic that needs further investigation.

Normal tissue or blood samples were not available but could have been sequenced and used for comparison of detected variants in the tumor, which would have assured classification of possible SNPs or germline mutations other than those SNPs we identified through Illumina Variant Studio. Moreover, with multiple intra-individual tumor samples and samples from different metastatic sites, we would have analyzed heterogeneity in mutation patterns, revealing possible subclones within the tumors. NGS test in plasma could have enabled detection of tumor heterogeneity.

Study III and IV: Blood-based markers in early stage lung cancer

The minimally invasive nature of a blood sample and the opportunity of serial sampling makes blood-based markers interesting for clinical purposes, such as early diagnostics, prognostication, treatment prediction, and disease monitoring. Study III and IV include surgically treated early-stage (I-IIIA) lung cancer patients with pre-operative serum and plasma analyzed for tumor markers and ctDNA respectively. Most patients with NSCLC are diagnosed in an advanced stage beyond surgery opportunity[51]. However, this might change due to screening which is glimpsed at the horizon. Unfortunately, many patients, also with early stage NSCLC, develop

later lung cancer recurrence and additional prognostic variables are therefore needed in order to refine lung cancer follow-up and treatment. In these studies, we investigated blood-based markers and their relation to lung cancer relapse.

In study III, five tumor markers, CEA, NSE, CA 125, HE4 and CA 19-9, were measured in serum and correlated to outcome. The 107 patients comprised 68% women and 32% men with median age 68 years. Fifteen percentage were never-smokers and the remaining 85% were current or former smokers at time of diagnosis. The tumors comprised 52% stage I, 29% stage II and 19% stage III. However, one patient with stage I was diagnosed with liver metastases a few weeks after surgery, probably present prior to surgery but despite efforts with radiology not correctly diagnosed. This patient has later been discussed again and questions regarding lung origin of the liver lesions have been raised. In this study, the patient was included in accordance with how the case was managed in the clinic, i.e. an early stage lung cancer patient. Forty patients (37%) were diagnosed with lung cancer recurrence through the study period. Analyses of the tumor markers were successfully performed for all except NSE which is particularly sensitive to hemolysis[138] and for which analysis failed in 35 cases. Subsequent analyses comparing relations between tumor markers and clinicopathological variables were performed, with and without NSE. In total, 64% of the patients had at least one elevated tumor marker in pre-operative serum. Frequency of positive markers are displayed in Table 8.

Table 8. Distribution of patients by number of elevated tumor markers.

No of elevated tumor markers	Frequency	No of elevated tumor markers when NSE was excluded	Frequency
0	36%	0	39%
1	39%	1	40%
2	19%	2	15%
3	4%	3	5%
4	1%	4	1%
5	1%		

Discrete categorization of the tumor markers revealed a tendency of more markers in higher stages and analysis of tumor markers' continuous levels in relation to stage revealed a higher level of CA 125 in higher stages (Jonkheere-Terpstra test, p=0.008). Furthermore, the tumor markers were evaluated in relation to *EGFR* and *KRAS* status. Thirteen patients had an *EGFR* mutation and 33 patients had a *KRAS* mutation of which one patient had both an *EGFR* and a *KRAS* mutation. Mutation status was unknown in six patients. *EGFR* and *KRAS* status showed no relation to the tumor markers as dichotomized variables but analysis of continuous levels of tumor markers revealed a higher level of CA 125 in patients without *EGFR*

mutation. When excluding three uncommon *EGFR* mutations which we thought had unsure clinical relevance (two P848L and one M766I) the result for CA 125 remained and a difference between *EGFR* positive and negative patients appeared for HE4 with lower values of HE4 associated with EGFR positive tumors. Due to the small number of patients, *EGFR* mutations and the varying results, although stable for CA 125, when excluding three cases the value of these results are questionable.

DFS differed between patients with more tumor markers and we found a significant difference between patients with ≥ 2 positive tumor markers compared to < 2 (Figure 15).

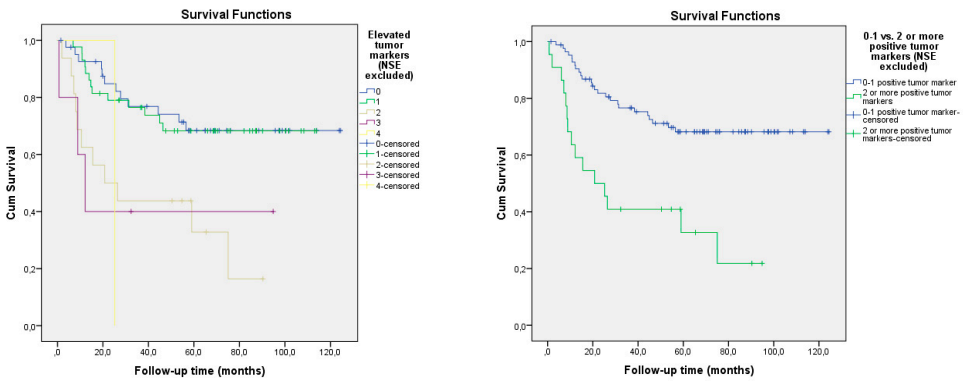


Figure 15. DFS in relation to number of positive tumor markers in pre-operative serum (NSE excluded). Patients grouped into A) 0-4 positive markers (log-rank trend test p=0.001) and B) < 2 markers and ≥ 2 markers, respectively (log-rank test p<0.001).

We further analyzed this relation between tumor markers and recurrence by Cox regression. In univariable analysis CA 125 (ten patients) and CA 19-9 (ten patients) were significantly associated with worse outcome. Furthermore, higher stage was associated with worse outcome, as expected, and adjuvant treatment probably since it is given to patients with higher stage and has a moderate effect on outcome which may not be detected in this small cohort. In multivariable Cox regression including CA 125, CA 19-9, stage and adjuvant treatment the two tumor markers remained associated with worse outcome but with a weak evidence (p=0.04 for each marker respectively). Patients with a positive CA 19-9 and/or CA 125 (18 patients in total, two with both positive CA 19-9 and CA 125) had a worse prognosis compared to patients without CA 19-9 and CA 125 adjusted to stage and adjuvant treatment (HR=2.8, 95% CI 1.3-5.7, p=0.006) adjusted to stage and adjuvant treatment. Due to few patients few positive tumor markers the results should be interpreted with caution.

In summary, the results point toward an association between stage and tumor markers, but we also demonstrated a possible addition of prognostic information beside stage in this small cohort. Some concerns of the potential utility of these markers are the few positive findings and the fact that they are not lung cancer-specific. Identifying lung cancer specific references and using serial monitoring of several markers, e.g. including markers positive in pre-operative serum in regular blood samples obtained at follow-up visits post-surgery to detect changes in marker levels, could be an option to use these markers. The potential value of these markers for early diagnostics was low, with only 64% cases positive for at least one tumor marker. Furthermore, we did not investigate these markers in a cohort without lung cancer for comparison of specificity and it is known that several factors might influence these markers such as different primary malignancies and benign conditions[147]. This should be taken into account and strengthen the hypothesis that a panel of markers should be necessary.

In study IV, we continued the investigation of blood-based markers and their possible prognostic role in early stage lung cancer by a combined analysis of the two most interesting serum markers from study III, CA 125 and CA 19-9, and cell-free circulating tumor-DNA (ctDNA). Fifty-eight patients were included in the study of which 25 patients were diagnosed with recurrence during our follow-up, 76% of these recurrences occurred within three years after surgery. Median follow-up time for patients alive and without lung cancer recurrence (n=33) was 3.3 years. Two tumors had mutations in two of the genes *EGFR*/*KRAS*/*BRAF*. Since the limited plasma volumes did not allow for multiple testing, the choice of mutation was based on availability of assays developed previously in other projects. In total, plasma was analyzed for *KRAS* mutations in 40 (69%) samples, *EGFR* mutations in 14 (24%) samples and *BRAF* mutations in 4 (7%).

Elevated tumor markers were detected in pre-operative serum from ten patients, four with CA 125, four with CA 19-9 and two with both markers present in serum. Eight of the patients with positive tumor markers were diagnosed with lung cancer recurrence. With low volumes of pre-operative plasma (median 1500µL), we detected seven patients with ctDNA, four of them in stage III and all of them with a *KRAS* mutation. Six of seven patients with pre-operative ctDNA detected in plasma had a lung cancer relapse. Two patients had both positive markers and ctDNA. Thus, 15 patients in total had either positive ctDNA and/or at least one positive tumor marker. Both tumor markers alone, ctDNA alone and the combination of ctDNA and markers as a dichotomized variable (ctDNA and/or at least one marker or neither ctDNA or tumor markers vs. neither ctDNA nor tumor markers) were associated with worse outcome when RFI was estimated with Kaplan-Meier curves and corresponding log-rank test (Figure 15).

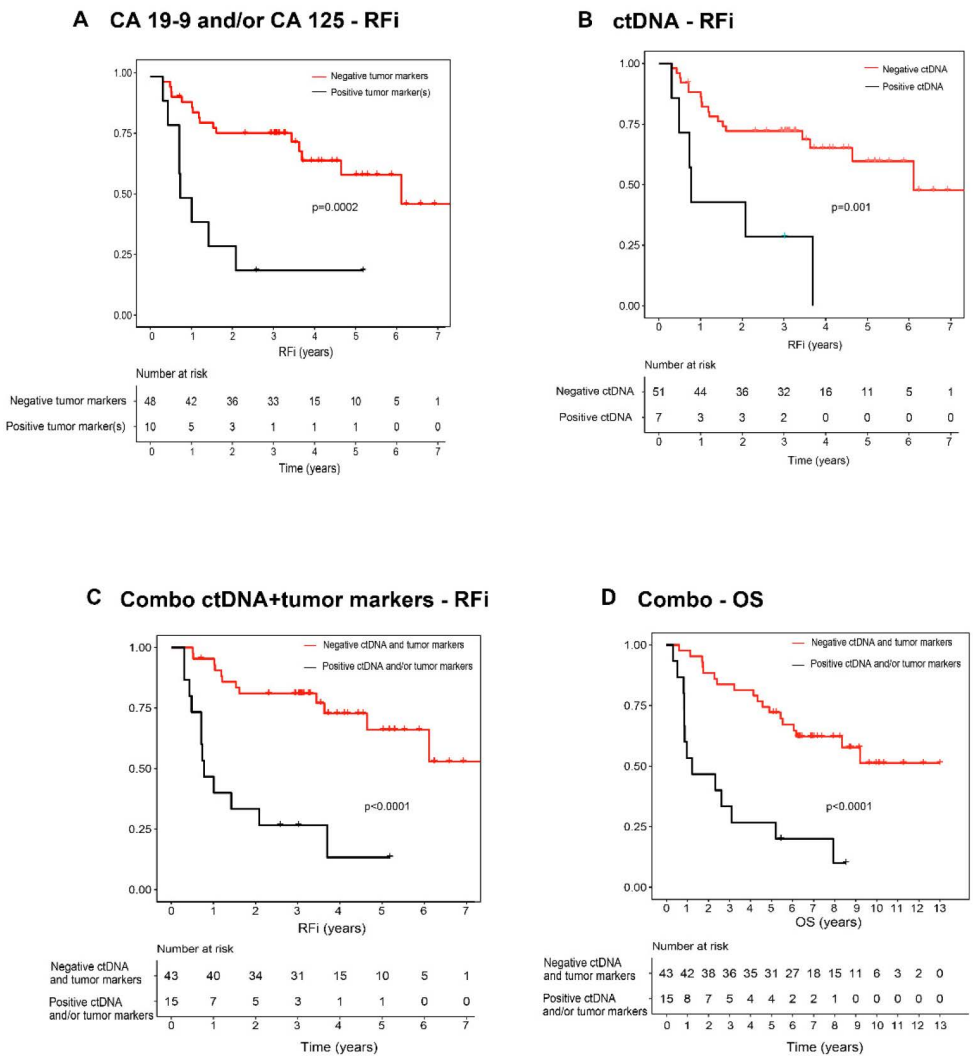


Figure 15. RFI and OS in relation to ctDNA and tumor marker status in pre-operative blood samples. A) RFI in patients with versus without tumor markers. B) RFI in patients with detected ctDNA versus without no detected ctDNA. C) RFI and D) OS in patients with ctDNA and/or tumor markers versus no detected ctDNA and no tumor markers. Log-rank test p-values are shown.

As in study III, positive findings of these blood-based markers tend to be more frequent in higher stages. We performed a Cox regression analysis considering ctDNA, adjuvant treatment, and stage. For ctDNA associated with recurrence and adjusted for stage and adjuvant treatment, we demonstrated HR 3.0 (95% CI 1.04-8.9), p=0.04, and for the combination (ctDNA and/or tumor markers), adjusted for stage and adjuvant treatment, HR 5.9 (CI 2.3-14.9, p<0.001). However, given the

small sample size and few events this analysis should be interpreted with great caution. Therefore, Figure 16, displaying RFI times, results of blood-based marker analyses and clinicopathological variables, is a more fair and perspicuous way of presenting these data. The extent of potentially stage-driven associations can be read from this overview of individual cases. Despite this, it is also seen that detection of blood-based markers is possible in tumors in stage I. This is an important finding in relation to screening which will probably be implemented in a near future and the detection of small lung cancers as well as benign lung nodules will increase. Blood-based markers could potentially be one of several tools to increase the accuracy in distinguishing nodules from lung tumors. The complexity of multiple variables contributing to lung cancer recurrence should be further studied and blood-based markers deserve to be investigated in larger cohorts and in particular using larger volumes of plasma for ctDNA detection.

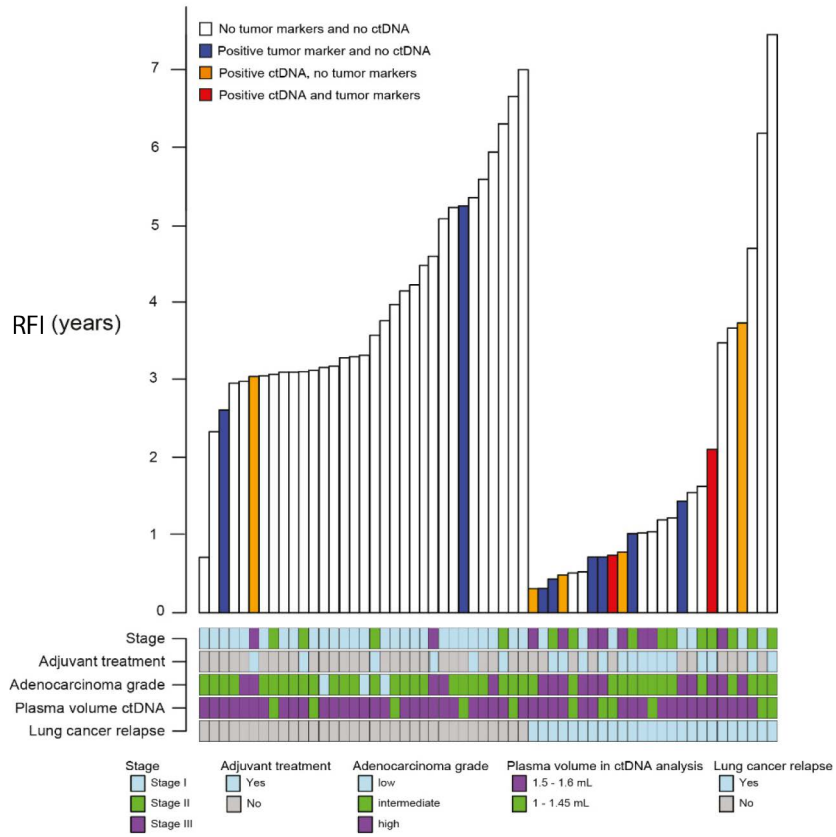


Figure 16. Clinicopathological factors and blood-based markers. Recurrence-free intervals (y-axis) are displayed in relation to results from ctDNA and tumor marker analyses (bar colors) in patients with subsequent lung cancer relapse (x-axis, to the right) or without lung cancer relapse (x-axis, left). Stage, plasma volume, adjuvant treatment and adenocarcinoma grade are described below each patient/bar.

Due to overlap with another study, one tumor was in in parallel re-classified as having a 10% large cell neuroendocrine tumor component and therefore classified as a combined LCNEC. Interestingly, this tumor had the highest amount of mutant ctDNA concentration (Figure 17).

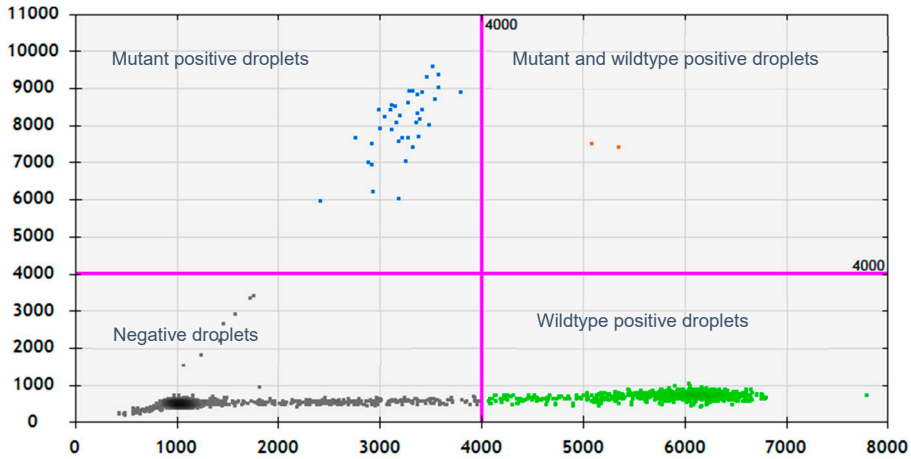


Figure 17. The case with the highest amount of ctDNA of the seven cases with positive ctDNA.

Conclusions

This thesis work comprises four studies with the overall aim to detect prognostic indicators based on analyses of blood samples and tumor specimens.

Study I and II

- Detection of *EGFR* mutations on the protein level by IHC staining is not feasible.
- A majority of lung tumors profiled by NGS, using a 26-gene pan-cancer panel, displays at least one mutation. Driver alterations are more frequent in AC compared to SqCC.
- Targetable alterations are common in never-smokers and extensive molecular profiling is recommended.
- No correlation between mutational status of the most frequently altered genes (*KRAS* or *TP53*) and PFS or OS can be demonstrated in advanced patients.

Study III and IV

- Blood-based (ctDNA or protein) markers can be measured pre-operatively in a subset of early stage lung cancer patients.
- There is a tendency of higher frequency of lung cancer recurrence in patients with pre-operative tumor markers and/or ctDNA, although this observation might derive from an association with higher stage and deserves to be further evaluated.

Future perspectives

The studies in this thesis cover only a small piece of the important lung cancer research field. Blood-based markers and mutational profiling in this work focused on prognostics. However, a lot of lung cancer research is ongoing, which is desirable in order to increase the knowledge and survival of this devastating disease. In this section, I present some views of future work and thoughts related to the projects within my thesis.

The need of accurate and broad mutational profiling and fusion gene detection in lung cancer might increase as knowledge of oncogene drivers, and the ways to target them, grow. This will require broader panels for mutations and fusion genes. The prognostic role of certain driver alterations, and not least the impact of co-occurring alterations, remain to be evaluated in order to individualize treatments further. In contrast to the growing field of targeted therapy, with corresponding predictive genetic alterations, chemotherapy lacks successful molecular markers of response. Although the role of chemotherapy as a first-line treatment is diminishing, it is still used both as a single treatment modality and in combination with for example immunotherapy and not the least, plays an important role in later treatment lines. The 101 patients, from within De novo-diagnosed cohort in study II, constitute a well-defined sub-cohort treated with platinum-doublet chemotherapy and with evaluation of chemotherapy response. We did not detect any association of PFS or OS with *KRAS* or *TP53* mutations (wildtype vs mutation) but this cohort will be used in future project in the Lung cancer research group, investigating e.g. possible associations between chemotherapy response and either gene expression signatures or DNA patterns caused by DNA repair deficiency.

The poor prognosis of lung cancer, much related to a large proportion of the patients being diagnosed in an inoperable stage, might change in favor of a larger proportion of lung cancers being detected at an early stage when lung cancer screening is implemented, which is probable considering the latest results from a screening trial. However, even early-stage patients unfortunately often experience lung cancer recurrence and more research is needed to improve this prognosis. Blood-based markers may here offer a minimally invasive way to monitor lung cancer worth studying in more detail. Indeed, despite the low number of tumors with detected ctDNA in study IV, and the suspected relation to stage for both tumor markers in serum and ctDNA, my results do not rule out a potential future clinical use for these

analyses. I therefore think that larger studies can yield more knowledge about the relation both to stage and to other variables, such as mutation status, adenocarcinoma grade, proliferation index in tumors and vascularization of the tumor, some of them included also in our studies. This could in turn result in a prognostic model with weighted variables adding prognostic information to stage. Furthermore, using a multiplexed assay for ctDNA detection could potentially make ctDNA more beneficial to use in early-stage lung cancer. Moreover, serial monitoring of ctDNA during follow-up after surgery could potentially be of value for early detection of lung cancer relapses and, in the end, inhibition of more advanced tumor dissemination by early treatment. Such an approach for longitudinal blood sampling is currently used both by us and by many other research groups.

To summarize, the subject of my thesis work is within one of the most rapidly evolving fields of oncology and I believe lung cancer management will continue to improve in a rapid pace.

Acknowledgement

This thesis work was done by the support from many persons and several collaborations. I would like to express my sincere gratitude to all of you who have supported me during this time. In particular, I would like to thank:

Maria Planck, my main supervisor. Thank you for introducing me to lung cancer research and for giving me this opportunity. I could not have had a better supervisor. Your enthusiasm and your wide and up-to-date knowledge in the field in combination with an ability to always achieve at least 100% both as a researcher and as a clinician is extraordinary and impressive. The translational Lung cancer research group you have founded is a platform for frontline research striving to improve lung cancer care. These years have been a pleasure!

Johan Staaf, my co-supervisor. Thank you for all sharp analyses and input to the studies and for helping me tackle statistical issues. As one of the primary investigators of the Lung cancer research group you have an ability to always think ahead and see future possible projects and collaborations.

Mats Jönsson, my co-supervisor. Thank you for patiently explaining pitfalls and necessities of laboratory work and sharing your wide knowledge in the field of molecular biology.

Åke Borg, my co-supervisor. Thank you for letting me be part of an impressive research environment.

All former and present members of the Planck/Staaf Lung cancer research group; *Anna Karlsson*, *Annette Salomonsson*, *Mats Jönsson*, *Christel Reuterswärd*, *Helena Cirenajwis, Elsa Arbajian*, *Bassam Hazem, Elin Nilsson, Marija Kotevska* and *Gudrun Oskarsdottir*. It has been a pleasure working with you and I'm looking forward to read about your future, exciting research. Thank you for all support and scientific discussions!

Hans Brunnström, for all valuable contributions to the studies and for the extensive work with the biobanks.

Kajsa Ericson Lindquist, for all the precise work with histology review in study II.

All *co-authors*, for valuable input on all projects. Especially, thank you *Pär-Ola Bendahl* for help with statistical issues, and thank you *Lao Saal* and *Anthony George* for fruitful collaborations with the Lung cancer research group.

All *lung cancer patients,* who generously contributed to the biobanks.

Klaus Bjerregaard, for administrative help with the Swedish Cancer Registry.

Björn Frostner and *Susanne André* for excellent administrative support.

Lena Mårtensson, for all work with Southern Lung Cancer Study.

Leif Johansson and *Per Jönsson*, for realizing early on the importance of biobanks and for help with the Southern Lung Cancer Study.

Carolina Andersson Rios, for always finding time for effective administrative help with the clinical parts of the projects.

Håkan Griph, for sharing your wide knowledge in lung cancer from the clinical view.

All funders of this work: *the Swedish Cancer Society*, *Governmental funding of clinical research within the national health services (ALF)*, *The Mrs Berta Kamprad foundation, BioCARE research program*, *Sjöberg foundation*, *Crafoord foundation*, *Maja and Hjalmar Leander foundation*, *Märta Winkler foundation*, *Anna-Lisa and Sven-Eric Lundgren foundation*, *The King Gustaf V Jubilee fund*, *John och Augusta Perssons stiftelse*, *the Gunnar Nilsson Cancer foundation,The Skåne University Hospital donation funds* and *Swedish Lung Cancer Study Group*.

My friends, for always being supportive. A special thanks to *Isabell*, *Emma*, *Viveka*, *Sofia* and *Klara* for chats and important fika.

Eva and *Jan*, for welcoming me into the Lundstedt family.

Mormor and *Farmor* for always believing in me.

My parents *Marie* and *Stefan*, and my brother *Erik* and his family for always supporting me through good and bad times.

Henrik, I would not have finished this without your support. I am looking forward to our future adventures!

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68(6): 394-424.
2. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J* 1950; 2(4682): 739-48.
3. Hartmann-Boyce J, Begh R, Aveyard P. Electronic cigarettes for smoking cessation. *Bmj* 2018; 360: j5543.
4. Newton JN, Dockrell M, Marczylo T. Making sense of the latest evidence on electronic cigarettes. *The Lancet* 2018; 391(10121): 639-42.
5. Hashim D, Boffetta P. Occupational and environmental exposures and cancers in developing countries. *Ann Glob Health* 2014; 80(5): 393-411.
6. Hamra GB, Guha N, Cohen A, et al. Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. *Environ Health Perspect* 2014; 122(9): 906-11.
7. Taylor R, Najafi F, Dobson A. Meta-analysis of studies of passive smoking and lung cancer: effects of study type and continent. *Int J Epidemiol* 2007; 36(5): 1048-59.
8. Samet JM, Avila-Tang E, Boffetta P, et al. Lung cancer in never smokers: clinical epidemiology and environmental risk factors. *Clin Cancer Res* 2009; 15(18): 5626-45.
9. Powell HA, Iyen-Omofoman B, Baldwin DR, Hubbard RB, Tata LJ. Chronic obstructive pulmonary disease and risk of lung cancer: the importance of smoking and timing of diagnosis. *J Thorac Oncol* 2013; 8(1): 6-11.
10. Koshiol J, Rotunno M, Gillison ML, et al. Assessment of human papillomavirus in lung tumor tissue. *J Natl Cancer Inst* 2011; 103(6): 501-7.
11. de Oliveira THA, do Amaral CM, de Franca Sao Marcos B, et al. Presence and activity of HPV in primary lung cancer. *J Cancer Res Clin Oncol* 2018; 144(12): 2367-76.

12. Digennaro M, Sambiasi D, Tommasi S, et al. Hereditary and non-hereditary branches of family eligible for BRCA test: cancers in other sites. *Hered Cancer Clin Pract* 2017; 15: 7.

13. Parry EM, Gable DL, Stanley SE, et al. Germline Mutations in DNA Repair Genes in Lung Adenocarcinoma. *Journal of Thoracic Oncology* 2017; 12(11): 1673-8.

14. Hwang SJ, Cheng LS, Lozano G, Amos CI, Gu X, Strong LC. Lung cancer risk in germline p53 mutation carriers: association between an inherited cancer predisposition, cigarette smoking, and cancer risk. *Hum Genet* 2003; 113(3): 238-43.

15. Bell DW, Gore I, Okimoto RA, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. *Nat Genet* 2005; 37(12): 1315-6.

16. Gazdar A, Robinson L, Oliver D, et al. Hereditary lung cancer syndrome targets never smokers with germline EGFR gene T790M mutations. *J Thorac Oncol* 2014; 9(4): 456-63.

17. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; 144(5): 646-74.

18. Travis WD BE, Burke AP, Marx A, Nicholson AG (ed). WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. 4th ed. Lyon, France: IARC Press; 2015.

19. Alavanja MC. Biologic damage resulting from exposure to tobacco smoke and from radon: implication for preventive interventions. *Oncogene* 2002; 21(48): 7365-75.

20. Phillips DH, Venitt S. DNA and protein adducts in human tissues resulting from exposure to tobacco smoke. *Int J Cancer* 2012; 131(12): 2733-53.

21. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 2012; 489(7417): 519-25.

22. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014; 511(7511): 543-50.

23. Luo SY, Lam DC. Oncogenic driver mutations in lung cancer. *Transl Respir Med* 2013; 1(1): 6.

24. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med* 2008; 359(13): 1367-80.

25. Naidoo J, Sima CS, Rodriguez K, et al. Epidermal growth factor receptor exon 20 insertions in advanced lung adenocarcinomas: Clinical outcomes and response to erlotinib. *Cancer* 2015; 121(18): 3212-20.

26. Zhang YL, Yuan JQ, Wang KF, et al. The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. *Oncotarget* 2016; 7(48): 78985-93.

27. Ohashi K, Sequist LV, Arcila ME, et al. Characteristics of lung cancers harboring NRAS mutations. *Clin Cancer Res* 2013; 19(9): 2584-91.

28. Odogwu L, Mathieu L, Blumenthal G, et al. FDA Approval Summary: Dabrafenib and Trametinib for the Treatment of Metastatic Non-Small Cell Lung Cancers Harboring BRAF V600E Mutations. *Oncologist* 2018; 23(6): 740-5.

29. Cardarella S, Ogino A, Nishino M, et al. Clinical, pathologic, and biologic features associated with BRAF mutations in non-small cell lung cancer. *Clin Cancer Res* 2013; 19(16): 4532-40.

30. Drilon A. MET Exon 14 Alterations in Lung Cancer: Exon Skipping Extends Half-Life. *Clin Cancer Res* 2016; 22(12): 2832-4.

31. Frampton GM, Ali SM, Rosenzweig M, et al. Activation of MET via diverse exon 14 splicing alterations occurs in multiple tumor types and confers clinical sensitivity to MET inhibitors. *Cancer Discov* 2015; 5(8): 850-9.

32. Drilon A, Cappuzzo F, Ou SI, Camidge DR. Targeting MET in Lung Cancer: Will Expectations Finally Be MET? *J Thorac Oncol* 2017; 12(1): 15-26.

33. Reungwetwattana T, Liang Y, Zhu V, Ou SI. The race to target MET exon 14 skipping alterations in non-small cell lung cancer: The Why, the How, the Who, the Unknown, and the Inevitable. *Lung Cancer* 2017; 103: 27-37.

34. Shaw AT, Engelman JA. ALK in lung cancer: past, present, and future. *J Clin Oncol* 2013; 31(8): 1105-11.

35. Nationellt kvalitetsregister för lungcancer. Available at: https://www.cancercentrum.se/vast/cancerdiagnoser/lunga-och-lungsack/kvalitetsregister/.

36. Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 2011; 6(2): 244-85.

37. Lungtumörer. Svensk Förening för Patologi - Svensk Förening för Klinisk cytologi. Available at: http://www.svfp.se/foreningar/uploads/L15178/kvast/thorax/Lungtumorer_v2.3_181105.pdf

38. Trawis WD BE, Müller-Hermelink HK et al. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart: World Health Organization Classification of Tumours. Lyon, France: IARC Press; 2004.

39. Nationellt vårdprogram Lungcancer 2018. Available at: https://www.cancercentrum.se.

40. Planchard D, Popat S, Kerr K, et al. Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2019.

41. Postmus PE, Kerr KM, Oudkerk M, et al. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2017; 28(suppl_4): iv1-iv21.

42. Rusch VW, Asamura H, Watanabe H, Giroux DJ, Rami-Porta R, Goldstraw P. The IASLC lung cancer staging project: a proposal for a new international lymph node map in the forthcoming seventh edition of the TNM classification for lung cancer. *J Thorac Oncol* 2009; 4(5): 568-77.

43. Lindeman NI, Cagle PT, Aisner DL, et al. Updated Molecular Testing Guideline for the Selection of Lung Cancer Patients for Treatment With Targeted Tyrosine Kinase Inhibitors: Guideline From the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. *J Thorac Oncol* 2018; 13(3): 323-58.

44. Kalemkerian GP, Narula N, Kennedy EB, et al. Molecular Testing Guideline for the Selection of Patients With Lung Cancer for Treatment With Targeted Tyrosine Kinase Inhibitors: American Society of Clinical Oncology Endorsement of the College of American Pathologists/International Association for the Study of Lung Cancer/Association for Molecular Pathology Clinical Practice Guideline Update. *J Clin Oncol* 2018; 36(9): 911-9.

45. Lindquist KE, Karlsson A, Leveen P, et al. Clinical framework for next generation sequencing based analysis of treatment predictive mutations and multiplexed gene fusion detection in non-small cell lung cancer. *Oncotarget* 2017; 8(21): 34796-810.

46. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011; 365(5): 395-409.

47. H. De Koning, C. Van Der Aalst, K. Ten Haaf, M. Oudkerk. Effects of volume CT lung cancer screening: mortality results of the NELSON randomised-controlled population based trial. Abstract PL02.05. IASLC 19th World Conference on Lung Cancer 2018.

48. International Association For the Study of Lung Cancer Staging Manual in Thoracic Oncology. USA: Editorial Rx Press; 2009.

49. Goldstraw P, Chansky K, Crowley J, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol* 2016; 11(1): 39-51.

50. Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The Eighth Edition Lung Cancer Stage Classification. *Chest* 2017; 151(1): 193-203.

51. Nationellt kvalitetsregister för lungcancer. Available at: https://www.cancercentrum.se/vast/cancerdiagnoser/lunga-och-lungsack/kvalitetsregister/.

52. Cuyun Carter G, Barrett AM, Kaye JA, Liepa AM, Winfree KB, John WJ. A comprehensive review of nongenetic prognostic and predictive factors influencing the heterogeneity of outcomes in advanced non-small-cell lung cancer. *Cancer Manag Res* 2014; 6: 437-49.

53. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982; 5(6): 649-55.

54. Nakamura H, Ando K, Shinmyo T, et al. Female gender is an independent prognostic factor in non-small-cell lung cancer: a meta-analysis. *Ann Thorac Cardiovasc Surg* 2011; 17(5): 469-80.

55. Avci N, Hayar M, Altmisdortoglu O, et al. Smoking habits are an independent prognostic factor in patients with lung cancer. *Clin Respir J* 2017; 11(5): 579-84.

56. Janjigian YY, McDonnell K, Kris MG, et al. Pack-years of cigarette smoking as a prognostic factor in patients with stage IIIB/IV nonsmall cell lung cancer. *Cancer* 2010; 116(3): 670-5.

57. Sun F, Ma K, Yang X, et al. A nomogram to predict prognosis after surgery in early stage non-small cell lung cancer in elderly patients. *Int J Surg* 2017; 42: 11-6.

58. Warth A, Muley T, Meister M, et al. The novel histologic International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification system of lung adenocarcinoma is a stage-independent predictor of survival. *J Clin Oncol* 2012; 30(13): 1438-46.

59. Russell PA, Wainer Z, Wright GM, Daniels M, Conron M, Williams RA. Does lung adenocarcinoma subtype predict patient survival?: A clinicopathologic study based on the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary lung adenocarcinoma classification. *J Thorac Oncol* 2011; 6(9): 1496-504.

60. Karlsson A, Jonsson M, Lauss M, et al. Genome-wide DNA methylation analysis of lung carcinoma reveals one neuroendocrine and four adenocarcinoma epitypes associated with patient outcome. *Clin Cancer Res* 2014; 20(23): 6127-40.

61. Karlsson A, Ringner M, Lauss M, et al. Genomic and transcriptional alterations in lung adenocarcinoma in relation to smoking history. *Clin Cancer Res* 2014; 20(18): 4912-24.

62. Ginsberg RJ, Rubinstein LV. Randomized trial of lobectomy versus limited resection for T1 N0 non-small cell lung cancer. Lung Cancer Study Group. *Ann Thorac Surg* 1995; 60(3): 615-22; discussion 22-3.

63. Blasberg JD, Pass HI, Donington JS. Sublobar resection: a movement from the Lung Cancer Study Group. *J Thorac Oncol* 2010; 5(10): 1583-93.

64. Fiorelli A, Caronia FP, Daddi N, et al. Sublobar resection versus lobectomy for stage I non-small cell lung cancer: an appropriate choice in elderly patients? *Surg Today* 2016; 46(12): 1370-82.

65. Veluswamy RR, Ezer N, Mhango G, et al. Limited Resection Versus Lobectomy for Older Patients With Early-Stage Lung Cancer: Impact of Histology. *J Clin Oncol* 2015; 33(30): 3447-53.

66. Altorki NK, Wang X, Wigle D, et al. Perioperative mortality and morbidity after sublobar versus lobar resection for early-stage non-small-cell lung cancer: post-hoc analysis of an international, randomised, phase 3 trial (CALGB/Alliance 140503). *Lancet Respir Med* 2018; 6(12): 915-24.

67. Nakamura K, Saji H, Nakajima R, et al. A phase III randomized trial of lobectomy versus limited resection for small-sized peripheral non-small cell lung cancer (JCOG0802/WJOG4607L). *Jpn J Clin Oncol* 2010; 40(3): 271-4.

68. Aokage K, Yoshida J, Hishida T, et al. Limited resection for early-stage non-small cell lung cancer as function-preserving radical surgery: a review. *Jpn J Clin Oncol* 2017; 47(1): 7-11.

69. Horn L, Johnson DH. Evarts A. Graham and the first pneumonectomy for lung cancer. *J Clin Oncol* 2008; 26(19): 3268-75.

70. Bendixen M, Jorgensen OD, Kronborg C, Andersen C, Licht PB. Postoperative pain and quality of life after lobectomy via video-assisted thoracoscopic surgery or anterolateral thoracotomy for early stage lung cancer: a randomised controlled trial. *Lancet Oncol* 2016; 17(6): 836-44.

71. Al-Ameri M, Bergman P, Franco-Cereceda A, Sartipy U. Video-assisted thoracoscopic versus open thoracotomy lobectomy: a Swedish nationwide cohort study. *J Thorac Dis* 2018; 10(6): 3499-506.

72. Kim MP. Robotic lobectomy leads to excellent survival in lung cancer patients. *J Thorac Dis* 2018; 10(Suppl 26): S3184-s5.

73. Oh DS, Reddy RM, Gorrepati ML, Mehendale S, Reed MF. Robotic-Assisted, Video-Assisted Thoracoscopic and Open Lobectomy: Propensity-Matched Analysis of Recent Premier Data. *Ann Thorac Surg* 2017; 104(5): 1733-40.

74. Subramanian MP, Meyers BF. Surgical Resection Versus Stereotactic Body Radiation Therapy for Stage I NSCLC: Can Randomized Trials Provide the Solution? *Cancers* 2018; 10(9): 310.

75. Lindberg K, Nyman J, Riesenfeld Kallskog V, et al. Long-term results of a prospective phase II trial of medically inoperable stage I NSCLC treated with SBRT - the Nordic experience. *Acta Oncol* 2015; 54(8): 1096-104.

76. Deng HY, Wang YC, Ni PZ, et al. Radiotherapy, lobectomy or sublobar resection? A meta-analysis of the choices for treating stage I non-small-cell lung cancer. *Eur J Cardiothorac Surg* 2017; 51(2): 203-10.

77. Auperin A, Le Pechoux C, Rolland E, et al. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer. *J Clin Oncol* 2010; 28(13): 2181-90.

78. Artal Cortes A, Calera Urquizu L, Hernando Cubero J. Adjuvant chemotherapy in non-small cell lung cancer: state-of-the-art. *Transl Lung Cancer Res* 2015; 4(2): 191-7.

79. Burdett S, Pignon JP, Tierney J, et al. Adjuvant chemotherapy for resected early-stage non-small cell lung cancer. *Cochrane Database Syst Rev* 2015; (3): Cd011430.

80. Non-small Cell Lung Cancer Collaborative Group. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *Bmj* 1995; 311(7010): 899-909.

81. Li M, Zhang Q, Fu P, et al. Pemetrexed plus platinum as the first-line treatment option for advanced non-small cell lung cancer: a meta-analysis of randomized controlled trials. *PLoS One* 2012; 7(5): e37229.

82. Scagliotti GV, Parikh P, von Pawel J, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. *J Clin Oncol* 2008; 26(21): 3543-51.

83. Socinski MA, Jotte RM, Cappuzzo F, et al. Atezolizumab for First-Line Treatment of Metastatic Nonsquamous NSCLC. *N Engl J Med* 2018; 378(24): 2288-301.

84. Paz-Ares L, Luft A, Vicente D, et al. Pembrolizumab plus Chemotherapy for Squamous Non-Small-Cell Lung Cancer. *N Engl J Med* 2018; 379(21): 2040-51.

85. Gandhi L, Rodriguez-Abreu D, Gadgeel S, et al. Pembrolizumab plus Chemotherapy in Metastatic Non-Small-Cell Lung Cancer. *N Engl J Med* 2018; 378(22): 2078-92.

86. Beatty GL, Gladney WL. Immune escape mechanisms as a guide for cancer immunotherapy. *Clin Cancer Res* 2015; 21(4): 687-92.

87. Antonia SJ, Villegas A, Daniel D, et al. Overall Survival with Durvalumab after Chemoradiotherapy in Stage III NSCLC. *N Engl J Med* 2018; 379(24): 2342-50.

88. Bylicki O, Barazzutti H, Paleiron N, Margery J, Assie JB, Chouaid C. First-Line Treatment of Non-Small-Cell Lung Cancer (NSCLC) with Immune Checkpoint Inhibitors. *BioDrugs* 2019.

89. Herbst RS, Soria JC, Kowanetz M, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* 2014; 515(7528): 563-7.

90. Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017; 357(6349): 409-13.

91. Takamochi K, Takahashi F, Suehara Y, et al. DNA mismatch repair deficiency in surgically resected lung adenocarcinoma: Microsatellite instability analysis using the Promega panel. *Lung Cancer* 2017; 110: 26-31.

92. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *N Engl J Med* 2015; 373(17): 1627-39.

93. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *N Engl J Med* 2015; 373(2): 123-35.

94. Reck M, Rodriguez-Abreu D, Robinson AG, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *N Engl J Med* 2016; 375(19): 1823-33.

95. Hellmann MD, Ciuleanu TE, Pluzanski A, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. *N Engl J Med* 2018; 378(22): 2093-104.

96. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004; 304(5676): 1497-500.

97. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004; 350(21): 2129-39.

98. Cohen MH, Williams GA, Sridhara R, et al. United States Food and Drug Administration Drug Approval summary: Gefitinib (ZD1839; Iressa) tablets. *Clin Cancer Res* 2004; 10(4): 1212-8.

99. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, et al. Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* 2005; 353(2): 123-32.

100. Yasuda H, Park E, Yun CH, et al. Structural, biochemical, and clinical characterization of epidermal growth factor receptor (EGFR) exon 20 insertion mutations in lung cancer. *Sci Transl Med* 2013; 5(216): 216ra177.

101. Herbst RS, Heymach JV, Lippman SM. Lung Cancer. *New England Journal of Medicine* 2008; 359(13): 1367-80.

102. Khozin S, Blumenthal GM, Jiang X, et al. U.S. Food and Drug Administration Approval Summary: Erlotinib for the First-Line Treatment of Metastatic Non-Small Cell Lung Cancer With Epidermal Growth Factor Receptor Exon 19 Deletions or Exon 21 (L858R) Substitution Mutations. *Oncologist* 2014; 19(7): 774-9.

103. Sequist LV, Yang JC, Yamamoto N, et al. Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. *J Clin Oncol* 2013; 31(27): 3327-34.

104. Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009; 361(10): 947-57.

105. Rosell R, Carcereny E, Gervais R, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *The Lancet Oncology* 2012; 13(3): 239-46.

106. Stewart EL, Tan SZ, Liu G, Tsao MS. Known and putative mechanisms of resistance to EGFR targeted therapies in NSCLC patients with EGFR mutations-a review. *Transl Lung Cancer Res* 2015; 4(1): 67-81.

107. Yu HA, Arcila ME, Rekhtman N, et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res* 2013; 19(8): 2240-7.

108. Soria JC, Ohe Y, Vansteenkiste J, et al. Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. *N Engl J Med* 2018; 378(2): 113-25.

109. Kim TM, Song A, Kim DW, et al. Mechanisms of Acquired Resistance to AZD9291: A Mutation-Selective, Irreversible EGFR Inhibitor. *J Thorac Oncol* 2015; 10(12): 1736-44.

110. Thress KS, Paweletz CP, Felip E, et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat Med* 2015; 21(6): 560-2.

111. Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 2007; 448(7153): 561-6.

112. Rikova K, Guo A, Zeng Q, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 2007; 131(6): 1190-203.

113. Shaw AT, Yeap BY, Mino-Kenudson M, et al. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. *J Clin Oncol* 2009; 27(26): 4247-53.

114. Inamura K, Takeuchi K, Togashi Y, et al. EML4-ALK fusion is linked to histological characteristics in a subset of lung cancers. *J Thorac Oncol* 2008; 3(1): 13-7.

115. Kazandjian D, Blumenthal GM, Chen HY, et al. FDA approval summary: crizotinib for the treatment of metastatic non-small cell lung cancer with anaplastic lymphoma kinase rearrangements. *Oncologist* 2014; 19(10): e5-11.

116. Solomon BJ, Mok T, Kim DW, et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N Engl J Med* 2014; 371(23): 2167-77.

117. Soria JC, Tan DSW, Chiari R, et al. First-line ceritinib versus platinum-based chemotherapy in advanced ALK-rearranged non-small-cell lung cancer (ASCEND-4): a randomised, open-label, phase 3 study. *Lancet* 2017; 389(10072): 917-29.

118. Peters S, Camidge DR, Shaw AT, et al. Alectinib versus Crizotinib in Untreated ALK-Positive Non-Small-Cell Lung Cancer. *N Engl J Med* 2017; 377(9): 829-38.

119. Camidge DR, Kim HR, Ahn MJ, et al. Brigatinib versus Crizotinib in ALK-Positive Non-Small-Cell Lung Cancer. *N Engl J Med* 2018; 379(21): 2027-39.

120. Gainor JF, Dardaei L, Yoda S, et al. Molecular Mechanisms of Resistance to First- and Second-Generation ALK Inhibitors in ALK-Rearranged Lung Cancer. *Cancer Discov* 2016; 6(10): 1118-33.

121. Shaw AT, Solomon BJ. Crizotinib in ROS1-rearranged non-small-cell lung cancer. *N Engl J Med* 2015; 372(7): 683-4.

122. Barlesi F, Mazieres J, Merlio JP, et al. Routine molecular profiling of patients with advanced non-small-cell lung cancer: results of a 1-year nationwide programme of the French Cooperative Thoracic Intergroup (IFCT). *Lancet* 2016; 387(10026): 1415-26.

123. Gautschi O, Milia J, Filleron T, et al. Targeting RET in Patients With RET-Rearranged Lung Cancers: Results From the Global, Multicenter RET Registry. *J Clin Oncol* 2017; 35(13): 1403-10.

124. Drilon A, Laetsch TW, Kummar S, et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *N Engl J Med* 2018; 378(8): 731-9.

125. Lima AB, Macedo LT, Sasse AD. Addition of bevacizumab to chemotherapy in advanced non-small cell lung cancer: a systematic review and meta-analysis. *PLoS One* 2011; 6(8): e22681.

126. Soria JC, Mauguen A, Reck M, et al. Systematic review and meta-analysis of randomised, phase II/III trials adding bevacizumab to platinum-based chemotherapy as first-line treatment in patients with advanced non-small-cell lung cancer. *Ann Oncol* 2013; 24(1): 20-30.

127. Reck M, Kaiser R, Mellemgaard A, et al. Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. *Lancet Oncol* 2014; 15(2): 143-55.

128. Revels SL, Lee JM. Anti-angiogenic therapy in nonsquamous non-small cell lung cancer (NSCLC) with tyrosine kinase inhibition (TKI) that targets the VEGF receptor (VEGFR): perspective on phase III clinical trials. *J Thorac Dis* 2018; 10(2): 617-20.

129. Consonni D, Pierobon M, Gail MH, et al. Lung cancer prognosis before and after recurrence in a population-based setting. *J Natl Cancer Inst* 2015; 107(6): djv059.

130. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009; 45(2): 228-47.

131. My cancer genome. Available at: https://mycancergenome.org

132. Hagemann IS, Devarakonda S, Lockwood CM, et al. Clinical next-generation sequencing in patients with non-small cell lung cancer. *Cancer* 2015; 121(4): 631-9.

133. Kobayashi Y, Togashi Y, Yatabe Y, et al. EGFR Exon 18 Mutations in Lung Cancer: Molecular Predictors of Augmented Sensitivity to Afatinib or Neratinib as Compared with First- or Third-Generation TKIs. *Clin Cancer Res* 2015; 21(23): 5305-13.

134. Figueredo A, Rumble RB, Maroun J, et al. Follow-up of patients with curatively resected colorectal cancer: a practice guideline. *BMC Cancer* 2003; 3: 26.

135. Doubeni CA, Doubeni AR, Myers AE. Diagnosis and Management of Ovarian Cancer. *Am Fam Physician* 2016; 93(11): 937-44.

136. Soletormos G, Duffy MJ, Othman Abu Hassan S, et al. Clinical Use of Cancer Biomarkers in Epithelial Ovarian Cancer: Updated Guidelines From the European Group on Tumor Markers. *Int J Gynecol Cancer* 2016; 26(1): 43-51.

137. Goonetilleke KS, Siriwardena AK. Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. *Eur J Surg Oncol* 2007; 33(3): 266-70.

138. Harmsma M, Schutte B, Ramaekers FC. Serum markers in small cell lung cancer: opportunities for improvement. *Biochim Biophys Acta* 2013; 1836(2): 255-72.

139. Scagliotti GV, Selvaggi G, Novello S, Hirsch FR. The biology of epidermal growth factor receptor in lung cancer. *Clin Cancer Res* 2004; 10(12 Pt 2): 4227s-32s.

140. Paz-Ares L, Socinski MA, Shahidi J, et al. Correlation of EGFR-expression with safety and efficacy outcomes in SQUIRE: a randomized, multicenter, open-label, phase III study of gemcitabine-cisplatin plus necitumumab versus gemcitabine-cisplatin alone in the first-line treatment of patients with stage IV squamous non-small-cell lung cancer. *Ann Oncol* 2016; 27(8): 1573-9.

141. Konig K, Peifer M, Fassunke J, et al. Implementation of Amplicon Parallel Sequencing Leads to Improvement of Diagnosis and Therapy of Lung Cancer Patients. *J Thorac Oncol* 2015; 10(7): 1049-57.

142. Planchard D, Remon J, Nowak F, Soria JC. Future Genetic/Genomic Biomarker Testing in Non-Small Cell Lung Cancer. *Am Soc Clin Oncol Educ Book* 2017; 37: 12-7.

143. Ibrahim U, Saqib A, Atallah JP. EGFR exon 18 delE709_T710insD mutated stage IV lung adenocarcinoma with response to afatinib. *Lung Cancer* 2017; 108: 45-7.

144. Arbour KC, Jordan E, Kim HR, et al. Effects of Co-occurring Genomic Alterations on Outcomes in Patients with KRAS-Mutant Non-Small Cell Lung Cancer. *Clin Cancer Res* 2018; 24(2): 334-40.

145. Tomasini P, Mascaux C, Jao K, et al. Effect of Coexisting KRAS and TP53 Mutations in Patients Treated With Chemotherapy for Non-small-cell Lung Cancer. *Clin Lung Cancer* 2018.

146. Pan W, Yang Y, Zhu H, Zhang Y, Zhou R, Sun X. KRAS mutation is a weak, but valid predictor for poor prognosis and treatment outcomes in NSCLC: A meta-analysis of 41 studies. *Oncotarget* 2016; 7(7): 8373-88.

147. Nakamura H, Nishimura T. History, molecular features, and clinical importance of conventional serum biomarkers in lung cancer. *Surgery Today* 2017; 47(9): 1037-59.