

Aus der Klinik für Nuklearmedizin, AB Klinische Nuklearmedizin,
der Medizinischen Fakultät Charité - Universitätsmedizin Berlin

DISSERTATION

***Die Asphärizität des präoperativen metabolischen Tumolvolumens in der
FDG-PET: Prognostische Bedeutung beim NSCLC und Einfluss der
Segmentierungsmethode***

***Asphericity of preoperative metabolic tumour volume in FDG-PET:
prognostic value in NSCLC and dependence on segmentation method***

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité - Universitätsmedizin Berlin

von

Christoph Chibolela

aus Berlin

Datum der Promotion: 30.11.2023

Vorwort

Vorwort

Teilergebnisse der vorliegenden Arbeit wurden veröffentlicht in: Rogasch, J.M.M., Furth, C., Chibolela, C., Hofheinz, F., Ochsenreither, S., Rückert, J.C., Neudecker, J., Böhmer, D., von Laffert, M., Amthauer, H., Frost, N.: Validation of Independent Prognostic Value of Asphericity of 18 F-Fluorodeoxyglucose Uptake in Non-Small-Cell Lung Cancer Patients Undergoing Treatment With Curative Intent. Clin Lung Cancer. 2020;21(3):264-272.e6.

Inhaltsverzeichnis

Tabellenverzeichnis.....	5
Abbildungsverzeichnis.....	6
Abkürzungen.....	7
Abstract.....	10
1 Einleitung.....	12
1.1 Epidemiologie des Lungenkarzinoms	12
1.2 Diagnostik, Staging und Prognose	13
1.2.1 Radiologisches Workup.....	13
1.2.2 Staging.....	13
1.2.3 Stadienbasierte Standardtherapie.....	15
1.2.4 Prognostische und prädiktive Faktoren.....	16
1.3 Technische Grundlagen der FDG-PET	18
1.3.1 Messprinzip.....	18
1.3.2 Datenkorrektur.....	18
1.3.3 Bildrekonstruktion.....	19
1.4 PET Parameter	20
1.4.1 MTV (Metabolisches Tumervolumen).....	20
1.4.2 SUV (Standardized Uptake Value).....	25
1.4.3 ASP (Asphärizität).....	26
2 Zielsetzung.....	29
3 Methoden.....	30
3.1 Patient*inneneinschluss und Follow-up	30
3.2 Akquisition der Bilddaten und Bildrekonstruktion	31
3.3 Bildauswertung und Segmentierung	33
3.3.1 Iterative Adaptive Thresholding (iAT).....	34
3.3.2 Fixed Thresholding.....	36
3.3.3 Regeln der Segmentierung.....	36
4 Statistische Analyse	38
4.1.1 Validierung der ASP als unabhängiger prognostischer Marker.....	38
4.1.2 Einfluss verschiedener Segmentierungsmethoden.....	43

Inhaltsverzeichnis

5 Resultate und Auswertung	44
5.1 Validität der ASP als prognostischer Marker	44
5.1.1 Datensatz.....	44
5.1.2 Cox Regression für DFS und OS.....	46
5.1.3 Subgruppenanalysen.....	51
5.1.4 Assoziationen von ASP mit anderen Variablen.....	53
5.2 Verschiedene Segmentierungsmethoden	58
5.2.1 Übereinstimmung von ASP und MTV.....	58
5.2.2 Erfolgsrate bei FT-Segmentierung.....	61
5.2.3 Prognostische Validität - FT vs. iAT.....	64
6 Diskussion	68
6.1 ASP als binarisierte vs. kontinuierliche Variable	69
6.2 Prognostische Validität der ASP in der Gesamtkohorte	70
6.3 Assoziation der ASP mit anderen Variablen	71
6.4 Prognostischer Wert der ASP nach Tumorstadium	73
6.5 Subgruppenanalyse: Chemotherapie	74
6.6 Assoziation der ASP mit anderen PET-Parametern	75
6.7 Segmentierungsmethode – Einfluss auf MTV und ASP	77
6.8 Segmentierungsmethode – Einfluss auf prognostischen Wert der ASP	79
6.9 Limitationen	82
7 Schlussfolgerung	84
Literaturverzeichnis	86
Eidesstattliche Versicherung	102
Lebenslauf	105
Publikationsliste	106
Danksagung	107
Bescheinigung Statistik	108

Tabellenverzeichnis

Tabelle 1: UICC-Stadien Einteilung, modifiziert nach UICC 8. Ausgabe.....	15
Tabelle 2: PET-CT Akquisition und Rekonstruktion.....	32
Tabelle 3: Dosimetrie.....	32
Tabelle 4: In dieser Arbeit verwendete Thresholding Methoden.....	36
Tabelle 5: Kovariablen für die univariable Cox Proportional Hazards Regression.....	39
Tabelle 6: Akaike Informationskriterium (AIC) für verschiedene Cox-Regressionsmodelle.....	41
Tabelle 7: Deskriptive Statistik.....	45
Tabelle 8: Übersicht: Überlebenszeiten nach UICC Stadium.....	46
Tabelle 9: Univariable Cox-Regression für das DFS.....	47
Tabelle 10: Multivariable Regression für das DFS.....	48
Tabelle 11: Univariable Regression für das OS.....	49
Tabelle 12: Multivariable Regression für das OS.....	50

Abbildungsverzeichnis

Abbildung 1: Vergleich zweier verschiedener Schwellenwerte (Threshold 1 und 2) bei homogener (a) und heterogener (b) Läsion.....	24
Abbildung 2: Verschiedene Körper und zugehörige ASP-Werte.....	29
Abbildung 3: Flowchart Patient*inneneinschluss.....	33
Abbildung 4: Tumorsegmentierung in ROVER am Beispiel eines zentral nekrotischen Lungenkarzinoms im rechten Unterlappen.	35
Abbildung 5: Segmentierungsalgorithmus nach Hofheinz et. al.	37
Abbildung 6: HR für DFS und OS: Stratifikation nach UICC Stadium.	53
Abbildung 7: HR für DFS und OS: Stratifikation nach T/N-Deskriptoren.	54
Abbildung 8: HR für DFS und OS: Gruppe der chemotherapeutisch behandelten Patient*innen, stratifiziert nach UICC Stadium.	55
Abbildung 9: Assoziation von ASP und UICC-Stadium sowie T- und N-Deskriptor.	56
Abbildung 10: Assoziation von ASP mit histologischem Subtyp und Resektionsstatus.	57
Abbildung 11: Korrelation von ASP mit MTV und SUV_{max}	58
Abbildung 12: PET-Parameter im Cox-Regressionsmodell.....	59
Abbildung 13: ASP und MTV verschiedener Segmentierungsmethoden.	60
Abbildung 14: Intraklassen-Korrelationskoeffizienten der ASP zwischen iAT und verschiedenen FT-Methoden.	61
Abbildung 15: Übereinstimmung des MTV zwischen iAT und verschiedenen FT-Methoden.	62
Abbildung 16: Erfolgsrate verschiedener Fixed Thresholding (FT) basierter Methoden.	63
Abbildung 17: Überstrahlungseffekte durch Weichteilnähe.	64
Abbildung 18: Entzündlicher Begleiterguss bei NSCLC im rechten Unterlappen.	65
Abbildung 19: MIP eines Patienten mit NSCLC im Stadium UICC III (Primärtumor im rechten Oberlappen).....	65
Abbildung 20: Minimal mögliches FT in Abhängigkeit des Verhältnisses von Tumorintensität (SUV_{max}) zu mittlerer Leberintensität (SUV_{mean}).	66
Abbildung 21: HR für DFS - verschiedene Segmentierungsmethoden.	67
Abbildung 22: HR für OS - verschiedene Segmentierungsmethoden.....	68
Abbildung 23: Segmentierbarkeit und Assoziation zu untersuchten Variablen.	69
Abbildung 24: Unterschiedliche Ausprägung der ASP in Abhängigkeit vom FT-Schwellenwert. ..	81

Abkürzungen

95%CI	95% Konfidenzintervall
Abb.	Abbildung
AC	Attenuation Correction (Schwächungskorrektur)
ACO	Ant Colony Optimization
ADC	Adenokarzinom
ADH	Antidiuretisches Hormon (Vasopressin)
adj.	adjustiert
AIC	Akaike Informationskriterium
ASP	Asphericity (Asphärizität)
AT	Adaptive Thresholding
CEA	Carcinoembryonales Antigen
CNN	Convolutional Neural Network
COPD	Chronisch obstruktive Lungenerkrankung
CT	Computertomographie
DFS	Disease Free Survival (erkrankungsfreies Überleben)
DICOM	Digital Imaging and Communications in Medicine
EANM	European Association of Nuclear Medicine
ECOG	Eastern Cooperative Oncology Group
EGFR	epidermal growth factor receptor
ESMO	The European Society for Medical Oncology
F18	Fluor-18
FDG	Fluordesoxyglucose
FLAB	Fuzzy Locally Adaptive Bayesian
FT	Fixed Thresholding
FWHM	Full Width at Half Maximum
GARAC	Gradient-aided region-based active contour
GLUT	Glucose Transporter
GPU	Graphics Processing Unit

Abkürzungen

HR	Hazard Ratio
iAT	iteratives, adaptives Thresholding
ICC	Intraklassen-Korrelationskoeffizient
IQR	Interquartilsabstand
keV	kilo Elektronenvolt
Kfz	Kraftfahrzeug
KM	Kontrastmittel
KRAS	Kirsten rat sarcoma virus
LK	Lymphknoten
LOR	Line Of Response
MIP	Maximum Intensity Projection
MRT	Magnetresonanztomographie
MTV	Metabolisches Tumervolumen
NSCLC	Non-small-cell lung cancer (nicht-kleinzelliges Lungenkarzinom)
OS	Overall Survival
OSEM	Ordered Subset Expectation Maximization
PET	Positronen-Emissions-Tomographie
PFS	Progression Free Survival (progressionsfreies Überleben)
PSF	Point Spread Function
ROI	Region of Interest
SCC	Plattenepithelkarzinom
SCLC	Small-cell lung cancer (kleinzelliges Lungenkarzinom)
SIADH	Syndrom der inadäquaten ADH-Sekretion
SUV	Standardized Uptake Value
SUV_{max}	maximale Standardized Uptake Value
SUV_{mean}	Mittlere Standardized Uptake Value
T/L-Ratio	Tumor-Leber-Verhältnis (SUV/SUV)
TLG	Total Lesion Glycolysis
TOF	Time Of Flight

Abkürzungen

UICC	The Union for International Cancer Control
VEGF	vascular epidermal growth factor
ZfKD	Das Zentrum für Krebsregisterdaten

Abstract

ENGLISCH

Background and Aim: Asphericity (ASP) is calculated from the surface area and volume of an object and is a measure of how the shape of an object differs from that of a sphere. Applied to [¹⁸F]FDG uptake in PET imaging ASP yields information on spatial heterogeneity of the metabolic tumour volume (MTV). Previous studies have demonstrated prognostic potential of this parameter in different types of cancer. The aim of the present thesis is to validate the ASP as a prognostic marker in non-small cell lung cancer (NSCLC) and to evaluate the stability of this parameter using different segmentation methods.

Patients and Methods: 361 Patients diagnosed with NSCLC in stages UICC I to IIIA who underwent preoperative PET and intended curative treatment at Charité-Universitätsmedizin Berlin between 2009 and 2017 were retrospectively included in the study. The ROVER-tool (ABX, Radeberg, Germany) was used for tumour delineation and for calculation of volumetric parameters including ASP. An iterative adaptive thresholding algorithm (iAT) was used as reference for MTV segmentation and was compared with 16 different thresholding segmentation methods. Cox regression for disease free survival (DFS) and overall survival (OS) was performed with respect to PET parameters (ASP, MTV, SUV_{max}), clinical (T/N descriptor, UICC stage), histological (ADC vs. SCC) and treatment related (Rx/1 vs. R0 resection, chemo- and radiotherapy y/n) variables.

Results: Out of 361 patients an event occurred in 178 and death occurred in 140 patients. Multivariable Cox regression demonstrated an independent prognostic value of ASP with a hazard ratio (HR) of 1,13 (95%CI: 1,02-1,25; p=0,017) and 1,13 (95%CI: 1,00-1,27; p=0,043) per 20% (Δ ASP) for DFS and OS, respectively. Additionally, male sex, ECOG performance status, UICC stage and age were shown to be independent predictors of survival. Stratified by UICC stage, ASP appears to be prognostic especially in stages I - II and also for patients who underwent chemotherapy. While some variation in the estimated effect sizes were observed for different segmentation algorithms, the direction of effect was overall consistent.

Conclusion: ASP was validated as prognostic factor for OS and DFS in patients with NSCLC stages I-IIIa undergoing resection with curative treatment intent. Prognostic effects were strongest in stages I and II and in patients who underwent chemotherapy. A possible predictive role of ASP has yet to be examined in prospective studies. Since ASP and therefore HRs can vary depending on the specific segmentation algorithm, spatial resolution of the scanner system and technique to estimate ASP, these factors have to be considered when applying the results of the present study to different settings.

Abstract

DEUTSCH

Hintergrund: Die Asphärizität (ASP) ist ein Parameter, der Oberfläche und Volumen eines Körpers miteinander in Beziehung setzt und ein Maß für die Formabweichung relativ zur idealen Kugel darstellt. Frühere Studien konnten das Potential der ASP des metabolischen Tumolvolumens (MTV) in der [¹⁸F]FDG-PET als prognostischer Marker für das Überleben bei verschiedenen Tumorerkrankungen demonstrieren. Ziel dieser Arbeit ist die Validierung der ASP als prognostischer Marker für das nicht-kleinzellige Lungenkarzinom (NSCLC) und die Evaluation der Robustheit dieses Parameters für verschiedene Segmentierungsalgorithmen.

Methoden: Eingeschlossen wurden 361 Patient*innen mit NSCLC im Stadium UICC I-IIIa mit präoperativer PET und kurativer Behandlungsintention, die im Zeitraum von 2009 bis 2017 an der Charité-Universitätsmedizin Berlin in Behandlung waren. Die Software ROVER (ABX, Radeberg, DE) wurde zur Auswertung der PET-Daten verwendet. Zur Segmentierung wurde dabei ein iterativer, Hintergrund-adaptiver Schwellenwert-basierter Algorithmus (iAT) als Referenzmethode verwendet und mit 16 weiteren Schwellenwert-basierten Algorithmen (FT) verglichen. Mittels Cox-Regression wurde der Einfluss von PET Parametern (ASP, MTV, SUV_{max}), klinischen (T/N Deskriptor, UICC Stadium), histologischen (ADC vs. SCC) und behandlungsbezogenen Variablen (R-Status, Chemotherapie, Radiotherapie) auf das erkrankungsfreie (DFS) und Gesamtüberleben (OS) untersucht.

Resultate: Von 361 Patient*innen zeigten 178 ein Event, 140 Patient*innen verstarben. In der multivariablen Cox Regression konnte die ASP als unabhängiger Prädiktor für das DFS und OS mit HR 1,13 (95%CI: 1,02-1,25; p=0,017) bzw. 1,13 (95%CI: 1,00-1,27; p=0,043) pro 20% (Δ ASP) bestätigt werden. Weiterhin waren Geschlecht, ECOG, UICC-Stadium, R-Status und das Alter unabhängige Prädiktoren für das OS und DFS. Stratifiziert nach UICC Stadium scheint die ASP v.a. in den Stadien I - II sowie für Patient*innen, die einer Chemotherapie zugeführt wurden, einen prognostischen Wert zu besitzen. Die geschätzten Effektgrößen weisen für verschiedene Segmentierungsmethoden kleinere Variationen auf, zeigen sich allerdings allesamt konsistent.

Schlussfolgerung: Die ASP konnte als unabhängiger prognostischer Marker für das OS und DFS bei Patient*innen mit NSCLC im Stadium UICC I - IIIa und kurativer Behandlungsintention validiert werden. Eine prädiktive Rolle der ASP in dieser Gruppe muss durch prospektive Studien weiter untersucht werden. Da die ASP und somit auch die mit ihr assoziierte Effektgrößen je nach verwendetem Segmentierungsalgorithmus, der räumlichen Auflösung des PET-Systems sowie der Schätzungsformeln zur Berechnung der ASP variieren können, müssen diese Faktoren bei der Übertragung der quantitativen Ergebnisse dieser Arbeit auf andere Zentren berücksichtigt werden.

1. Einleitung

1.1. Epidemiologie des Lungenkarzinoms

Das Lungenkarzinom stellt in Deutschland laut Daten des Zentrums für Krebsregisterdaten (ZfKD) des Robert Koch-Instituts mit einer Inzidenz von 21.500 Fällen/Jahr die dritthäufigste Krebsneuerkrankung von Frauen und mit 35.960 Fällen/Jahr die zweithäufigste der Männer dar (1). Gleichzeitig weist das Lungenkarzinom mit niedrigen 5-Jahresüberlebensraten von durchschnittlich 21% (Frauen) bzw. 15% (Männer) eine ungünstige Prognose auf und ist im Hinblick auf die Krebssterbefälle bei Frauen die zweithäufigste, bei Männern sogar die häufigste Tumorentität (1). Ähnliche Zahlen finden sich auch auf internationaler Ebene in der industrialisierten Welt (2). Der häufigste histologische Subtyp ist dabei mit ca. 80% das NSCLC, gegenüber dem deutlich selteneren kleinzelligen Lungenkarzinom (SCLC, small cell lung cancer) mit ca. 15-20% (3).

Die wichtigsten Risikofaktoren für die Entwicklung eines Lungenkarzinoms sind ein hohes Alter und der Raucherstatus (4). Nieraucher machen etwa 25% der Lungenkrebsfälle aus, hier stehen Passivrauchen und die berufliche oder umweltbedingte Exposition mit Radon, Kfz-Emissionen oder verschiedenen industriellen Karzinogenen im Vordergrund (5).

Patient*innen mit NSCLC werden in ca. 90% der Fälle vor Diagnosestellung symptomatisch. Zum klassischen Bild gehören pulmonale Symptome wie Husten bzw. veränderter Hustencharakter bei COPD, Dyspnoe und Hämoptysen durch Arrosion von Blutgefäßen. Je nach Lokalisation des Tumors und intrathorakaler Ausbreitung können verschiedene weitere Symptome auftreten. Eine begleitende B-Symptomatik mit Gewichtsverlust, Fieber, Schwäche und Anämie ist oft Ausdruck einer fortgeschrittenen Erkrankung. Bei Metastasierung können je nach Befallsmuster u.a. Knochenschmerzen oder neurologische Symptome auftreten (4). Weiterhin sind zahlreiche paraneoplastischen Syndrome wie Hypercalciämie (v.a. beim Plattenepithelkarzinom) und das Syndrom der inadäquaten ADH-Sekretion (SIADH) (v.a. beim SCLC) beschrieben. Sie treten bei ca. 10% der Erkrankungen auf, gelegentlich auch als Erstmanifestation (6). Bei asymptomatischen Patient*innen fällt das Lungenkarzinom als Zufalls- oder als Screeningbefund zumeist als pulmonaler Rundherd in Röntgen- oder Computertomographie (CT) auf.

1.2. Diagnostik, Staging und Prognose

1.2.1. Radiologisches Workup

Bei karzinomverdächtigen klinischen Untersuchungsbefunden und Risikofaktoren erfolgt in aller Regel schnell eine radiologische Diagnostik, die eine Röntgenuntersuchung und eine CT des Thorax umfasst. Letztere ist sehr sensitiv für das Vorliegen von Läsionen und erlaubt anhand morphologischer Kriterien bereits eine Einschätzung der Dignität (7). Für die histologische Diagnosesicherung ist zusätzlich die Bronchoskopie (v.a. bei zentralen Tumoren) oder die transthorakale Nadelaspiration (v.a. bei peripheren Tumoren) erforderlich. Bei Vorliegen eines Pleuraergusses kann auch durch Pleurapunktion eine zytologische Diagnose erreicht werden. Die diagnostische Bedeutung der FDG-PET liegt aktuell vor allem in der Malignitätsbeurteilung von Lungenrundherden, dem mediastinalen Lymphknoten-Staging und dem Nachweis von Fernmetastasen. Bei potentiell kurativer Therapieintention ist die kombinierte [¹⁸F]FDG-PET/CT inzwischen diagnostischer Standard. So konnte bereits in den frühen 2000er Jahren in prospektiven, randomisierten Studien eine signifikante Reduktion der Anzahl von Mediastinoskopien und eine Senkung der Rate vermeidbarer Operationen (d.h. bei Patient*innen mit bereits metastasierter Erkrankung) durch Verwendung der [¹⁸F]FDG-PET gezeigt werden (8,9).

1.2.2. Staging

Wie bei allen anderen soliden Tumoren hat die *Union for International Cancer Control* (UICC)-TNM-Klassifikation, aktuell in der 8. Auflage (2017), auch beim NSCLC eine maßgebliche Bedeutung in der Patient*innenbehandlung. Prognose und Therapieentscheidungen sind stark vom Ausmaß des Tumorleidens abhängig. Die TNM-Klassifikation ermöglicht eine kategorielle, anatomische Umschreibung der Tumorerkrankung anhand der Deskriptoren T (Tumor - Größe und lokale Invasivität des Primärtumors), N (Node - regionale Lymphknoten) und M (Metastasierung) (10). Auf Basis der TNM-Klassifikation lässt sich das anatomische Ausmaß der Tumorerkrankung einem UICC-Stadium zuordnen. Es stellt heutzutage die wichtigste Variable für die primäre Therapieplanung und Prognose dar (11).

Für das klinische **T-Staging** spielt die Kontrastmittel-CT die wichtigste Rolle. Hier können Tumordurchmesser und Nähe zu bzw. Infiltration von benachbarten Strukturen wie Hauptbronchus, Pleura bzw. Thoraxwand, Zwerchfell, Mediastinum, große Gefäße, Herz und Ösophagus

Einleitung

identifiziert und Aussagen über die technische Resektabilität getroffen werden. Höhere T-Stadien kennzeichnen größere und infiltrativ wachsende Tumoren. Eine Kombination der CT mit der [¹⁸F]FDG-PET kann hier Zusatzinformationen liefern, z.B. bei der Unterscheidung von Tumorgewebe und Atelektase (12).

Für das **N-Staging** liefert die kombinierte FDG-PET/CT einen erwiesenen Informationsgewinn. Eine CT-graphische Beurteilung der Lymphknotengröße allein ist nicht sehr spezifisch, da Lymphknotenvergrößerungen auch ohne Krankheitswert und reaktiv bedingt sein können, etwa infolge akuter oder chronisch-inflammatorischer Prozesse wie z.B. im Rahmen der Anthrakose bei Rauchern. Patient*innen mit NSCLC weisen sehr häufig chronische Lungengerüstveränderungen auf, die die radiologische Diagnostik erschweren. Die FDG-PET erhöht hier die diagnostische Aussagekraft, und führt zu einem robusteren N-Staging mit Konsequenzen bezüglich der weiteren Diagnostik und Therapie (8,13–15). So führten Shim et al. 2005 in einer prospektiven Untersuchung an 106 Patient*innen einen direkten Vergleich der PET/CT gegenüber der alleinigen CT durch und konnten einen signifikanten Zuwachs von Sensitivität (85% vs. 70%) und Spezifität (84% vs. 69%) für die Erkennung von Lymphknotenmetastasen zeigen (16). Die große Stärke in der kombinierten PET/CT liegt dabei vor allem in der erhöhten Spezifität. Problematisch ist allerdings weiterhin die nur geringe Sensitivität insbesondere für kleinere Lymphknotenmetastasen, diese lag beispielsweise in einer Untersuchung von Billé et al. 2009 bei nur 32% für Lymphknoten < 10mm (17).

Für das **M-Staging** wird in der aktuellen S3-Leitlinie des Lungenkarzinoms die Ganzkörper-[¹⁸F]FDG-PET/CT aufgrund ihrer hohen Sensitivität insbesondere für Metastasen im Skelettsystem (Sensitivität 92%, Spezifität 98%) und den Nebennieren (Sensitivität 89%, Spezifität 91%) empfohlen (4,18,19). Zur Detektion von zerebralen Metastasen kommt hier die Magnetresonanztomographie (MRT) zum Einsatz, da die PET aufgrund des physiologisch bedingt hohen cerebralen Glucosestoffwechsels hier eine geringere Sensitivität aufweist (20).

Die TNM-Klassifikation für maligne Tumoren wird regelmäßig aktualisiert und befindet sich zum Zeitpunkt der Veröffentlichung dieser Arbeit in der 8. Auflage (2017) (10). Wesentliche Änderungen zur vorherigen Ausgabe (7. Auflage, 2010) betreffen den T-Deskriptor (u.a. strengere Wahl der Größenschwellwerte und Hochstufung von Zwerchfellbeteiligung auf T4), sowie den M-Deskriptor (Unterscheidung zwischen solitärer und multipler Fernmetastasierung) (21). Die Einteilung in die einzelnen UICC-Stadien ist in Tabelle 1 zusammengefasst.

Einleitung

Tabelle 1: UICC-Stadien Einteilung, modifiziert nach UICC 8. Ausgabe (21)

T/M Stadium		N0	N1	N2	N3
T1	T1a	I A1	II B	III A	III B
	T1b	I A2	II B	III A	III B
	T1c	I A3	II B	III A	III B
T2	T2a	I B	II B	III A	III B
	T2b	II A	II B	III A	III B
T3		II B	III A	III B	III C
T4		III A	III A	III B	III C
M1	M1a,b	IV A	IV A	IV A	IV A
	M1c	IV B	IV B	IV B	IV B

1.2.3. Stadienbasierte Standardtherapie

Nach Richtlinien der European Society For Medical Oncology (ESMO) von 2017 erfolgt im sehr frühen Stadium UICC IA mit kleinem Primärtumor (T1 N0) eine alleinige operative Therapie, wobei die Lobektomie gegenüber limitierten Resektionsformen (z.B. Segmentresektion, Wedgeresektion) im Allgemeinen empfohlen wird (22). Im Stadium IB kann zusätzlich eine adjuvante Chemotherapie erwogen werden, wobei die Datenlage aufgrund der großen Heterogenität dieser Patient*innengruppe nicht eindeutig ist (23). Es scheinen Patient*innen mit größeren Tumoren und bestimmten histologischen Kriterien von einer adjuvanten Chemotherapie zu profitieren (24,25). Bei mangelnder kardiopulmonalen Reserve mit hohem perioperativen Risiko oder anderen Gründen, die gegen eine Operation sprechen, kann auch eine kurative primäre Strahlentherapie angeboten werden (22,26).

Im Stadium UICC II ist die Lobektomie mit systematischer Lymphadenektomie die Therapie der Wahl, wobei ggf. bei Infiltration von Thoraxwand, parietalem Perikard und Zwerchfell (T3 bzw. T4) zusätzlich eine erweiterte Resektion entsprechender Anteile erfolgt. Eine adjuvante, platinbasierte Chemotherapie (4 Zyklen) wird im Allgemeinen empfohlen. In der LACE-Studie führte die Behandlung mit Cisplatin und Vinorelbin bei R0-resezierten Patient*innen zu einer absoluten Erhöhung des 5-Jahresüberlebens um 11,6% im Vergleich zur Kontrollgruppe (27). Für den Zeitpunkt der Chemotherapie (adjuvant vs. neoadjuvant) konnte in einer Metaanalyse von Lim et al. aus dem Jahr 2009 kein Unterschied für das Gesamtüberleben gezeigt werden (28).

Einleitung

Patient*innen im lokal fortgeschrittenen Stadium UICC III stellen eine sehr heterogene Gruppe dar und erfordern häufig differenziertere, multimodale Behandlungskonzepte (29). Maßgeblich für die Therapieentscheidung ist dabei das Ausmaß der Lymphknotenbeteiligung. Im N0/1 Stadium ist primär die operative Versorgung indiziert, gefolgt von sequenzieller, adjuvanter Chemo- und Strahlentherapie. Dies gilt bei gegebener Resektabilität prinzipiell auch für das N2-Stadium bei Befall von nur einer Lymphknotenstation (single-level), ggf. erfolgt zuvor eine präoperative Induktionschemotherapie oder -strahlentherapie zur Verringerung der Tumorlast (22). Bei erweitertem Befall (multi-level) sowie im N3-Stadium erfolgt in multidisziplinärer Absprache häufig die definitive simultane Strahlenchemotherapie (4).

Im Falle einer Fernmetastasierung (UICC IV) wird im Allgemeinen von einer kurativen Therapie abgesehen. Die Indikation zur Lokaltherapie wird hier symptombezogen im Sinne eines palliativen Therapiekonzepts gestellt (4).

1.2.4. Prognostische und prädiktive Faktoren

Trotz der anerkannten, hohen prognostischen Aussagekraft der TNM/UICC-Einteilung kommt es dennoch auch innerhalb der einzelnen Stadien zu teilweise größeren Variationen der Überlebensraten. Das klinische Tumorstadium erklärt nur etwa 65% der Varianz im 5-Jahresüberleben. Unter Einschluss von Alter, Geschlecht und histologischem Subtyp als Kovariablen verbessert sich die prognostische Aussagekraft des Modells nur unwesentlich ($R^2 = 68\%$) (30). Damit stellt sich die Frage nach weiteren Einflussgrößen, die die Vorhersagegüte erhöhen können. Angesichts der hohen und wachsenden Zahl verschiedener Therapieformen und des unterschiedlichen Therapieansprechens auch innerhalb einzelner Tumorstadien stellt sich darüberhinaus die Frage nach Faktoren, die den Erfolg einer Therapie vorhersagbar machen. Erst dadurch wird eine individualisierte Beratung und Therapieentscheidung möglich. Aus diesem Grund ist die Suche nach und Identifikation von zusätzlichen, prognostischen und prädiktiven Faktoren von großer klinischer Bedeutung (31–34).

Weitgehend etabliert sind demographische Einflussfaktoren wie das Patient*innenalter, Geschlecht und der physische Allgemeinzustand, häufig gemessen mithilfe des ECOG-Performance Status oder des Karnofsky-Index (30,35). Auch die Charakterisierung des histologischen Subtyps (z.B. Adenokarzinom, Plattenepithelkarzinom) besitzt eine Bedeutung, da sie weiterführende diagnostische und therapeutische Strategien wesentlich beeinflussen kann (36). Im Zuge

Einleitung

diagnostisch-methodischer Fortschritte bei der Untersuchung von Lungentumoren wurden in den letzten Jahrzehnten jedoch zunehmend weitere potentielle Biomarker mit prognostischem und prädiktivem Wert identifiziert (31,35,37). Dazu gehören u.a. einfache Seromarker (z.B. Calcium) (38), Tumormarker (z.B. Carcinoembryonales Antigen, CEA), Genmutationen (z.B. Kirsten rat sarcoma virus, KRAS) (39), sowie auch verschiedene radiologische und nuklearmedizinische Marker (40,41). Neben Einzelparametern werden anhand großer Datensätze zunehmend auch Muster bzw. Signaturen (u.a. „Radiomics“, „Genomics“) aus einer Vielzahl an verschiedenen Parametern mithilfe von künstlicher Intelligenz untersucht, um auf statistischem Weg zu präziseren Prognosemodellen und individualisierten Therapien zu gelangen (42,43).

Der klare Vorteil bildgebender Verfahren gegenüber histologischen und molekularbiologischen Verfahren besteht in der Non-Invasivität und heutzutage auch breiten Verfügbarkeit in der industrialisierten Welt. Die Suche nach bildmorphologischen Tumoreigenschaften mit prognostischer oder prädiktiver Bedeutung erhält daher eine hohe Aufmerksamkeit in der gegenwärtigen Forschung (44–46). Neben statischen Faktoren zum Zeitpunkt der Diagnosestellung können mithilfe bildgebender Verfahren auch leicht Verlaufskontrollen erfolgen und somit die Dynamik einzelner Faktoren zu Aussagen über Prognose und Therapieansprechen herangezogen werden.

Zusammenfassend sieht sich die Onkologie mit einer zunehmenden Zahl an Variablen, die das TNM-Modell in seiner prognostischen Aussagekraft potenziell bereichern können, konfrontiert. Gleichwohl werden sich viele dieser Variablen als irrelevant oder redundant herausstellen. Daher ist eine Validierung und kritische Betrachtung der jeweiligen Prognosefaktoren essentiell und es gilt eben solche zu identifizieren, die tatsächlich den größten prognostischen und prädiktiven Nutzen mit sich bringen. Sie sollten möglichst unabhängig von bereits etablierten prognostischen Variablen sein, d.h. einen tatsächlichen informativen Zugewinn darstellen.

1.3. Technische Grundlagen der FDG-PET

1.3.1. Messprinzip

Grundprinzip der PET ist die Anreicherung einer radioaktiven Substanz in Geweben entsprechend ihrer Avidität für diese. Im Falle der FDG-PET wird als Radiopharmakon [¹⁸F]-Fluordesoxyglucose ([¹⁸F]FDG) verwendet, ein Glucose-Molekül, bei dem eine Hydroxylgruppe durch das Radionuklid Fluor-18 substituiert wurde. In stoffwechselaktiven Geweben mit hohem Glucoseumsatz und entsprechend hoher Expression von Glucosetransportern (GLUT) wird [¹⁸F]FDG vermehrt angereichert, so z.B. in Gehirn und Leberzellen, aber eben auch in Tumoren, Metastasen oder entzündlichem Gewebe mit üblicherweise hoher Teilungsaktivität und hochreguliertem Glucosestoffwechsel (47,48).

Das radioaktive Fluor-18 ist ein Positronen-Strahler und zerfällt mit einer Halbwertszeit von 110 min unter Aussendung von β^+ -Strahlung (Positronen). Positronen treffen praktisch sofort nach ihrer Entstehung auf ein Elektron wobei es zur sogenannten Annihilation kommt: Die beiden Teilchen vernichten einander und es entstehen zwei energiereiche Photonen (γ -Quanten), welche sich aufgrund des Gesetzes der Impulserhaltung diametral voneinander entfernen und jeweils eine Energie von 511 keV (die Ruheenergie eines Elektrons bzw. Positrons) aufweisen (49). Detektiert werden die Photonen von Detektoren (i.d.R. Szintillationskristalle), welche wie beim CT-Scanner ringförmig angeordnet sind (48).

Aufgrund der Lichtgeschwindigkeit ($c = 299\,792\,458\text{ m/s} \approx 30\text{cm pro Nanosekunde}$) treffen die zwei Photonen nahezu zeitgleich, d.h. mit nur wenigen Nanosekunden Unterschied, auf je einen Detektor und werden als „koinzident“ betrachtet. D.h. das zugehörige Annihilationsereignis wird auf der Verbindungslinie der beiden Detektoren (engl. *line of response*, LOR) vermutet. Gespeichert wird diese Information in Form eines Datenpunktes in einem Sinogramm (48,50).

1.3.2. Datenkorrektur

In der Realität liegen Annihilationsereignisse tatsächlich nicht immer in der detektierten LOR, etwa im Falle von Streuung von Photonen im Gewebe. Ferner kann es zur vollständigen Absorption eines Photons kommen. Maßgeblich für die Wahrscheinlichkeit, ob ein Photon ungestört den Weg zum Detektor findet, ist die Dicke und die Dichte der Gewebsschicht, durch die es sich bewegen muss. Während z.B. Lungengewebe den Weg der Photonen nur geringfügig beeinflusst, kommen nach Durchquerung von 7 cm weichteildichten Gewebes nur etwa die Hälfte der Photonen am Detektor

Einleitung

an (51). Unter Verwendung der Dichteinformation der CT wird daher eine Korrektur für diese Abschwächung (engl. *attenuation correction*, AC) vorgenommen.

Annihilationsereignisse können auch aus dem Zufall heraus gleichzeitig auftreten und somit fälschlicherweise derselben LOR zugeordnet werden. Umgekehrt kann bei sehr klein gewählten Koinzidenzzeiten auch ein Annihilationsereignis übersehen werden, wenn sich die Strecken der zugehörigen Photonen zum Detektor deutlich voneinander unterscheiden. Eine Korrektur kann unter Berücksichtigung der sog. *time-of-flight* (TOF) erfolgen (52). Weitere messphysikalische Korrekturverfahren stellen u.a. die *random correction* und *dead time correction* dar (50).

1.3.3. Bildrekonstruktion

Bei der Bildrekonstruktion erfolgt die Rückprojektion der korrigierten Rohdaten (Sinogramm) in die Form eines Tomogramms. Als Rekonstruktionsalgorithmus wird hierzu meist die *ordered subset expectation maximization* (OSEM) verwendet. Dabei wird in mehreren Iterationen die wahrscheinlichste Aktivitätsverteilung geschätzt (48,53). Mit zunehmender Zahl an Iterationsschritten erhöht sich die Bildinformation, allerdings nimmt auch das Bildrauschen zu. Das Rauschverhalten kann durch Verwendung der TOF-Information (s.o.) sowie der *point spread function* (PSF) verbessert werden (54). Die PSF beschreibt, wie ein ideales, punktförmiges Signal durch den Scanner dargestellt wird. Eine wichtige Kenngröße der PSF ist die *full width at half maximum* (FWHM), welche das räumliche Auflösungsvermögen maßgeblich bestimmt. Es konnte gezeigt werden, dass die Verwendung von TOF und PSF zu signifikanten Veränderungen quantitativer PET-Parameter, insbesondere zu einer Erhöhung des *standardized uptake value* (SUV, s.u.) führen kann (55,56).

Neben Rekonstruktionsalgorithmus und Korrekturverfahren sind die Schichtdicke und Schichtabstand weitere relevante Parameter der Rekonstruktion. Somit besteht hier insgesamt ein hohes Maß an Variation. Die z.T. sehr ausgeprägten Auswirkungen unterschiedlicher Bildrekonstruktion auf PET-Parameter sind bereits detailliert in der Literatur beschrieben (44,57–61). Die rekonstruierten Bilder können abschließend durch Verwendung von Filtern weiter prozessiert werden. Sogenanntes Glätten kann beispielsweise ein Bildrauschen verringern, birgt allerdings die Gefahr, dass kleine Details übersehen werden.

1.4. PET Parameter

Neben der bloßen Lokalisierung von Läsionen in der PET können auch quantitative Bestimmungen vorgenommen werden um bestimmte Eigenschaften (engl. *features*) der Läsion hervorzuheben. Die Disziplin, die sich mit der gezielten Suche nach relevanten Form- und Texturfeatures in der Bildgebung befasst, wird auch als Radiomics bezeichnet. Neben Dignitätsbeurteilung, Staging und Verlaufskontrolle, sollen sie dabei unterstützen prognostische und prädiktive Aussagen im Hinblick auf Therapieansprechen, Rezidivrisiko und Mortalität zu treffen.

Intuitiv stellt sich in der PET zunächst die Frage nach der Aktivität und dem Volumen einer identifizierten Läsion. Diese Eigenschaften werden durch die Parameter SUV und MTV abgebildet. Die Asphärizität (ASP) ist demgegenüber ein relativ neuer Parameter und ist ein Maß für die Formunregelmäßigkeit einer Läsion. Zahlreiche Faktoren können Einfluss auf die quantitative Bestimmung dieser Parameter nehmen.

1.4.1. MTV (Metabolisches Tumervolumen)

Das MTV beschreibt das Radionuklid-anreichernde Volumen einer Läsion, meist ausgedrückt in ml bzw. cm^3 . In einer 2016 veröffentlichten Metaanalyse von Liu et al. konnte der Nutzen des MTV als prognostischer Faktor für *disease-free survival* (DFS) und *overall survival* (OS) beim NSCLC gezeigt werden, wobei höhere Werte mit einem kürzeren Überleben assoziiert waren (62). Neben der prognostischen Bedeutung gibt es auch Evidenz für einen prädiktiven Nutzen des MTV, etwa im Hinblick auf das Ansprechen auf Chemo- und Immuntherapien (63,64).

Für die volumetrische Charakterisierung einer Läsion muss diese im PET-Bilddatensatz räumlich definiert werden (Segmentierung), d.h. für jedes Voxel muss entschieden werden, ob es zur Läsion gehört oder nicht. Die Gesamtheit aller eingeschlossenen Voxel stellt die sogenannte *region of interest* (ROI), das Abbild der Läsion, dar. Eine manuelle Segmentierung von Voxel zu Voxel ist zwar möglich, jedoch sehr zeitaufwendig und anwenderabhängig, daher existieren zahlreiche (semi-)automatische Segmentierungsalgorithmen für die PET (61). Das MTV wird maßgeblich durch die verwendete Segmentierungsmethode beeinflusst. Eine einheitliche Konvention für die Läsionsabgrenzung in der PET ist bisher nicht flächendeckend etabliert.

Manuelle Segmentierung des MTV

Die manuelle Segmentierung erfordert die voxelweise Bestimmung der ROI durch den Untersucher, ist sehr aufwendig und stark untersucherabhängig (65). Neben Faktoren der Bildqualität (räumliche Auflösung, Rauschen) kann auch die Darstellung der PET-Daten (Fensterung, Farbcodierung) bei ein und demselben Untersucher zu einer unterschiedlichen Identifikation der Tumorgrenzen führen (66). Um eine bessere Reproduzierbarkeit zu gewährleisten, sollten daher standardisierte Protokolle verwendet werden, die im Konsens der jeweiligen Fachgesellschaft entwickelt wurden (67,68). Deren Verwendung muss allerdings regelmäßig trainiert werden (69).

Schwellenwert-basierte Segmentierung des MTV

Bei Schwellenwert-basierten (engl. *Threshold based*) Verfahren wird eine ROI anhand eines Aktivitätsschwellenwerts bestimmt, welcher wiederum eine Isokontur als Rand der Läsion definiert. Es existieren verschiedene Variationen, wobei der konstant festgelegte Schwellenwert (engl. *fixed threshold*, FT) die einfachste Methode darstellt. Komplexere Algorithmen passen den Schwellenwert unter Berücksichtigung umgebender Voxel an (*adaptive thresholding*, AT). Beim *fixed thresholding* wird die ROI auf Basis eines festgelegten Schwellenwerts definiert. Zu hohe FT-Werte haben eine zu enge Wahl der Tumorgrenzen zur Folge (weniger intensive, periphere Tumoranteile werden ignoriert), zu kleine FT Werte können hingegen zu einer fälschlichen Miterfassung von nicht-läsionalen Strukturen führen. Problematisch ist auch das Vorliegen starker Heterogenität des Tumors. Der FT-Wert kann als Absolutwert (in SUV) angegeben werden (FT_{abs}), oder relativ zu einer bestimmten, tumorbezogenen Referenz; häufig dem sog. SUV_{max} , also dem SUV des „heißesten“ Voxel (engl. *hottest voxel*). Beim **absoluten FT** liegen typische Schwellenwerte im Falle des NSCLC bei ca. 2,5 - 6,0 SUV, wobei der optimale Wert von verschiedenen technischen Faktoren sowie von FDG-Avidität und Lage des Tumors (peripher vs. zentral) abhängt (62). Auch ist die klinische Fragestellung, also die Konsequenz einer Über- bzw. Unterschätzung des MTV miteinzubeziehen (70–72). **Relatives FT** arbeitet beim NSCLC häufig mit Werten zwischen 20 - 50% des SUV_{max} (62,73–76). Relatives FT hat den Vorteil, dass es weniger von der Genauigkeit der quantitativen SUV-Bestimmung abhängt und somit in diesem Hinblick weniger technischen Einflüssen unterliegt als das absolute FT. Auf der anderen Seite kommt es gerade bei großen, FDG-aiden Tumoren mit entsprechend hohem SUV_{max} oft zu unvollständiger Tumorsegmentierung, sodass einige Autoren eine Abstufung empfehlen, mit Werten von 42% des SUV_{max} für Tumoren < 3 cm und 15% des SUV_{max} für Tumoren > 5 cm (75).

Einleitung

Im Weiteren kann zur Berechnung eines FT-Wertes die Hintergrundintensität mit einbezogen werden, welche wiederum mithilfe einer separaten ROI in einem geeigneten läsionsfreien Bildabschnitt gemessen werden kann. In der Regel ergibt sich der Schwellenwert aus einer Formel der Form:

$$I_{\text{Threshold}} = \alpha \cdot (I_{\text{Referenz}} - I_{\text{Background}}) + I_{\text{Background}}$$

wobei α ein i.d.R. empirisch ermittelter Wert (z.B. 40%) darstellt, I_{Referenz} die höchste Aktivität (SUV_{max}) oder die über die Läsion gemittelte Aktivität (SUV_{mean}), und $I_{\text{Background}}$ aus einer geeigneten Hintergrundregion ermittelt wird. Ein problematischer Aspekt der FT-Verfahren ist, dass gerade bei sehr heterogener und asymmetrischer Verteilung des MTV ein fixer Schwellenwert nicht immer zu einer idealen Abgrenzung der Läsion führt. Abb. 1 illustriert dies exemplarisch: Während bei symmetrischer Läsion sowohl Threshold 1 als auch Threshold 2 zu einem relativ ähnlichen MTV führen, ist dies im Falle der asymmetrischen Läsion nicht der Fall. **Adaptives Thresholding** nimmt sich dieses Problems an, in dem es den verwendeten Schwellenwert in Abhängigkeit von der konkreten Umgebung des betrachteten Voxels anpasst (77).

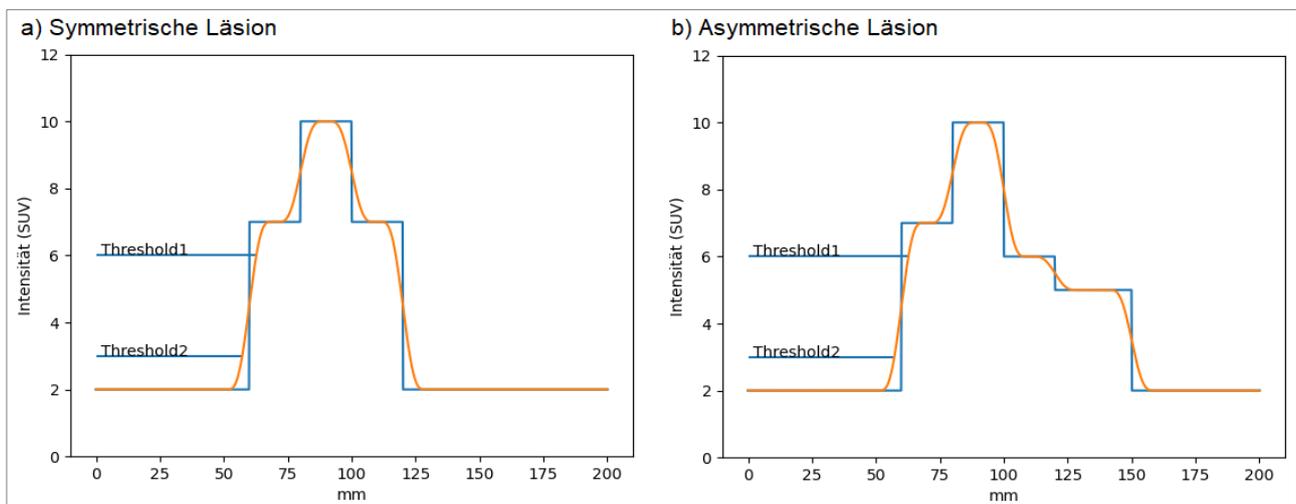


Abbildung 1: Vergleich zweier verschiedener Schwellenwerte (Threshold 1 und 2) bei homogener (a) und heterogener (b) Läsion. In (b) kommt es je nach Schwellenwert zu deutlichen Unterschieden im abgegrenzten Tumolvolumen.

Weitere Verfahren zur Segmentierung des MTV

Sowohl die manuelle als auch die Thresholding-basierten Methoden können allgemeiner als voxelbasierte Segmentierungsverfahren aufgefasst werden, da hier primär direkt für jedes Voxel entschieden wird, ob es zur ROI gehört oder nicht. Daneben existieren zahlreiche weitere Algorithmen zur Segmentierung, die auf ganz anderen Ansätzen fußen. So versuchen **kantendetektierende** Verfahren die Konturen einer Läsion zu identifizieren (sog. *edge detection*), etwa durch Berechnung von Intensitätsgradienten oder mittels statistischer Verfahren wie z.B. Markov-Modellen. Obgleich methodisch sehr spitzfindig und kreativ haben sich viele dieser Methoden, im Gegensatz zur manuellen oder Thresholding-basierten Segmentierung, im klinischen Alltag aufgrund ihrer komplexeren Implementierung oder mangelnden Robustheit nicht bewährt (78). Sehr moderne Verfahren der Tumorsegmentierung nutzen sogenannte **convolutional neural networks (CNNs)**, welche sich konzeptionell am schichtweisen Aufbau und der Funktionsweise der Sehrinde von Säugetieren orientiert (79,80). Unter Einsatz großer Trainingsdatensätze und hohem rechnerischen Aufwand können CNNs darauf trainiert werden Läsionen zu erkennen und zu segmentieren. Limitierend ist dabei allerdings die begrenzte Verfügbarkeit ausreichend großer und adäquater Trainingsdatensätze. Zahlreiche unterschiedliche CNNs wurden bereits entwickelt und auf verschiedene radiologische Fragestellungen in verschiedenen Modalitäten (CT, MR, PET, Ultraschall) angewendet. Dabei zeigt die Literatur vielversprechende Ergebnisse, nicht nur für die Segmentierung von Lungentumoren, sondern auch u.a. für Tumoren von Hirn, Prostata und des Kopf-Hals-Bereichs (81–84). Zum aktuellen Zeitpunkt sind CNN-basierte Verfahren allerdings noch weitgehend auf den Forschungsbereich beschränkt (80). Anzumerken ist zudem, dass sich unterschiedliche CNNs in ihrer Architektur und Hyperparameter unterscheiden und damit prinzipiell auch zu unterschiedlichen Ergebnissen führen können (85).

Konsequenzen unterschiedlicher Segmentierungsalgorithmen

In der Literatur finden sich unzählig viele Publikationen zu verschiedenen, neuartigen Segmentierungsalgorithmen. Der Vorteil automatisierter Algorithmen ist die weitgehende Untersucherunabhängigkeit und die dadurch bessere Reproduzierbarkeit. Dennoch ist die visuelle Kontrolle der Plausibilität des Resultats durch einen PET-erfahrenen Untersucher essenziell. Sowohl die Tumoridentifikation als auch die Differenzierung zwischen Tumorbestandteilen und benachbartem, nicht-tumorösen Gewebe ist fehleranfällig, da auch anderweitig pathologisch verändertes Gewebe (z.B. bei Inflammation) aber auch gesundes Gewebe (z.B. Myokard, Leber)

Einleitung

aus physiologischen Gründen eine hohe Traceravidität aufweisen kann. Für die klinische Praxis wäre eine standardisierte Vorgehensweise bei der Segmentierung wünschenswert. In der Literatur finden sich zahlreiche Vergleiche der Performance verschiedener Algorithmen (74,86), wobei hier auch stets die Frage nach dem Goldstandard (engl. *ground truth*) zu stellen ist, anhand derer die Performance evaluiert werden soll. Diese können z.B. visuell durch einen oder mehrere Expert*innen, durch Verwendung der CT, oder anhand des pathologischen Präparats erfolgen. Auch hier ist bisher kein einheitliches Vorgehen etabliert (87). Vielmehr kann die geeignete *ground truth* abhängig von der klinischen Fragestellung sein. So ist z.B. im Kontext einer zu planenden Strahlentherapie beim NSCLC eine allein CT-basierte *ground truth* nicht zwangsläufig zielführend, da sie die atemabhängigen Bewegungen des Tumors nicht miterfasst. Hier ist die Integration der PET-basierten *ground truth* daher sinnvoll (88–90). Generell ist eine große Schwäche bildgebender Verfahren der systematische Mangel einer idealen *ground truth*, denn selbst die aufwendige, visuelle Segmentierung durch Expert*innen ist variabel.

Weitere Einflussfaktoren

Wie dargestellt, hängt das MTV maßgeblich vom verwendeten Segmentierungsalgorithmus ab, da dieser letztlich die Tumorränder definiert. Darüber hinaus können zahlreiche technische und klinische Faktoren zu einer Beeinflussung des segmentierten MTV führen. Grundsätzlich können alle Faktoren, die mit einer Änderung des Bildrauschens einhergehen, auch die Darstellung der Tumorränder beeinflussen. Dabei handelt es sich insbesondere um scanner- und rekonstruktionsspezifische Faktoren. So führte in einer Untersuchung von Huang et al. 2021 die kombinierte Verwendung von TOF und PSF zu einer deutlichen Verbesserung des Rauschverhaltens bei Lungenläsionen in der PET wobei sich das MTV um 24% reduzierte (56).

Neben technischen Faktoren spielen auch Übergewicht oder eine zu geringe Traceranreicherung im Zielgewebe, sei es durch einen hohen Blutzucker, Unregelmäßigkeiten bei der Radionuklid-Applikation oder eine unzureichende Scandauer eine wichtige Rolle für die Bildstatistik und können die Differenzierung der exakten Tumorgrenzen erschweren.

1.4.2. SUV (Standardized Uptake Value)

Die Signalintensität und damit die metabolische Aktivität wird durch den SUV (engl. *standardized uptake value*) gemessen. Sie berechnet sich aus dem Verhältnis

$$SUV = \frac{\text{Aktivität pro Volumen [gemessen]}}{\text{Aktivität der injizierten Dosis pro kgKG [berechnet]}}$$

zu einem gegebenen Zeitpunkt und stellt eine einfache Normierung der PET-Signalintensität dar, mit dem Ziel eine Vergleichbarkeit zwischen verschiedenen Untersuchungen und Patient*innen zu ermöglichen. Wichtige Messgrößen sind die maximale Aktivität in der Läsion (SUV_{max}), also der SUV des sog. *hottest voxel*, sowie die über die gesamte Läsion gemittelte, durchschnittliche Aktivität (SUV_{mean}). Insgesamt ist die prognostische Bedeutung des SUV_{max} bereits in verschiedenen Metaanalysen gezeigt worden (62,91). Seine Dokumentation wird in der aktuellen S3-Leitlinie zur Diagnostik des Lungenkarzinoms explizit empfohlen (4).

Einflussfaktoren

Die berechneten SUVs sind von zahlreichen Faktoren abhängig. Hierzu zählen scannerspezifische Faktoren (Detektoren und die erreichte räumliche Auflösung) und Bildrekonstruktionsverfahren. Insbesondere der Partialvolumeneffekt spielt bei der relativ begrenzten Auflösung von heutigen PET-Scannern (ca. 4-5mm) eine große Rolle (60). Insgesamt kann es allein dadurch zu Schwankungen von bis zu 30% des SUV kommen (92). Die Verwendung von TOF und PSF führte in der zuvor zitierten Untersuchung von Huang et. al. bei Lungenläsionen zu einer Erhöhung des SUV_{max} um 23% (56).

Neben technischen Faktoren können auch klinische Einflüsse vor und während der Bildakquisition die adäquate Berechnung des SUV stören. Hier seien z.B. Unregelmäßigkeiten bei der Applikation des Tracers, Verzögerungen bei der Messung, oder ausgeprägte Bewegungsartefakte erwähnt, wodurch es zu Diskrepanzen zwischen berechnetem SUV und tatsächlicher lokaler Aktivität kommt. Auch eine diabetische Stoffwechsellage oder Nahrungsaufnahme vor der Untersuchung kann physiologisch bedingt zu einer geringeren Traceraufnahme im Zielgewebe führen und zu einer ungenauen Charakterisierung von Lungenherden in der PET führen (93,94). Für Lungenläsionen scheinen Diabetes und Übergewicht die gemessenen SUVs jedoch nicht signifikant zu beeinflussen (95,96).

Einleitung

Da die Tumorränder keine Rolle für die Berechnung des SUV_{max} , spielen, ist der verwendete Algorithmus zur Segmentierung des MTV i.d.R. unerheblich, solange er nicht zu einem Ausschluss des *hottest voxel* führt. Für andere SUV-basierte Parameter wie der über das gesamte MTV gemittelte SUV (SUV_{mean}) oder die *total lesion glycolysis* (TLG) als Produkt von SUV_{mean} und MTV ist dagegen naturgemäß ein direkter Einfluss der Tumorsegmentierung anzunehmen, wobei im Falle zu eng gewählter Tumorgrenzen eine Überschätzung des wahren SUV_{mean} zu erwarten ist und umgekehrt (97).

1.4.3. ASP (Asphärizität)

Die ASP ist ein Formparameter, der aus einem segmentierten Volumen abgeleitet werden kann. Im Kontext der PET-Bildgebung wird damit die Unregelmäßigkeit der Oberfläche des MTV charakterisiert. Die ASP entspricht hierbei der relativen Abweichung der Oberfläche des MTV in Relation zur Oberfläche einer isovolumetrischen Kugel und kann somit vereinfacht als Maß für die „Entrundung“ verstanden werden (98). Der Parameter inkorporiert dadurch morphologische Charakteristika, die im Allgemeinen mit der Malignität eines Tumors assoziiert sind, wie z.B. asymmetrisches Wachstum, Spikulae und zentrale Nekrosen (4,7).

Definition und Beispiele

Das mathematische Konzept der *Sphärizität* Ψ ist sehr alt und dokumentierte naturwissenschaftliche Anwendungen gehen auf Wadell et al. aus dem Jahr 1935 zurück, der diesen Parameter zur Formbeschreibung von Kristallen verwendete (99).

Die charakteristische Eigenschaft einer Kugel ist, dass ihre Form ein optimales (kleinstes) Verhältnis von Oberfläche zu Volumen darstellt. Jegliche Verformung einer Kugel (bei konstant gehaltenem Volumen) wird zu einer Vergrößerung der Oberfläche (A) führen. Die ASP ist daher wie folgt definiert:

$$ASP = \frac{A - A_{Kugel}}{A_{Kugel}} = \frac{A}{A_{Kugel}} - 1 = \left(\frac{1}{36\pi} \cdot \frac{A^3}{V^2} \right)^{\frac{1}{3}} - 1$$

Eine Asphärizität von 0 entspricht einer idealen Kugel. Sie steigt mit zunehmender Größenabweichung der Oberfläche von der „idealen“ Oberfläche.

Einleitung

Abbildung 2 zeigt 6 verschiedene Formen und die dazugehörigen ASP-Werte. Die Beispiele sind an mögliche morphologischer Erscheinungsformen von Lungentumoren in der Bildgebung angelehnt. So können Tumoren z.B. als rundliche Knoten (a), spikuliert (c), zentral eingeschmolzen (f), der Pleura aufsitzend (e), oder mit Satellitenherden (d) auftreten.

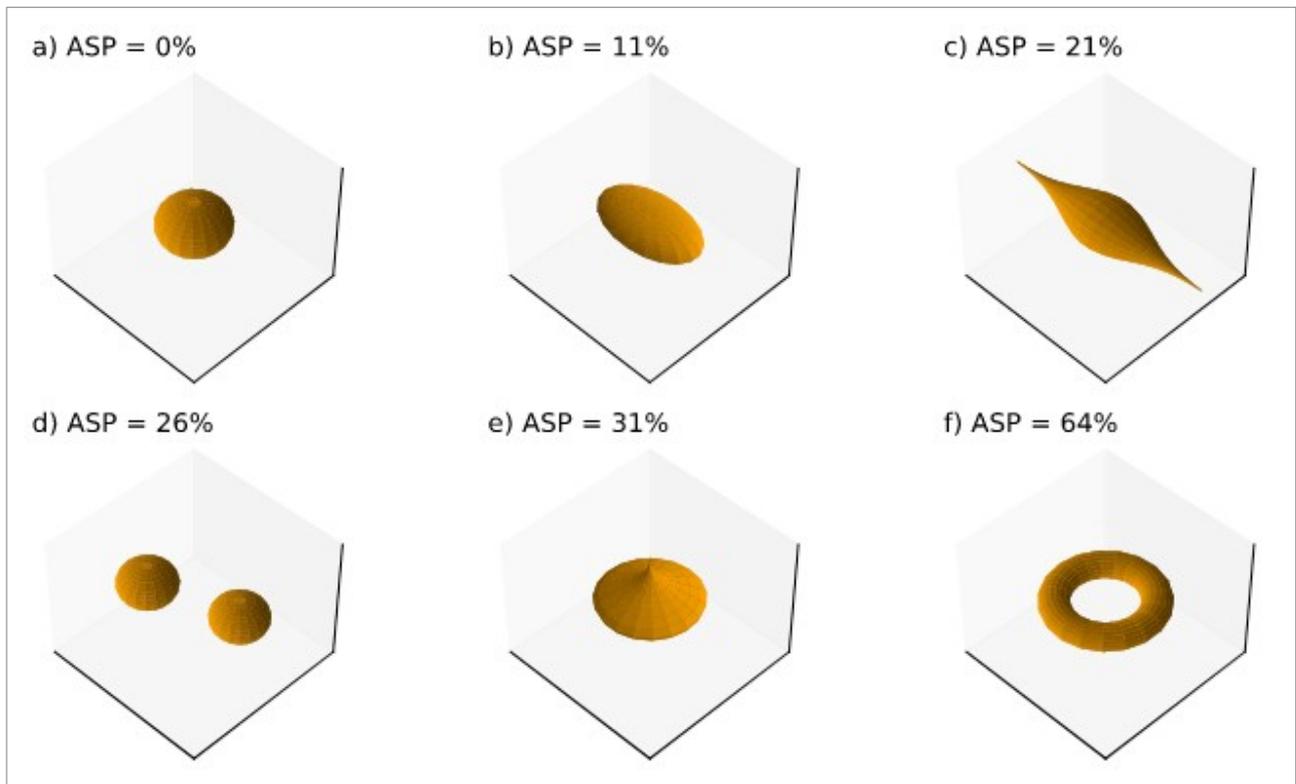


Abbildung 2: Verschiedene Körper und zugehörige ASP-Werte. Sämtliche Körper mit Ausnahme von (d) haben das gleiche Volumen. In Beispiel (d) ist das Volumen gleichmäßig auf zwei Kugeln verteilt wurden. Berechnet wurde hier die Gesamt-ASP des Ensembles aus beiden Kugeln. Beispiel (e) ist eine konstruierte Form, die nach unten hin einer planen Fläche aufsitzt und nach oben hin spitz ausläuft.

Bisherige Datenlage zur prognostischen Bedeutung

Im Jahr 2014 konnten Apostolova et al. in einer kleineren retrospektiven Untersuchung von 60 Patient*innen mit NSCLC in den Stadien UICC I - III eine signifikante prognostische Bedeutung der ASP im Hinblick auf Progress und Sterblichkeit zeigen. Dabei waren die beobachteten Effektgrößen deutlich stärker als die anderer untersuchter PET-Parameter (SUV_{max} , MTV) (100). Vielversprechende Ergebnisse konnten in den Folgejahren auch für Tumoren des Kopf-Hals-Bereichs gezeigt werden (98,101).

Weitere Untersuchungen demonstrierten einen Zusammenhang zwischen der ASP und dem histopathologischen und molekularen Phänotyp von Tumoren:

In einer retrospektiven Studie mit 83 Patient*innen mit NSCLC zeigte sich eine Korrelation der ASP (aber auch von SUV_{max} und MTV) mit der Expression des Proteins Ki-67, das als wichtiger tumorbiologischer Marker für Proliferation bekannt ist. Interessanterweise wies jedoch unter diesen 3 Parametern nur die ASP eine signifikante Korrelation mit dem Angiogenese-Marker VEGF, sowie dem M-Stadium (Metastasierung) auf (102). In einer Untersuchung am Prostatakarzinom war die ASP im Vergleich zu anderen PET-Parametern am stärksten mit dem Gleason-Score korreliert (103). Insgesamt scheint damit die Vermutung naheliegend, dass die ASP ein PET-Parameter sein könnte, der die Übersetzung der mikroskopischen Tumorbiologie hin zum makroskopischen Tumorkorrelat bisher am deutlichsten verkörpert. Belastbare Daten diesbezüglich fehlen allerdings.

2. Zielsetzung

Bisherige kleinere Studien ergaben Hinweise auf ein mögliches Potential der Asphärizität (ASP) des in der [¹⁸F]FDG-PET dargestellten metabolischen Tumervolumens des Primärtumors als prognostischer Marker für das nicht-kleinzellige Lungenkarzinom (NSCLC) (100).

Hauptziel der folgenden Arbeit ist die Validierung der ASP als unabhängiger prognostischer Marker für Patient*innen mit NSCLC in den Stadien UICC I bis IIIA mit primär kurativer Behandlungsintention. Anhand einer retrospektiven Untersuchung an einer größeren Studienpopulation soll mithilfe der multivariablen Cox Regression die Frage beantwortet werden, ob die ASP unabhängig von etablierten prognostischen Faktoren einen prognostischen Wert für das krankheitsfreie Überleben (DFS) und Mortalität (OS) in diesem Patient*innenkollektiv aufweist.

Zusätzlich soll in dieser Arbeit die Robustheit der ASP untersucht werden, da diese eine Voraussetzung für ihren zentrumsübergreifenden Einsatz darstellt. Die Frage nach dem Einfluss technischer Parameter des PET-Systems ist bereits zuvor untersucht worden (57). In bisherigen Arbeiten zur ASP bei NSCLC wurde dieselbe Segmentierungsmethode verwendet (iAT). In dieser Arbeit soll nun die Frage beantwortet werden, ob die ASP zwischen verschiedenen Segmentierungsalgorithmen vergleichbar ist und ob der prognostische Wert der ASP auch bei Verwendung anderer, üblicher Segmentierungsalgorithmen grundsätzlich erhalten bleibt.

3. Methoden

3.1. Patient*inneneinschluss und Follow-up

Betrachtet wurden Patient*innen mit gesichertem NSCLC in den Stadien UICC I - IIIA, die im Zeitraum von 2009 bis 2017 an der Charité-Universitätsmedizin Berlin in Behandlung waren und bei denen eine kurative Therapie (einschließlich Operation) vorgesehen war. Von 661 erhielten initial 396 Patient*innen eine prätherapeutische [¹⁸F]FDG-PET/CT in der Klinik für Nuklearmedizin der Charité. Von diesen Patient*innen erfüllten 361 die weiteren Einschlusskriterien (s. Abb. 3). Patient*innen mit prognostisch relevanten Zweittumoren wurden ausgeschlossen, um diese Faktoren als störende Einflüsse bei der Überlebensanalyse zu vermeiden. Das Follow-up erfolgte initial in Zusammenschau aller relevanten Daten des Krankenhausinformationssystem der Charité (SAP). Da sich allerdings ein großer Teil der Patient*innen in ambulanter Weiterbetreuung außerhalb der Charité befand, oder inzwischen gar keine weitere Nachbeobachtung benötigte, wurden die jeweiligen Hausärzte bzw. ambulant tätigen Onkologen postalisch kontaktiert und zur Ausfüllung eines einheitlichen Follow-up-Formblatts gebeten. In Zweifelsfällen wurde telefonischer Kontakt mit Patient*innen bzw. Angehörigen aufgenommen.

In jedem Fall lag die individuelle Einwilligung der Patient*innen zur wissenschaftlichen Datenerhebung und -auswertung sowie zur Kontaktaufnahme mit den Behandler*innen vor. Die wissenschaftliche Untersuchung erfolgte in Übereinstimmung mit der Ethikkommission der Charité (Votum EA1/373/16) und unter Wahrung des Datenschutzes.

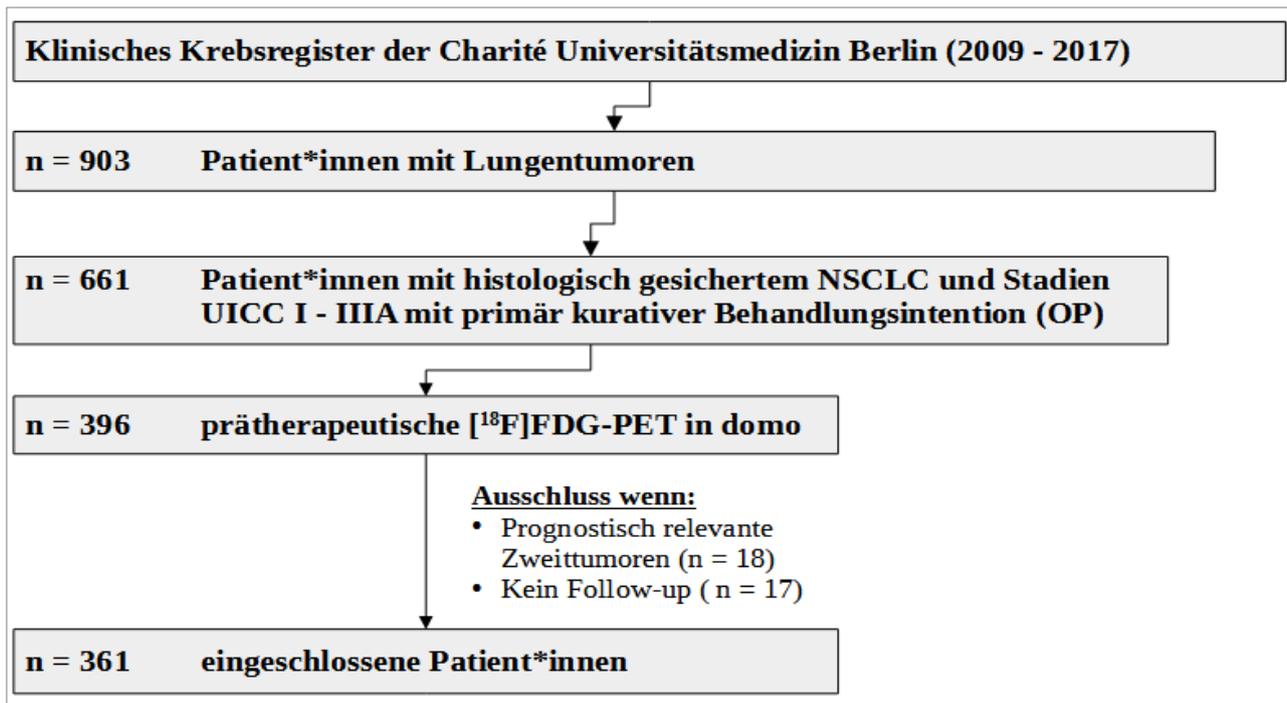


Abbildung 3: *Flowchart Patient*inneneinschluss.*

3.2. Akquisition der Bilddaten und Bildrekonstruktion

Das Vorliegen einer initialen $[^{18}\text{F}]$ FDG-PET war ein notwendiges Kriterium für den Einschluss der Patient*innen in die Untersuchung. Die PET wurde in der Klinik für Nuklearmedizin, allesamt mit dem gleichen Scanner (Philips Gemini TF 16, Amsterdam, Niederlande) vorgenommen. Alle Patient*innen waren für mindestens 6h vor der Injektion des $[^{18}\text{F}]$ FDG nüchtern. Ein Blutzucker von ≤ 190 mg/dl wurde vor der Injektion mittels Blutzuckermessgerät sichergestellt.

Die Aufnahme umfasste den Bereich von Schädelbasis bis proximales Femur. Tabelle 2 listet die verwendeten technischen Parameter auf. Die dosimetrischen Daten der Studienpopulation sind in Tabelle 3 zusammengefasst. Die Rekonstruktion der PET-Rohdaten erfolgte durch OSEM (ordered subset expectation maximization), wobei ab Juni 2013 die zusätzliche TOF Analyse integriert wurde. Die rekonstruierte Schichtdicke der PET-Daten betrug 4mm (Matrix 144×144 , Voxelgröße $4 \times 4 \times 4 \text{mm}^3$).

Methoden

Tabelle 2: PET-CT Akquisition und Rekonstruktion. Technische Parameter des verwendeten PET-CT Systems (Gemini TF 16; Philips, Amsterdam (NL) mit TOF Technologie (Philips Astonish TF))

Parameter	Einstellung	
PET	Modus:	3D, Emission
	Scanzeit:	90-180s pro Bettposition
	Überlappung der Bettpositionen:	53,3%
	Rekonstruktion:	OSEM (ab 06/2013 zudem mit TOF Analyse: BLOB-OS-TF, 3 Iterationen, 33 Subsets) 4 mm Schichtdicke, Matrix 144x144
CT	<u>2009 - 2015</u>	<u>2016 - 2017</u>
	Kontrastmittel (KM)-gestützte CT	Low-Dose CT
	Automatische Modulation des Röhrenstroms	Automatische Modulation des Röhrenstroms
	max. mAs-Produkt: 200mAs	max. mAs-Produkt: 50mAs
	Röhrenspannung: 120 kV	Röhrenspannung: 120 kV
	KM-Bolusrate: 3mL/s, Verzögerungszeit: 80s	
	Schichtdicke 5mm	Schichtdicke 5mm

Tabelle 3: Dosimetrie

	Median	IQR	Range
Gewicht (kg)	75	65 - 85	42 - 130
i.v. Aktivität (MBq)	300	254- 343	150 – 417
Uptakezeit (min)	68	58 – 86	21 – 160
Zerfallskorrigierte Aktivität (MBq/kg)	2,5	1,9 – 3,1	0,9 – 5,6

3.3. Bildauswertung und Segmentierung

Ziel der Bildauswertung war die Bestimmung von ASP, MTV und SUV_{max} aus den segmentierten Primärtumoren für jede Patientin und jeden Patienten. Dazu wurde der im DICOM-Format vorliegende PET-Datensatz mithilfe der Software ROVER (Version 3.0.34, ABX advanced biochemical compounds GmbH, Radeberg, Germany) eingelesen und ausgewertet. Die Lokalisierung des Primärtumors wurde durch einen PET-erfahrenen Nuklearmediziner nach Facharztstandard kontrolliert. Es wurde eine Maske um den Tumor gelegt, die diesen vollständig umschließt, und dann mithilfe verschiedener Algorithmen die Tumor-ROI segmentiert.

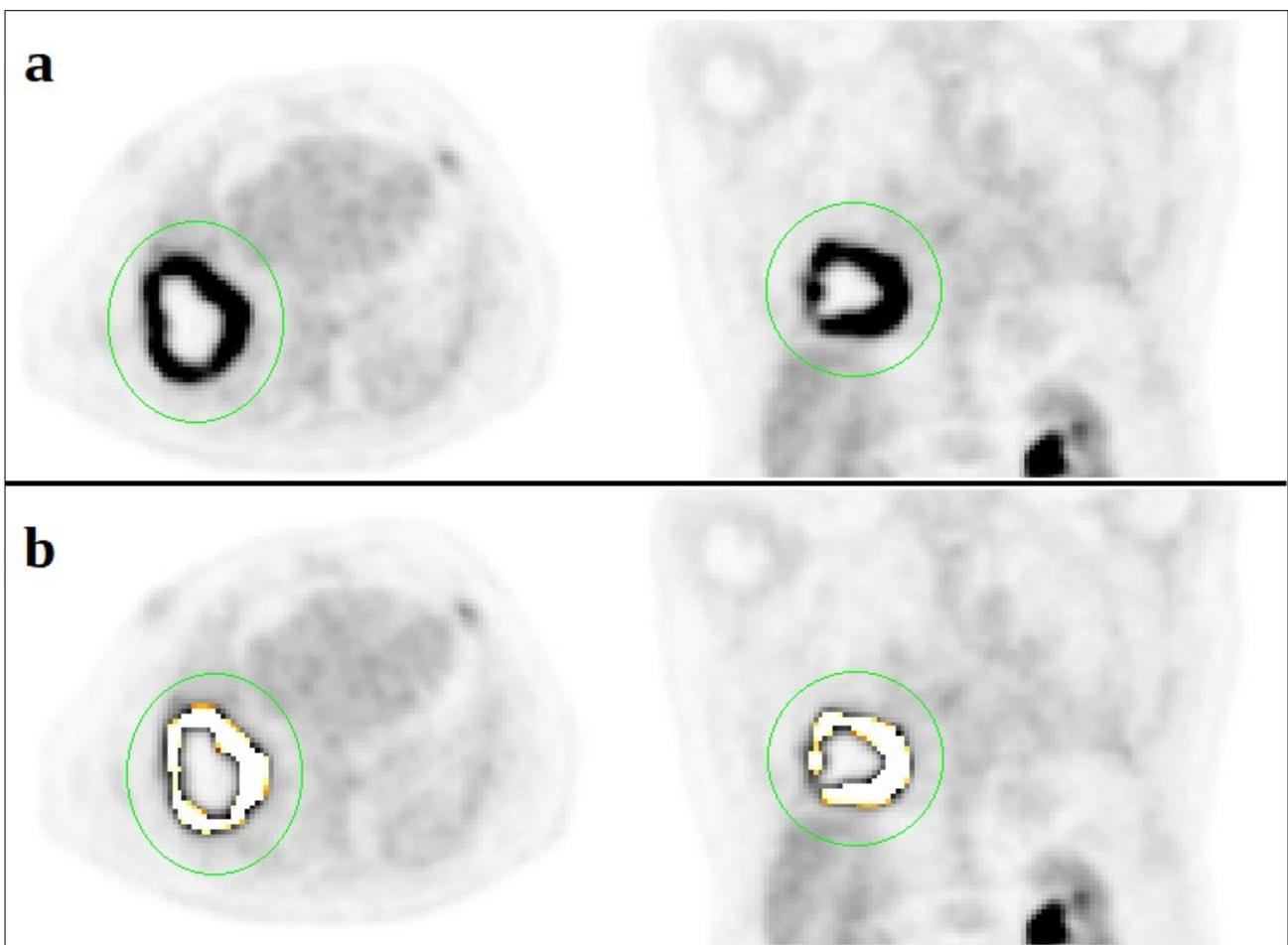


Abbildung 4: Tumorsegmentierung in ROVER am Beispiel eines zentral nekrotischen Lungenkarzinoms im rechten Unterlappen. PET in axialer (links) und koronarer (rechts) Schichtung. Zunächst erfolgt die Platzierung der 3D-Maske (a; grün). Nach Auswahl des Segmentierungsalgorithmus und ggf. weiterer Parameter wird die resultierende ROI ausgegeben (goldene Region).

3.3.1. Iterative Adaptive Thresholding (iAT)

In der vorliegenden Arbeit wird zur Tumorsegmentierung ein iterativer, Hintergrund-adaptiver Thresholding Algorithmus (iAT) nach Hofheinz (77) verwendet, mit dem auch in den vorangegangenen Publikationen zur ASP gearbeitet wurde (98,100–102). Er ermöglicht im Vergleich zum Fixed Thresholding eine robustere Abgrenzung bei schwacher Traceraufnahme und höhergradiger Heterogenität des Tumors (77).

Der Untersucher definiert eine grobe Maske um die betrachtete Läsion. In jedem Iterationsschritt finden nun automatisiert die folgenden Prozesse statt, die zur besseren Nachvollziehbarkeit in Abb. 5 veranschaulicht sind:

0. Der heißeste Voxel in der Maske bildet zu Beginn die erste Schätzung der Läsions-ROI.
1. Um die aktuell geschätzte Läsions-ROI (rot) wird eine dreidimensionale äußere Hülle (blaue/grüne Bereiche) mit einem festen Abstand zu dieser gelegt.
2. Es folgt eine Iteration über alle Voxel innerhalb der eingehüllten Region.
 - 2.a. Definition der Voxelumgebung: Um jedes Voxel wird ein kugelförmiger Raum mit festem Radius definiert.
 - 2.b. Ermittlung der voxelspezifischen Hintergrund-Intensität (I_{Bg}): Der Bereich, wo sich Voxelumgebung und die Hülle schneiden (grün) definiert den Grenzbereich (in Abb. 5 als *background region* bezeichnet). I_{Bg} ergibt sich aus der gemittelten Intensität aller Voxel im Grenzbereich. Sollte der Grenzbereich leer sein (d.h. die Voxelumgebung liegt vollständig im Inneren der eingehüllten Region), so wird für I_{Bg} stellvertretend die Durchschnittsintensität der gesamten Hülle (blau) verwendet.
 - 2.c. Definition der voxelspezifischen Referenzintensität (I_{ref}): Diese entspricht dem maximalen Intensitätswert im Schnittbereich von Voxelumgebung und geschätzter Läsions-ROI.
 - 2.d. Mithilfe der Formel:

$$I_{Threshold} = T \cdot (I_{Ref} - I_{Bg}) + I_{Bg}$$

wird nun ein voxelspezifischer Schwellenwert $I_{Threshold}$ ermittelt. Dieser wird zur Segmentierung innerhalb der Voxelumgebung angewendet und die im vorherigen

Methoden

Iterationsschritt ermittelte ROI entsprechend aktualisiert. Der Prozess beginnt nun wieder bei 1.

3. Mit zunehmender Zahl an Iterationen werden die Änderungen der ROI-Schätzung immer geringer, bis schließlich keine Änderung mehr stattfindet („Konvergenz“) und die finale Schätzung erreicht ist.

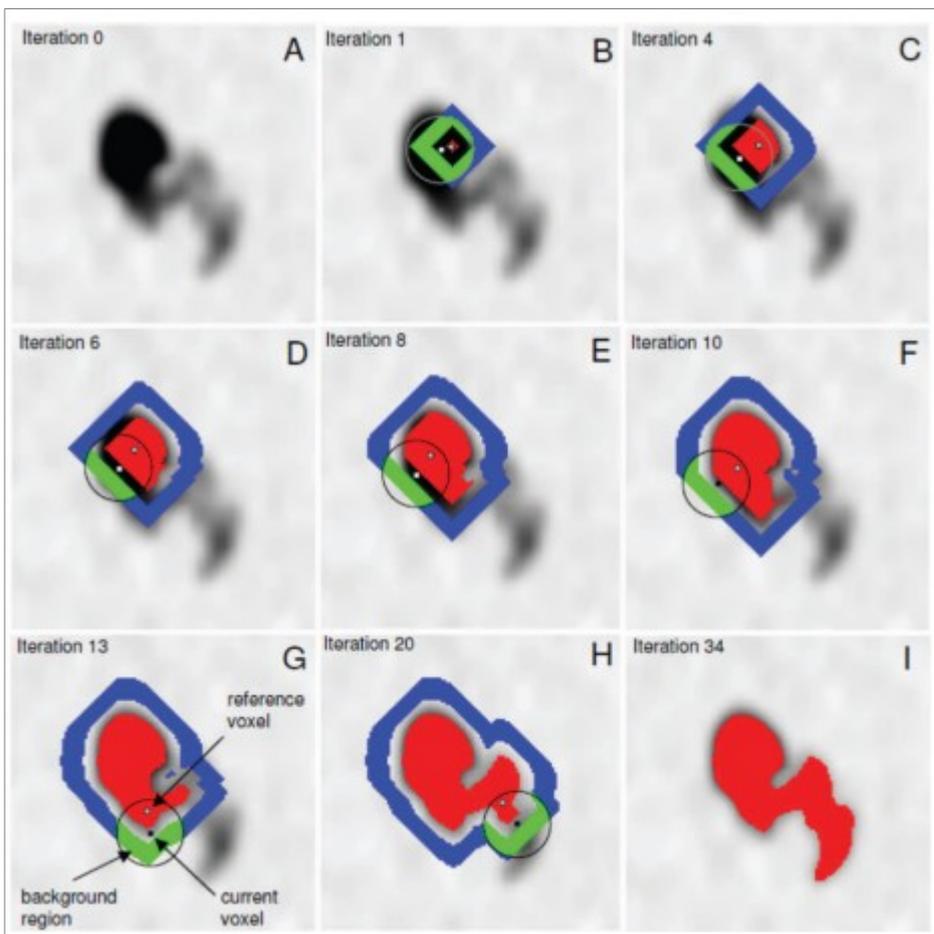


Abbildung 5: **Segmentierungsalgorithmus nach Hofheinz et. al.** Mit freundlicher Genehmigung von John Wiley and Sons (77)

3.3.2. Fixed Thresholding

Sämtliche Tumore wurden zusätzlich mithilfe verschiedener Fixed Thresholding Methoden und einer visuell gesteuerten Thresholding Methode segmentiert. Tabelle 4 gibt einen Überblick über die insgesamt 17 verschiedenen Methoden.

Tabelle 4: In dieser Arbeit verwendete Thresholding Methoden

Kategorie	Methoden
Adaptiv (AT)	iAT (iteratives adaptives Thresholding)
Fix (FT)	absolut absolute SW in SUV: 2,3 , 2,5 , 2,7 , 3,0 , 3,3 , 4,0
	relativ relative SW in %SUV _{max} : 20, 25, 30, 35, 40, 42, 45, 50, 60
	visuell „visuell optimierter SW“ (durch Ausprobieren ermittelter FT-SW, der in den Augen des Autors zur besten Erfassung des Tumors führt)

SW - Schwellenwerte

Für jede Segmentierung erfolgte eine visuelle Prüfung der ROI um völlig unplausible Segmentierungen zu identifizieren. War dies der Fall und eine manuelle Korrektur nicht ohne weiteres bzw. nicht in reproduzierbarer Weise möglich, wurde dies entsprechend als Versagen der Methode gewertet.

3.3.3. Regeln der Segmentierung

Für eine standardisierte Vorgehensweise bei der Bestimmung der ROI des Primärtumors mithilfe des iterativen adaptiven Thresholding (iAT) wurden folgende Regeln aufgestellt:

1. Alle Bestandteile des Primärtumors werden als eine zusammenhängende ROI aufgefasst. Dies gilt also auch bei Vorliegen mehrerer ipsilateraler Satelliten-Läsionen.
2. Kontralaterale Läsionen oder Fernmetastasen werden nicht in die ROI des Primärtumors einbezogen (da diese nicht Teil des T-Deskriptors sind).
3. FDG-avide Lymphknoten werden möglichst von der ROI ausgeschlossen. Ist eine Differenzierung nicht sinnvoll möglich, etwa bei Kontinuität von Primärtumor mit dem Lymphknoten, wird dieser in die ROI miteingeschlossen.
4. FDG-avides, nicht-tumoröses Gewebe in der Umgebung des Tumors (z.B. postobstruktive Pneumonie) wird ebenfalls nicht in die ROI eingeschlossen werden, sofern dies möglich ist.

Methoden

5. Ist eine Differenzierung zu umgebenden, FDG-aviden Strukturen wie z.B. Leber, Milz und Zwerchfell nicht möglich, wird die ROI manuell bearbeitet, allerdings nur dann, wenn tatsächlich alle anderen Methoden (z.B. Anpassung der Maske) versagen.
6. Bei Versagen des iAT wird auf das Fixed Thresholding zurückgegriffen, wobei der Schwellenwert visuell optimiert wird.
7. Jede Segmentierung ist visuell auf grundsätzliche Plausibilität zu überprüfen.

3.4. Statistische Analyse

Die statistische Analyse erfolgte in Python (Version 3.8). Für parametrische und nicht-parametrische Tests wurde die *scipy*-Bibliothek (Version 1.6.0) verwendet. Bei Annahme nicht normalverteilter Daten (Shapiro-Wilk-Test) wurden Parameter durch Median, IQR und Spannweite (Range) beschrieben. Auf Unterschiede hinsichtlich der Verteilung zwischen zwei oder mehreren Gruppen wurde mithilfe des Mann-Whitney-U-Tests bzw. des Kruskal-Wallis-Tests geprüft.

Überlebenszeitanalysen erfolgten unter Verwendung der *lifelines*-Bibliothek (Version 0.25.9). Zur Berechnung medianer Überlebenszeiten via der Kaplan-Meier-Methode oder des medianen Follow-ups via der umgekehrten Kaplan-Meier-Methode (104) wurde dabei das *KaplanMeierFitter*-Modul verwendet. Cox-Regressionsanalysen erfolgten mithilfe des *CoxPHFitter*-Moduls. Die Prüfung der Cox Proportional Hazards Assumption erfolgte dabei auf Grundlage der Schoenfeld-Residuen (Prüfung derer Zeitunabhängigkeit) via der *check_assumptions* Funktion im selbigen Modul. Berichtet wurden das Hazard-Ratio (HR) sowie die 95%-Konfidenzintervalle (95% CI) und die p-Werte der Einzelparameter im Wald-Test.

Statistische Signifikanz wurde bei einem p-Wert $\leq 0,05$ angenommen.

3.4.1. Validierung der ASP als unabhängiger prognostischer Marker

Es erfolgte eine uni- und multivariable Cox Proportional Hazards Regressionsanalyse der ASP für das krankheitsfreie (DFS) und Gesamtüberleben (OS). Die ASP wurde auf Grundlage einer Segmentierung des MTV nach dem iAT-Algorithmus berechnet.

Variablenselektion

Neben den PET Parametern ASP, MTV und SUV_{max} wurden demographische (Alter, Geschlecht, ECOG-Status), tumorbezogene (UICC-Stadium, T/N-Deskriptor, Histologie) sowie behandlungsbezogene Variablen (Resektions- bzw, R-status, Chemotherapie, Bestrahlung) untersucht. Eine Übersicht aller Variablen ist in Tabelle 5 gezeigt.

Methoden

Tabelle 5: Kovariablen für die univariable Cox Proportional Hazards Regression

PET-Parameter	
ASP	kontinuierlich
MTV	kontinuierlich
SUV _{max}	kontinuierlich
Demographie	
Geschlecht	männlich vs. weiblich
Alter bei Diagnosestellung	kontinuierlich
ECOG-Status	≥2 vs. < 2
Tumorstadium	
TNM T-Deskriptor*	T2 vs. T1; T3 vs. T1-2; T4 vs. T1-3
TNM N-Deskriptor*	N0 vs. N1; N0/1 vs. N2
UICC Stadium (UICC 7. Edition) **	II vs. I; IIIA vs. I-II
Histologie	
Histologischer Subtyp	SCC vs. ADC sonstige vs. SCC/ADC
Behandlungsvariablen	
Resektionsstatus	R0 vs. R1/x
Chemotherapie	ja vs. nein
Bestrahlung	ja vs. nein

* In der Mehrheit der Fälle war eine histologische Sicherung gegeben. In diesen Fällen wurde das pTNM-Stadium verwendet. Andernfalls wurde auf das klinische cTNM-Stadium zurückgegriffen.

** Die Diagnosezeitpunkte der untersuchten Kohorte lagen zum Großteil im Gültigkeitszeitraum der 7. Auflage der UICC TNM-Klassifikation, teilweise jedoch auch in dem der 6. und 8. Auflage. Zur Vereinheitlichung erfolgte in letzterem Fall eine Umcodierung des Tumorstadiums entsprechend der 7. Auflage.

Aufgrund des hohen Grades an Multikollinearität mussten einzelne Variablen aus der multivariablen Analyse ausgeschlossen werden, um die Stabilität der Cox-Regression nicht zu gefährden. Unter den PET Parametern wurden das MTV und der SUV_{max} ausgeschlossen, da sie signifikant mit der ASP korrelierten (s. Kapitel 4.1.4). Weiterhin wurden T- und N-Deskriptor ausgeschlossen, da sie das prognostisch maßgebliche UICC-Stadium definieren und somit neben diesem redundant werden.

Methoden

Nach Einschluss der ASP in das multivariable Cox-Modell erfolgte die weitere Variablenauswahl unter Zuhilfenahme des Akaike'schen Informationskriterium (AIC). Für alle Variablenkombinationen wurde das AIC des zugehörigen Modells berechnet. Modelle mit einem $\Delta AIC \leq 2$ im Vergleich zum besten Modell (minimales AIC) wurden in Anlehnung an Empfehlungen aus der Literatur als gleichwertig betrachtet (105,106). Aus diesen Modellen wurde zur Reduktion von Redundanzen dasjenige mit der geringsten Parameterzahl ausgewählt. Dabei wurde in nicht-eindeutigen Fällen den demographischen Variablen Geschlecht, Alter und ECOG-Status Vorzug gegeben, da für diese als etablierte Einflussgrößen und mögliche Confounder a priori eine Relevanz für die Überlebenszeitanalyse zu erwarten ist. Eine Übersicht der betrachteten Modelle ist in Tabelle 6 ausgewiesen. Diese Vorgehensweise resultierte in den Einschluss der sechs Variablen *ASP*, *Geschlecht*, *Alter*, *ECOG-Status*, *UICC-Stadium* und *R-Status* sowohl für das DFS als auch für das OS.

Methoden

Tabelle 6: Akaike Informationskriterium (AIC) für verschiedene Cox-Regressionsmodelle

AIC des Modells	Anzahl Variablen	Zusätzlich zur ASP eingeschlossen Modellvariablen							
		Geschlecht	ECOG-Status	Alter	UICC-Stadium	Histologie	R-Status	Chemotherapie	Strahlentherapie
DFS									
1806,56	9	•	•	•	•	•	•	•	•
1806,65	8	•	•	•	•	•	•	•	
1806,96	7	•	•		•	•	•	•	
1807,23	8	•	•		•	•	•	•	•
1807,72	7	•	•	•	•		•	•	
1807,78	7	•	•	•	•	•	•		
1807,85	8	•	•	•	•		•	•	•
1807,99	6	•	•		•		•	•	
1808,07	6	•	•	•	•		•		
1808,52	6	•	•		•		•	•	•
...									
OS									
1421,83	7	•	•	•	•		•	•	
1423,09	6	•		•	•		•	•	
1423,24	6	•	•	•	•		•		
1423,38	8	•	•	•	•		•	•	•
1424,29	7	•	•	•	•		•		•
1424,80	7	•		•	•		•	•	•
1425,04	8	•	•	•	•	•	•	•	
1426,00	7	•		•	•	•	•	•	
...									

Erläuterung: Für sämtliche 2^8 (= 256) Variablenkombinationen erfolgte die Berechnung der AIC. Die Modelle wurden anschließend nach ihren AIC-Werten aufsteigend sortiert. Kleinere AIC-Werte entsprechen einer besseren Modellanpassung. Dargestellt sind die Modelle mit den kleinsten AIC-Werten d.h. die Modelle mit der besten Modellanpassung. Das Symbol • indiziert den Einschluss der jeweiligen Variable im Modell. In grün ist das in dieser Arbeit verwendete Modell markiert.

Assoziation der ASP mit tumorbezogenen Variablen

Die Assoziation der ASP mit den einzelnen UICC Stadien (I, II und IIIA) wurden mithilfe des Kruskal-Wallis Tests untersucht. Auf gleiche Weise wurde innerhalb der einzelnen UICC-Stadien auf Unterschiede der ASP zwischen den einzelnen T- und N-Deskriptoren getestet. Die Assoziation der ASP mit dem histologischen Subtyp sowie mit dem R-Status wurde mithilfe des Mann-Whitney U Tests untersucht. Die Nullhypothese war in jedem Fall das Fehlen eines Unterschieds der Verteilung der ASP zwischen den jeweiligen Gruppen.

Assoziation der ASP mit MTV und SUV_{max}

Zur Identifikation signifikanter Korrelationen zwischen den PET-Parametern erfolgte die paarweise Bestimmung des Pearson-Korrelationskoeffizienten. Die Berechnung erfolgt mithilfe der *scipy* Bibliothek (Version 1.8.0) in Python.

Subgruppenanalysen

Die prognostische Bedeutung der ASP wurde innerhalb der einzelnen UICC Stadien (I, II und IIIA) und T/N-Stadien untersucht. Ferner wurde gesondert diejenige Patient*innengruppe in den Stadien II und IIIA betrachtet, die einer prä- oder postoperativen Chemotherapie zugeführt wurde.

Es erfolgte hierzu eine multivariable Cox-Regression mit Adjustierung für Alter, Geschlecht, ECOG-Status und R-Status.

3.4.2. Einfluss verschiedener Segmentierungsmethoden

Übereinstimmung von MTV und ASP verschiedener Segmentierungsmethoden

Es erfolgte die paarweise Prüfung der Übereinstimmung von MTV bzw. ASP der untersuchten Segmentierungsmethoden mit der Referenzmethode (iAT). Hierzu wurde der Intra-Class Correlation Coefficient (ICC) für *absolute agreement* im *single fixed raters* Modell mithilfe der *pingouin* Bibliothek (Version 0.5.0) in Python berechnet.

Prognostische Validität – Vergleich zur iAT-Methode

Für jede Segmentierungsmethode erfolgte eine multivariable Cox-Regression für DFS und OS. Dabei wurde, wie zuvor für die Referenzmethode (iAT) beschrieben, nach Geschlecht, Alter, ECOG-Status, UICC-Stadium und R-Status adjustiert (s. Kapitel 3.4.1, Variablenselektion). Fälle, in denen eine Segmentierung mit der jeweiligen Methode nicht möglich war, wurden aus der Analyse ausgeschlossen. Für jede Methode erfolgte ein Vergleich mit der Referenzmethode.

4. Resultate und Auswertung

4.1. Validität der ASP als prognostischer Marker

4.1.1. Datensatz

Sämtliche erhobenen Variablen werden in Tabelle 7 dargestellt.

Bei einem medianen Alter von 67 Jahren und hohem Anteil männlicher Patienten (65%) ist die demographische Verteilung der Studienpopulation vergleichbar mit den Daten des ZfKD 2014 (1). Von den 361 untersuchten Patient*innen war etwa die Hälfte (47%) dem Tumorstadium UICC I zuzuordnen. Die weiteren Patient*innen teilten sich in annähernd gleichen Anteilen auf die Stadien UICC II (27%) und UICC IIIA (26%) auf.

In je etwa der Hälfte der Fälle lag ein Adenokarzinom (ADC, 51%) oder ein Plattenepithelkarzinom (SCC, 42%) vor. Ein kleiner Teil der Patient*innen wiesen andere (z.B. ein großzelliges oder adenosquamöses Karzinom) oder nicht weiter spezifizierte histologische NSCLC-Entitäten auf (7%).

Ein Ereignis bezüglich des DFS (definiert als lokale oder ferne Progression, Rezidiv oder Tod jeglicher Ursache) trat in 178 (49%) der insgesamt 361 Fälle nach im Median 13 Monaten auf (IQR: 8 - 24 Monate; Tabelle 8). Zudem traten 137 (44%) Todesfälle nach im Median 20 Monaten auf (IQR: 10 - 35 Monate). Das Follow-up der Überlebenden betrug im Median 37 Monate (IQR: 23 - 63 Monate).

Resultate und Auswertung

Tabelle 7: Deskriptive Statistik

Demographie			
Alter bei Diagnosestellung (Jahre)	Median: 67, IQR: 59 - 73, Range: 41 - 87		
Geschlecht	weiblich:	125 (35%)	
	männlich:	236 (65%)	
ECOG Performance Status	ECOG < 2:	321 (89%)	
	ECOG ≥ 2:	40 (11%)	
Tumorstadium			
UICC Stadium (7. Auflage)	<u>UICC I</u>	<u>UICC II</u>	<u>UICC III</u>
	IA 92 (25%)	IIA 49 (14%)	IIIA 93 (26%)
	IB 80 (22%)	IIB 47 (13%)	
Histologie			
NSCLC - Subtyp	Adenokarzinom (ADC)		183 (51%)
	Plattenepithelkarzinom (SCC)		141 (42%)
	weitere / nicht angegeben		17 (7%)
Behandlung			
Resektion	R0:	339 (94%)	R1/x: 22 (6%)
	Chemotherapie	nein:	223 (62%)
			adjuvant: 114 (32%)
			neoadjuvant: 36 (10%)
Radiotherapie	nein:	312 (86%)	ja: 49 (14%)
			adjuvant: 43 (12%)
			neoadjuvant: 6 (25%)
PET-Parameter			
ASP in %	Median: 10 (IQR: 4 - 25 ; Range: 0 ,0- 143)		
MTV in ml	Median: 8 (IQR: 3 - 19 ; Range: 0,5 - 313)		
SUV _{max} in SUV	Median: 11 (IQR: 6 - 16 ; Range: 1,0 - 55)		

Tabelle 8: Übersicht: Überlebenszeiten nach UICC Stadium

	UICC Stadium			Gesamt
	I	II	IIIA	
Anzahl Fälle	172	96	93	361
Anzahl Ereignisse	67	50	61	178
Anzahl Todesfälle	51	45	44	140
Überlebenszeiten in Monaten mit 95%-Konfidenzintervall:				
medianes DFS	63 (50 - 76)	32 (10 - 53)	18 (13 - 23)	39 (29 - 50)
medianes OS	nicht erreicht	52 (42 - 61)	46 (23 - 70)	63 (47 - 80)
medianes Follow-up (DFS)	43 (33 - 54)	46 (23 - 68)	42 (36 - 49)	44 (35 - 52)
medianes Follow-up (OS)	45 (35 - 54)	60 (39 - 80)	48 (34 - 62)	46 (37 - 54)

4.1.2. Cox Regression für DFS und OS

Es erfolgte eine univariable und multivariable Cox-Regression für das DFS und OS. In das multivariable Cox-Modell wurden die Variablen ASP, Geschlecht, Alter, ECOG-Status, UICC-Stadium sowie R-Status eingeschlossen. Angegeben sind die Hazard Ratio (HR), das 95%-Konfidenzintervall (95%CI) sowie die Signifikanzwerte (p) der Einzelparameter im Wald-Test.

Die berichteten Einheiten der kontinuierlichen Variablen (PET-Parameter und Alter) wurden so gewählt, dass sowohl die Einheit selbst als auch die resultierenden Effektgrößen leicht zu interpretieren sind. So wurde z.B. im Falle der ASP die Einheit 20% anstelle 1% gewählt, da eine Änderung von 20% (Δ ASP) intuitiv fassbarer ist als eine Änderung um 1% (s. Abb. 2) und die resultierenden Hazard Ratios („1,23 pro 20%“ anstelle „1,01 pro 1%“) leichter zu handhaben sind.

Disease Free Survival

Tabellen 9 und 10 zeigen univariable und multivariable Cox-Regression für das krankheitsfreie Überleben (DFS).

In der univariablen Analyse (s. Tabelle 9) erwiesen sich die PET-Parameter ASP, MTV und SUV_{max} , das Tumorstadium (UICC Stadium und im Einzelnen auch die T- und N-Deskriptoren), sowie die Therapievariablen R-Status und Radiotherapie, nicht jedoch die Chemotherapie als prognostisch signifikante Variablen. Für den histologischen Subtyp wurde keine Assoziation mit dem krankheitsfreien Überleben beobachtet. Männer wiesen im Vergleich zu Frauen ein kürzeres DFS

Resultate und Auswertung

auf. Prognostisch ungünstig waren zudem ein reduzierter Allgemeinzustand (ECOG-Status ≥ 2) sowie ein hohes Alter bei Diagnosestellung.

In der multivariablen Analyse (s. Tabelle 10) zeigen sich weiterhin ASP, Geschlecht, Alter, ECOG-Status, UICC-Stadium sowie R-Status als signifikante prognostische Faktoren. Die zugehörigen geschätzten Effektgrößen fallen im Vergleich zur univariablen Analyse schwächer aus, wobei die Unterschiede für UICC-Stadium (IIIA vs. I/II) sowie R-Status deutlich ausgeprägter sind als für ASP und demographische Variablen.

Tabelle 9: Univariable Cox-Regression für das DFS

Variablen		HR	95%CI	p-Wert
Demographie	Geschlecht (männlich vs. weiblich)	1,77	1,27 - 2,47	0,001
	Alter (pro 10 Jahre)	1,22	1,03 - 1,44	0,020
	ECOG ≥ 2 vs. < 2	2,07	1,31 - 3,28	0,002
Tumor	ASP (pro 20 %)	1,23	1,12 - 1,34	< 0,001
	MTV (pro 30 ml)	1,13	1,03 - 1,23	0,007
	SUV _{max} (pro 10 SUV)	1,24	1,02 - 1,50	0,031
	UICC Stadium			< 0,001
	Stadium II vs. I	1,74	1,21 - 2,50	0,003
	Stadium IIIA vs. I/II	1,81	1,32 - 2,47	< 0,001
	T-Deskriptor			0,003
	T2 vs. T1	1,04	0,73 - 1,46	0,842
	T3 vs. T1/2	1,79	1,25 - 2,58	0,002
	T4 vs. T1/2/3	1,60	0,84 - 3,04	0,149
	N-Deskriptor			0,002
	N1 vs. N0	1,61	1,12 - 2,32	0,010
	N2 vs. N0/1	1,31	0,91 - 1,90	0,149
Histologie			0,796	
SCC vs. ADC	1,23	0,91 - 1,66	0,185	
Sonstige vs. SCC/ADC	0,53	0,27 - 1,05	0,067	
Behandlung	Rx/1 vs. R0 Resektion	3,49	2,18 - 5,57	< 0,001
	Chemotherapie: ja	1,05	0,78 - 1,42	0,737
	Strahlentherapie: ja	1,59	1,08 - 2,34	0,018

Resultate und Auswertung

Tabelle 10: Multivariable Regression für das DFS

Variable		HR	95%CI	p-Wert
Demographie	Geschlecht (männlich vs. weiblich)	1,61	1,14 – 2,26	0,007
	Alter (pro 10 Jahre)	1,19	1,00 – 1,41	0,045
	ECOG ≥ 2 vs. < 2	2,32	1,45 – 3,71	< 0,001
Tumor	ASP (pro 20 %)	1,13	1,02 – 1,25	0,017
	UICC Stadium			0,001
	Stadium II vs. I	1,81	1,24– 2,65	0,002
	Stadium IIIA vs. I/II	1,45	1,03 – 2,04	0,034
Behandlung	Rx/1 vs. R0 Resektion	2,05	1,23 – 3,39	0,006

Overall Survival

Wie für das DFS zeigt sich auch für das OS in der univariablen Analyse (s. Tabelle 11) eine prognostische Relevanz von demographischen Variablen (Alter, Geschlecht, ECOG-Status) sowie von UICC-Stadium und T- bzw. N-Deskriptoren. Unter den PET-Parametern zeigten sich die ASP und der SUV_{max} , nicht jedoch das MTV, signifikant. Im Gegensatz zum DFS konnte für das OS eine Assoziation mit dem histologischen Subtyp nachgewiesen werden, wobei ein SCC im Vergleich zum ADC mit einem höheren Sterberisiko einherging. Unter den Behandlungsvariablen erschien allein der R-Status prognostisch bedeutsam.

In der adjustierten Analyse (s. Tabelle 12) zeigen sich weiterhin demographische Variablen (Geschlecht, Alter und ECOG-Status), ASP und R-Status prognostisch bedeutsam. Für das UICC-Stadium ergab sich ein signifikanter Unterschied im Überlebensrisiko zwischen Stadium I und II, der Unterschied zwischen den Stadien I/II und IIIA war jedoch nicht statistisch signifikant.

Resultate und Auswertung

Tabelle 11: Univariable Regression für das OS

Variablen		HR	95%CI	p-Wert
Demographie	Geschlecht (männlich vs. weiblich)	1,90	1,30 - 2,78	0,001
	Alter (pro 10 Jahre)	1,44	1,19 - 1,75	< 0,001
	ECOG \geq 2 vs. < 2	1,80	1,08 - 3,00	0,024
Tumor	ASP (pro 20 %)	1,21	1,10 - 1,34	< 0,001
	MTV (pro 30 ml)	1,10	0,99 - 1,23	0,081
	SUV _{max} (pro 10 SUV)	1,25	1,01 - 1,55	0,043
	UICC Stadium			< 0,001
	Stadium II vs. I	1,95	1,31 - 2,92	0,001
	Stadium IIIA vs. I/II	1,52	1,07 - 2,18	0,021
	T-Deskriptor			0,089
	T2 vs. T1	0,80	0,54 - 1,17	0,244
	T3 vs. T1/2	1,61	1,06 - 2,44	0,026
	T4 vs. T1/2/3	1,79	0,87 - 3,67	0,113
	N Deskriptor			0,008
	N1 vs. N0	1,84	1,25 - 2,72	0,002
	N2 vs. N0/1	1,02	0,66 - 1,57	0,923
	Histologischer Subtyp			0,124
	SCC vs. ADC	1,55	1,09 - 2,19	0,014
Sonstige vs. SCC/ADC	0,86	0,45 - 1,63	0,639	
Behandlung	Rx/1 vs. R0 Resektion	3,34	1,97 - 5,65	< 0,001
	Chemotherapie: ja vs. nein	0,87	0,61 - 1,22	0,416
	Strahlentherapie: ja vs. nein	0,98	0,61 - 1,59	0,943

Resultate und Auswertung

Tabelle 12: Multivariable Regression für das OS

Variable		HR	95%CI	p-Wert
Demographie	Geschlecht (männlich vs. weiblich)	1,76	1,19 – 2,60	0,005
	Alter (pro 10 Jahre)	1,44	1,18 – 1,76	< 0,001
	ECOG (≥ 2 vs. < 2)	1,92	1,14 – 3,22	0,014
Tumor	ASP (pro 20%)	1,13	1,00 – 1,27	0,043
	UICC Stadium			< 0,001
	Stadium II vs. I	2,10	1,38 – 3,19	0,001
	Stadium IIIA vs. I/II	1,20	0,82 – 1,78	0,347
Behandlung	Rx/1 vs. R0 Resektion	2,52	1,45 – 4,37	0,001

4.1.3. Subgruppenanalysen

UICC-Stadium und TNM-Klassifikation

Die beobachteten unadjustierten und adjustierten HR für DFS und OS innerhalb der einzelnen UICC Stadien I - IIIA sind in Abb. 6 dargestellt. Für die Stadien UICC I und II werden durchschnittliche (bezogen auf die Gesamtkohorte) bis überdurchschnittliche HR geschätzt. Für das Stadium IIIA ist das Gegenteil der Fall. Die Ergebnisse der adjustierten Analyse sind bei nur kleiner Anzahl von Fällen bzw. Ereignissen in den einzelnen Subgruppen nicht statistisch signifikant.

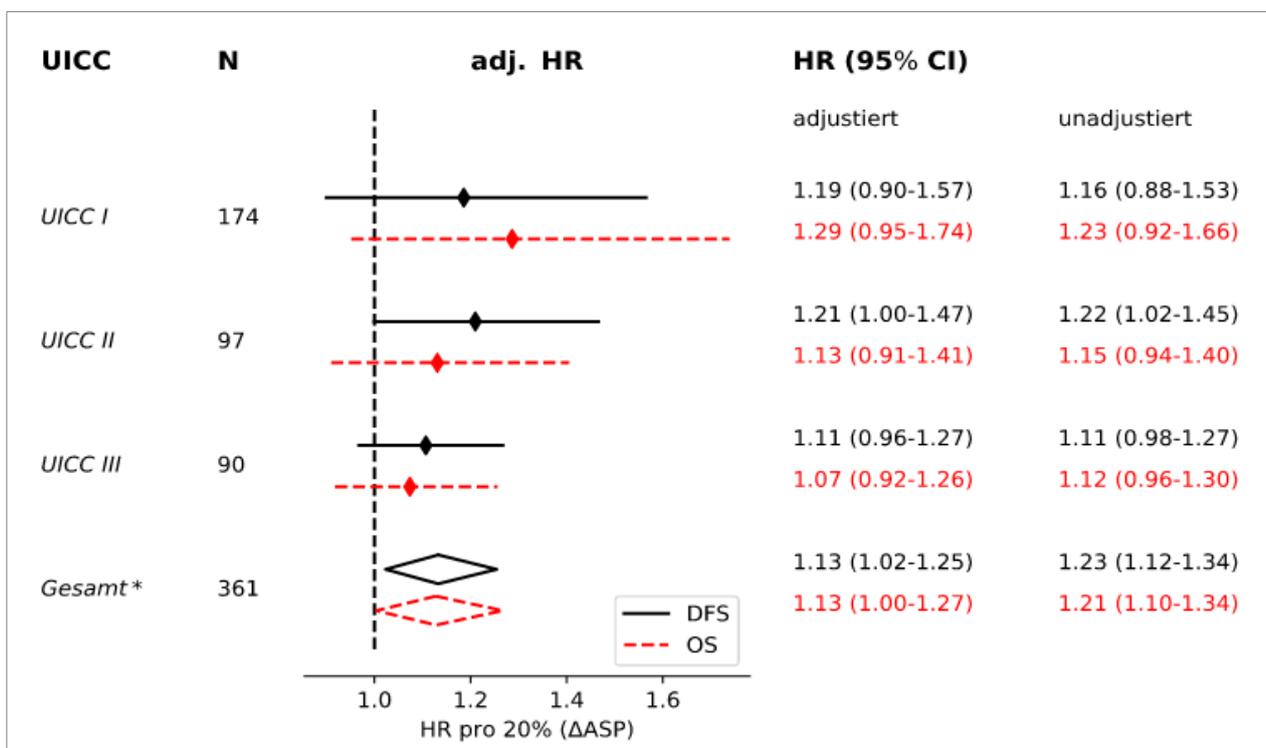


Abbildung 6: **HR für DFS und OS: Stratifikation nach UICC Stadium.** Die adjustierte Gesamtanalyse (*) berücksichtigt das UICC-Stadium und entspricht somit den Ergebnissen aus Kapitel 4.1..2.

Abb. 7 zeigt die unadjustierten und adjustierten HR für DFS und OS bei Stratifikation nach T/N-Deskriptor. Aufgrund niedriger Fallzahlen wurden T3- und T4-Stadium sowie N1- und N2-Stadium jeweils zusammengefasst. Wie auch im Falle des UICC-Stadiums werden in frühen Stadien (bis T1 N1/2) überdurchschnittlich hohe HR geschätzt, in höheren Stadien ist das Gegenteil der Fall. Für die Stadien T2-4 N0 sind die Ergebnisse statistisch signifikant.

Resultate und Auswertung

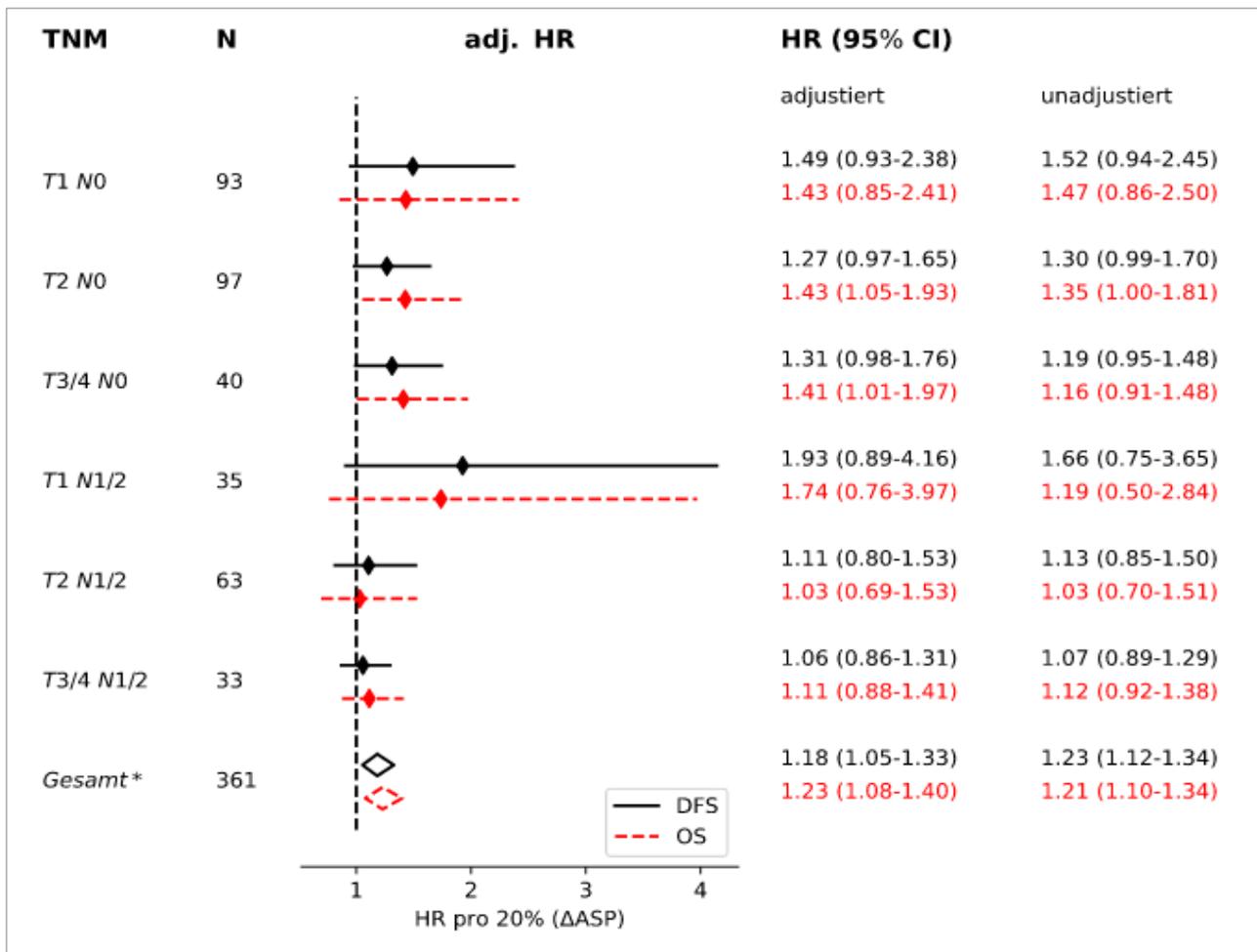


Abbildung 7: HR für DFS und OS: Stratifikation nach T/N-Deskriptoren. Die Gesamtanalyse (*) berücksichtigt eine Adjustierung für das T- und N-Stadium.

Chemotherapie

Betrachtet wurde die Subgruppe von 138 Patient*innen aller Stadien, die einer prä- und/oder postoperativen Chemotherapie unterzogen wurden. Hier zeigte sich die ASP als signifikanter prognostischer Marker mit einem adjustierten HR von 1,24 (95%CI: 1,07 - 1,46) für DFS und 1,23 (95%CI: 1,05 – 1,45) für OS pro 20% Δ ASP. Bei Stratifizierung in die UICC-Stadien bewährt sich die ASP im Stadium UICC II als prognostisch für das DFS und OS. Die Ergebnisse sind in Abb. 8 veranschaulicht.

Resultate und Auswertung

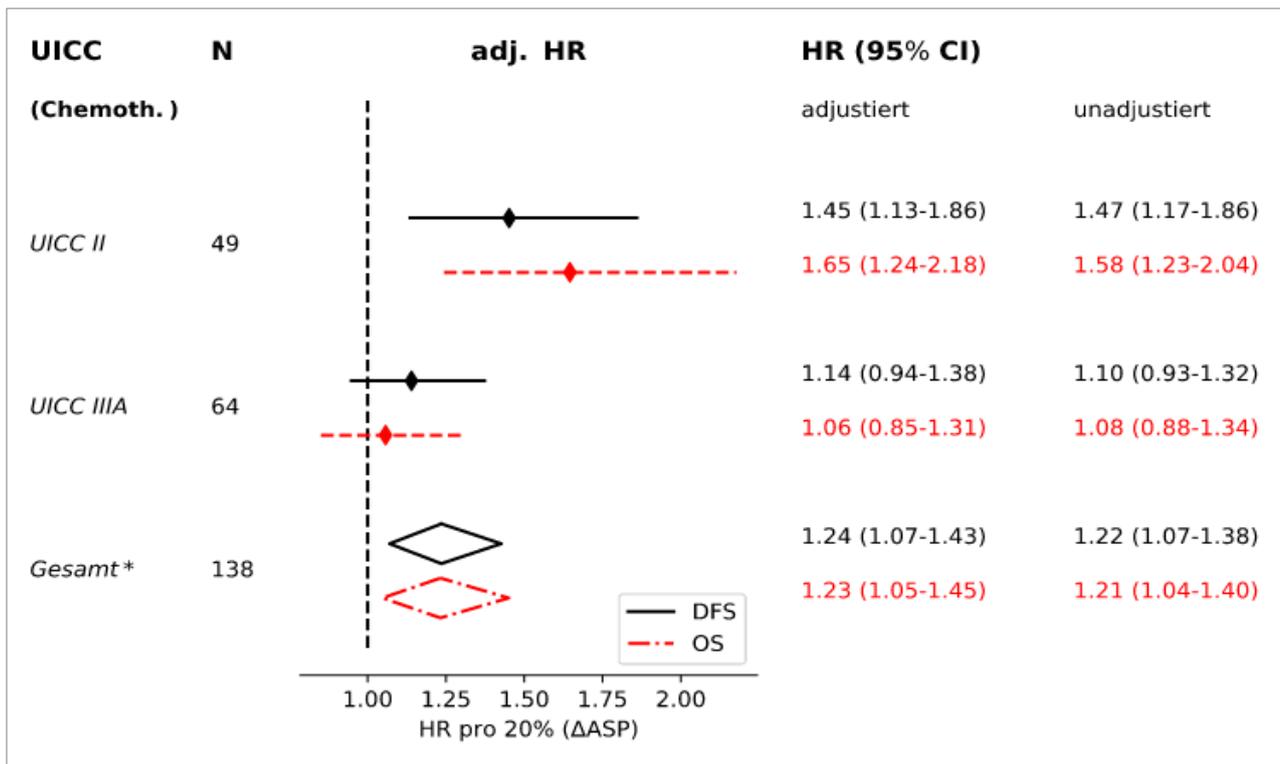


Abbildung 8: HR für DFS und OS: Gruppe der chemotherapeutisch behandelten Patient*innen, stratifiziert nach UICC Stadium. Eine isolierte Betrachtung der Gruppe UICC I war bei niedrigen Fall- und Ereigniszahlen nicht möglich. Die Gesamtanalyse (*) beinhaltet eine Adjustierung für das UICC Stadium. Sie bezieht auch die Patient*innen im Stadium UICC I mit ein.

4.1.4. Assoziationen von ASP mit anderen Variablen

Die folgenden Analysen zeigen die Assoziation der ASP zu anderen, tumorbezogenen Variablen mit dem Ziel die Beziehung der Variablen zueinander nachzuvollziehen und die Grenzen einer isolierten Betrachtung der jeweiligen prognostischen Komponenten aufzuzeigen. Mit Blick auf die multivariable Regression dienen sie der Erkennung hochgradiger Multikollinearität, welche die Stabilität des Regressionsmodells einschränken könnte.

UICC Stadium und T/N-Deskriptoren

Verglichen mit dem Stadium UICC I ist das Stadium UICC II mit höheren ASP-Werten assoziiert (Median: 13%; IQR: 7 bis 34% vs. Median: 7%; IQR: 3 bis 14%, Mann-Whitney U Test: $p < 0,001$). Für die ASP konnte zwischen den Stadien UICC II und UICC III (Median: 19%; IQR: 7 bis 42%) hingegen kein signifikanter Unterschied gezeigt werden (Mann-Whitney U Test: $p = 0,46$).

Resultate und Auswertung

Innerhalb der einzelnen UICC Stadien liegt auch eine signifikante Assoziation der ASP-Werte mit dem T-Deskriptor vor (Kruskal-Wallis-Test: jeweils $p < 0,001$), wobei höhere ASP-Werte mit höheren T-Stufen einhergehen (s. Abb. 9). In Stadium UICC II ist der N1-Status im Vergleich zum N0 Status mit einer *niedrigeren* ASP assoziiert (Mann-Whitney U Test: $p < 0,001$). Ursache hierfür ist der Umstand, dass gemäß UICC (7. Ausgabe) ein N0-Status nur dann zu einer UICC II Klassifizierung führen kann, wenn mindestens ein T2b Stadium vorliegt. Dies ist aber gerade mit einer höheren ASP assoziiert (s.o.).

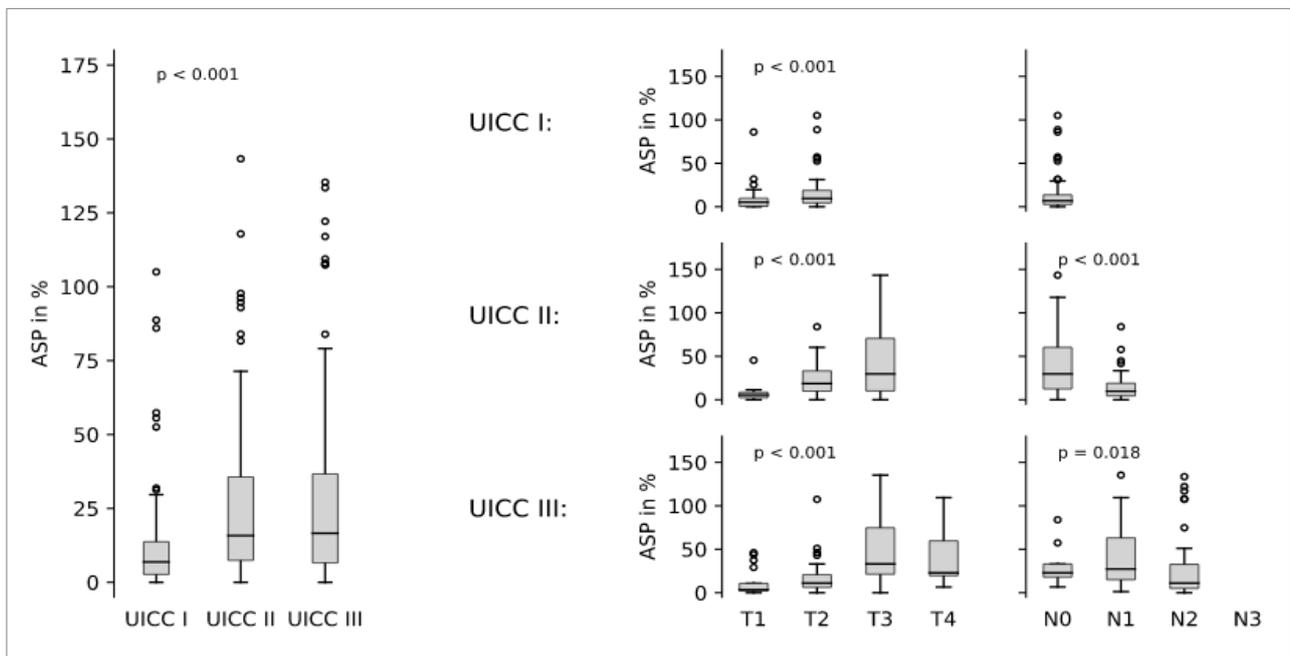


Abbildung 9: **Assoziation von ASP und UICC-Stadium sowie T- und N-Deskriptor. Darstellung als Boxplots.** Gezeigt ist der p-Wert für Unterschiede innerhalb der Gruppen im Mann-Whitney U-Test (2 Gruppen) bzw. Kruskal-Wallis-Test (> 2 Gruppen).

Histologischer Subtyp und R-Status

Verglichen mit dem Adenokarzinom (ADC) fanden sich für das Plattenepithelkarzinom signifikant höhere ASP-Werte (SCC: Median 13,7%; IQR 5,5 - 28,8%, vs. ADC: 8,0%; IQR, 3,3 - 18,4%, Mann-Whitney U Test: $p = 0,008$). Für die Gruppe der nicht weiter spezifizierten („sonstigen“) histologischen Entitäten (Median 9,8%; IQR 4,7 - 26,0%) zeigte sich im Mann-Whitney U Test kein signifikanter Unterschied der medianen ASP zum ADC ($p = 0,246$) oder SCC ($p = 0,119$). Wird die ASP zwischen der Gruppe tumorfrei resezierter (R0) und nicht-tumorfrei resezierter (R1) Patient*innen verglichen, so zeigt sich auch hier ein signifikanter Unterschied mit höheren ASP-

Resultate und Auswertung

Werten in der R1/x-Gruppe (s. Abb. 10). Keiner der R1/x resezierten Patient*innen wies eine ASP < 5% auf.

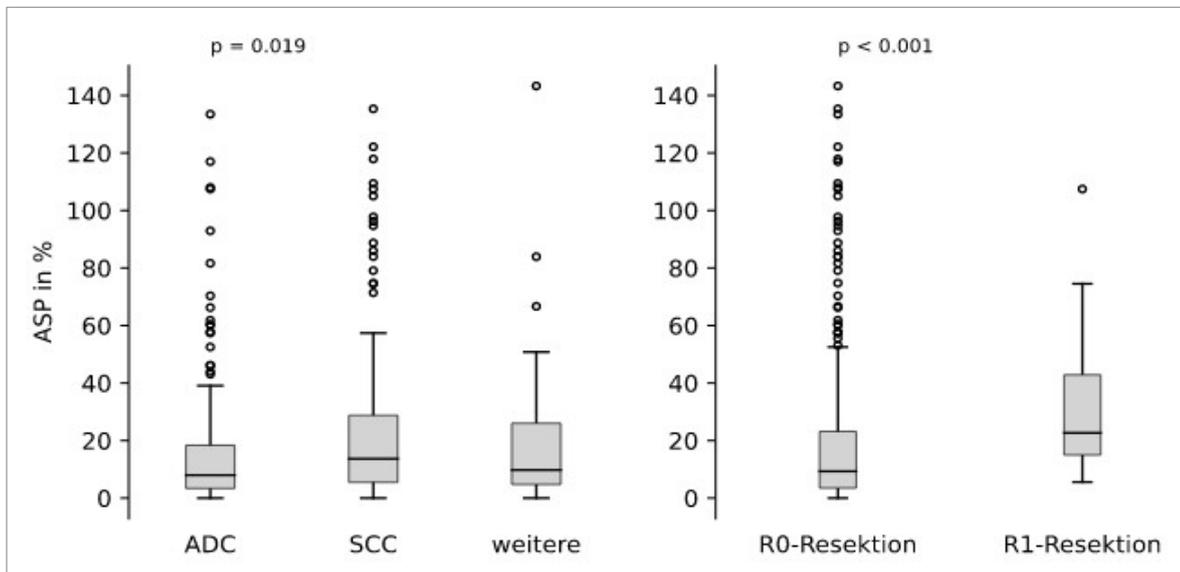


Abbildung 10: Assoziation von ASP mit histologischem Subtyp und Resektionsstatus. Gezeigt ist der p-Wert für Unterschiede innerhalb der Gruppen im Mann Whitney U-Test (2 Gruppen) bzw. Kruskal-Wallis-Test (> 2 Gruppen).

ASP, MTV und SUV_{max}

Korrelation der PET-Parameter miteinander

Abb. 11 zeigt die paarweise Abhängigkeit der ASP mit dem MTV und der SUV_{max} . In beiden Fällen liegt eine signifikante positive Korrelation vor, wobei diese für das MTV deutlich stärker ausfällt (Pearson $r = 0,568$, $p < 0,001$) als für den SUV_{max} (Pearson $r = 0,235$, $p < 0,001$). Weiterhin besteht auch zwischen dem MTV und SUV_{max} eine signifikante Korrelation (Pearson $r = 0,255$, $p < 0,001$) (nicht visualisiert).

Resultate und Auswertung

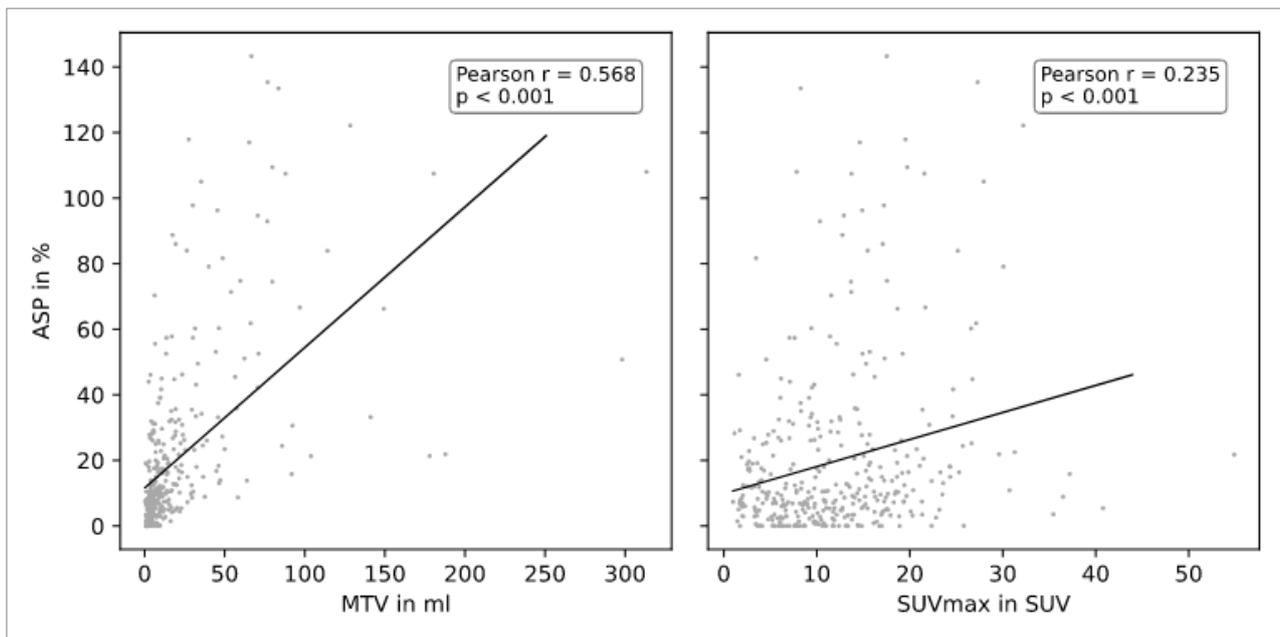


Abbildung 11: **Korrelation von ASP mit MTV und SUV_{max}**

Einfluss von MTV und SUV_{max} auf die Regressionsanalyse

In Abb. 12 ist der Effekt eines Einschlusses von MTV und/oder SUV_{max} als zusätzliche PET-Parameter neben der ASP gezeigt. Dazu erfolgte eine Cox-Regression für jede weitere PET-Parameterkombination, wobei analog zu Kapitel 5.1.2 auch hier für Geschlecht, Alter, ECOG-Status, UICC-Stadium und R-Status adjustiert wurde.

Für die ASP zeigen sich bei Einschluss weiterer PET-Parameter die geschätzten HR sehr robust mit nur minimalen Schwankungen des HR im Bereich zwischen 1,12 - 1,15 für DFS und OS. Die Ergebnisse zeigen sich jeweils bei Einschluss des MTV oder der SUV_{max} (für DFS) weiterhin signifikant. Bei gleichzeitigem Einschluss von MTV und SUV_{max} wird das Signifikanzniveau allerdings nicht mehr erreicht.

Im Gegensatz dazu ist das Verhalten des MTV bei Einschluss zusätzlicher Parameter deutlich heterogener. Während sich die Schätzungen für das DFS noch annähernd konsistent zeigen, ist dies für das OS nicht mehr der Fall. Es fällt auf, dass das MTV bei gleichzeitiger Adjustierung für ASP zu einer Umkehr des Sterberisikos zu führen scheint.

Im Falle der SUV_{max} zeigen sich konsistente Schätzungen für das HR, wobei hohe SUV-Werte mit einem höheren Risiko in Bezug auf DFS und OS assoziiert sind. Die Konfidenzintervalle sind im Vergleich zu den Schätzungen für ASP und MTV allerdings sehr groß.

Resultate und Auswertung

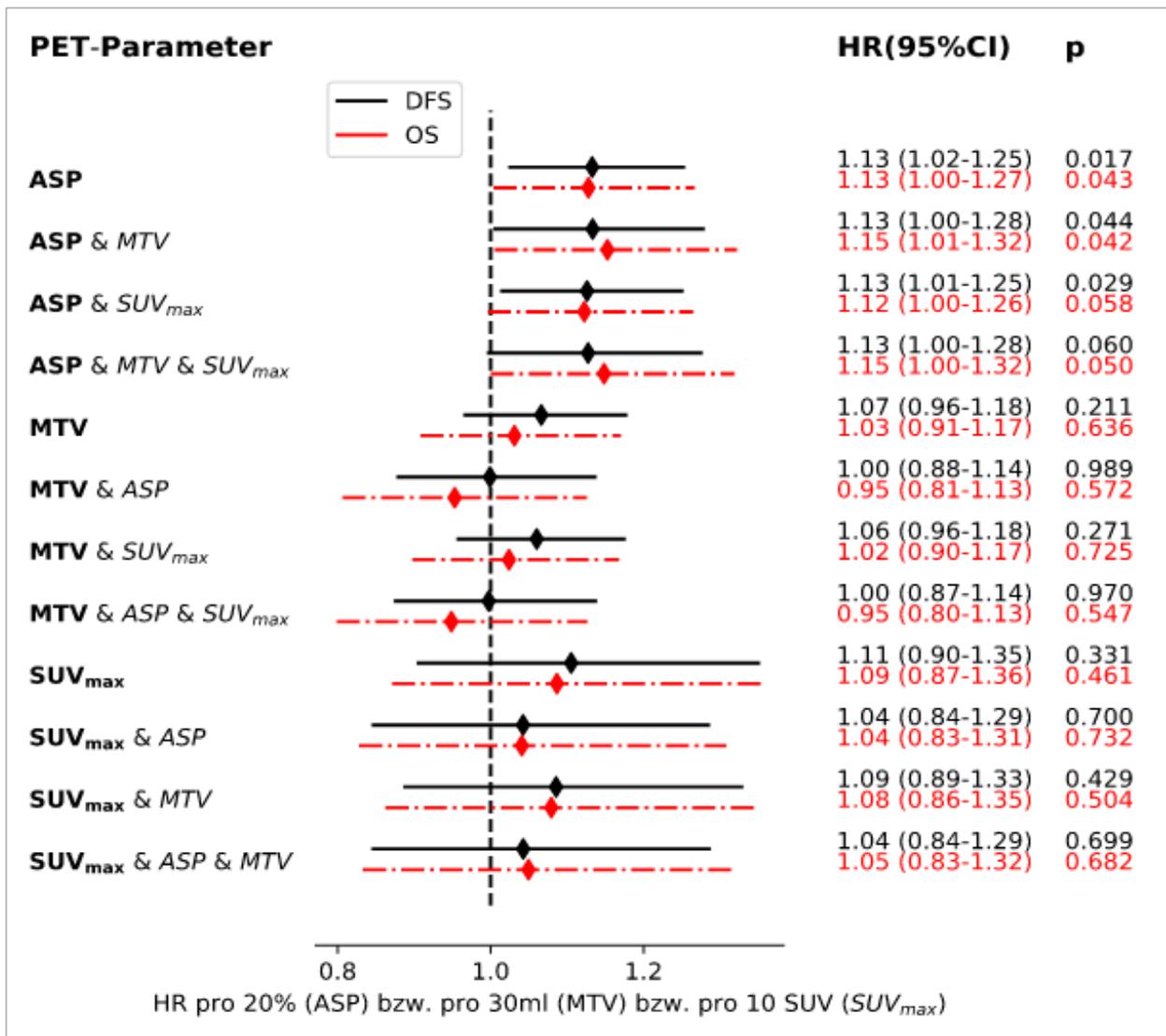


Abbildung 12: **PET-Parameter im Cox-Regressionsmodell: adjustierte HR der einzelnen PET-Parameter (fett gedruckt) für DFS und OS.** Dargestellt ist der Einfluss von nur einem, zweier oder aller PET-Parameter gleichzeitig als Kovariablen im Regressionsmodell.

4.2. Verschiedene Segmentierungsmethoden

Neben der iAT-Methode wurden 16 verschiedene Fixed Thresholding (FT)-Methoden zur Tumorsegmentierung untersucht. Abb. 13 gibt einen deskriptiven Überblick über die erhobene Verteilung von ASP und MTV.

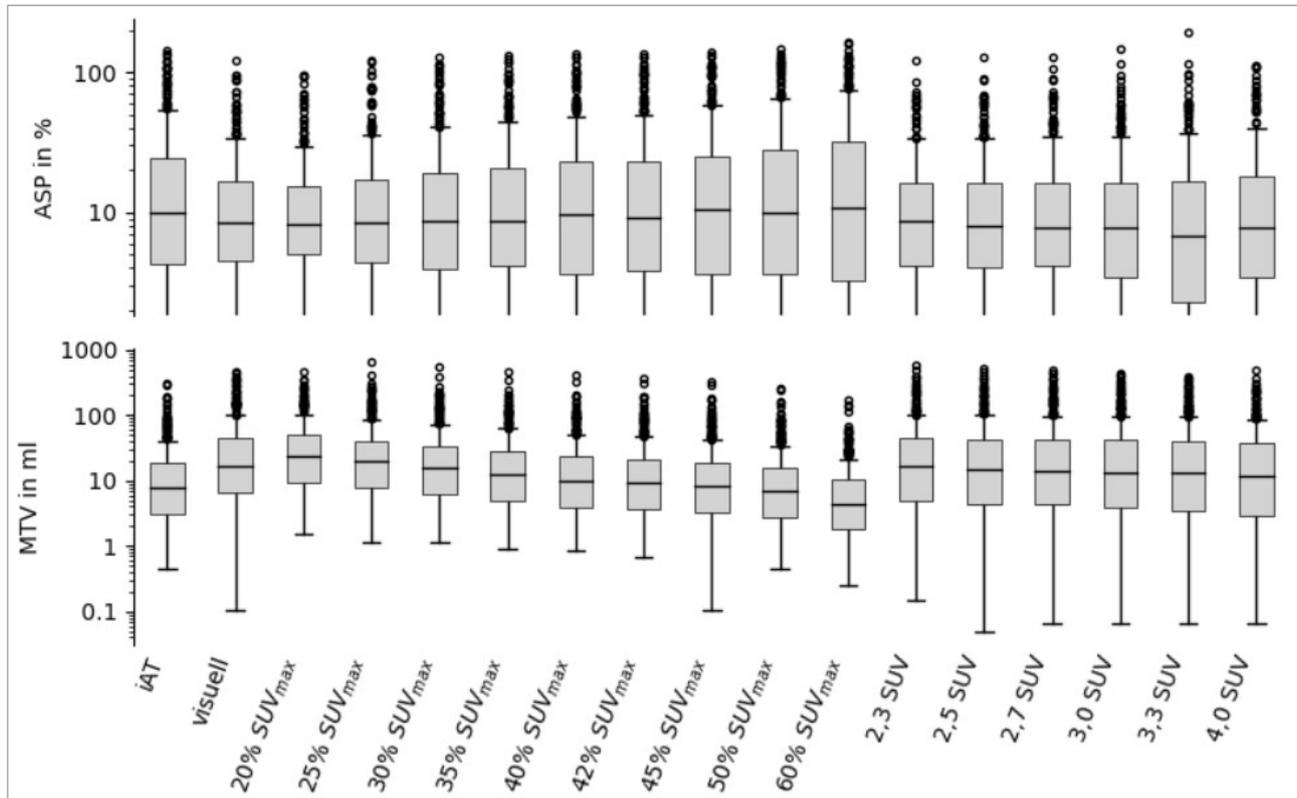


Abbildung 13: **ASP und MTV verschiedener Segmentierungsmethoden.** Logarithmische Darstellung. Deutlich zu erkennen ist, dass der Median des MTV stark mit der verwendeten Segmentierungsmethode variiert, während er im Falle der ASP vergleichsweise konstant bleibt. Eine höherschwellige Segmentierung führt allerdings zur Verbreiterung der Verteilung.

4.2.1. Übereinstimmung von ASP und MTV

Abbildungen 14 und 15 fassen die Übereinstimmung von ASP und MTV verschiedener Segmentierungsalgorithmen mit dem Referenzalgorithmus (iAT) zusammen, ausgedrückt durch Intraklassenkorrelations-Koeffizienten. Die höchste Übereinstimmung findet sich in beiden Fällen für relatives Thresholding mit Werten von 40% bis 45% vom SUV_{max} (ICC jeweils ≥ 0.98).

Resultate und Auswertung

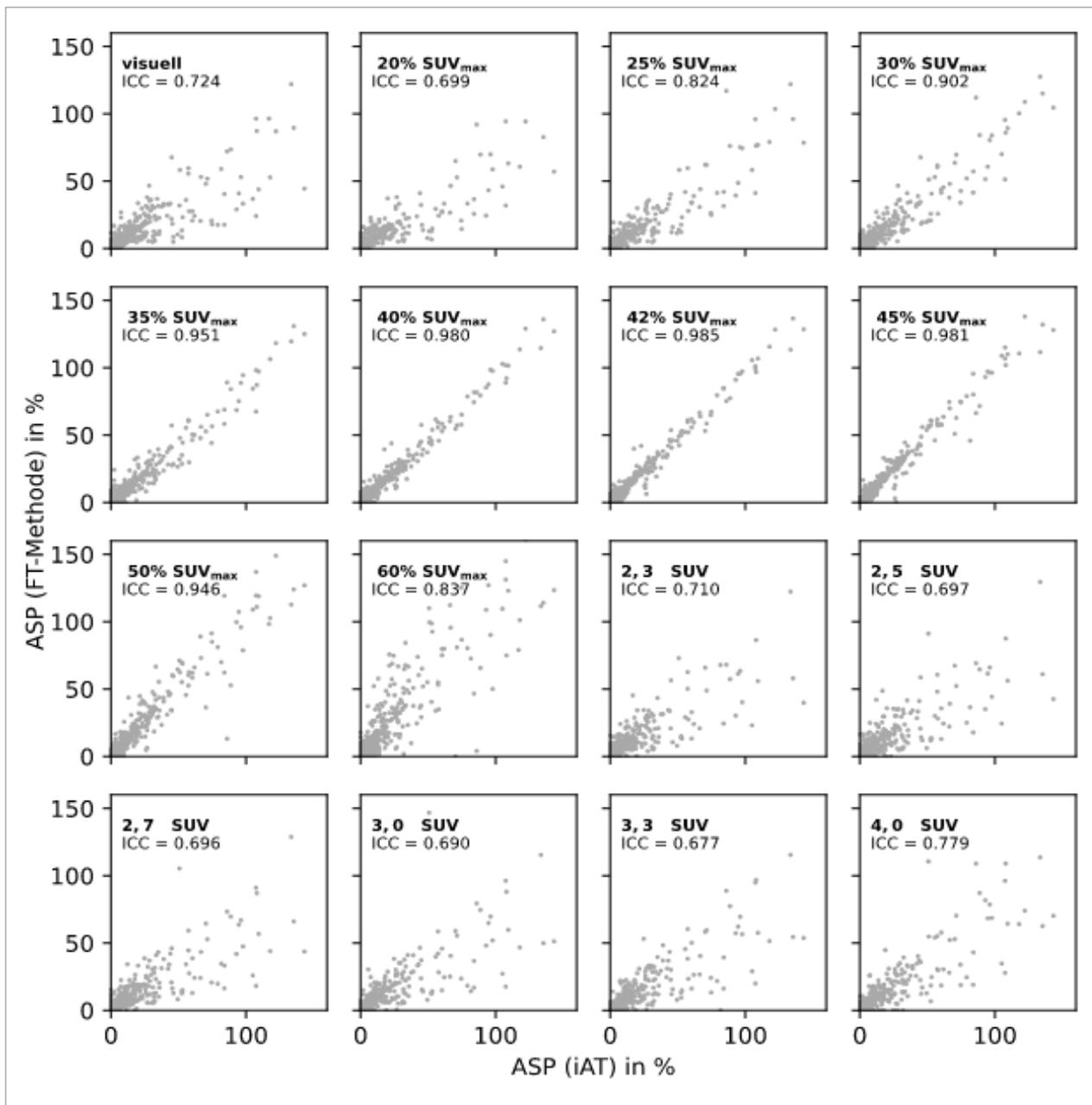


Abbildung 14: **Intraklassen-Korrelationskoeffizienten der ASP zwischen iAT und verschiedenen FT-Methoden.** Eine sehr gute Übereinstimmung liegt für relatives FT mit 40-45% des SUV_{max} vor. Die mit absoluten Schwellenwerten arbeitenden FT-Segmentierungen führen zu systematisch niedrigeren ASP-Werten im Vergleich zur iAT-Methode.

Resultate und Auswertung

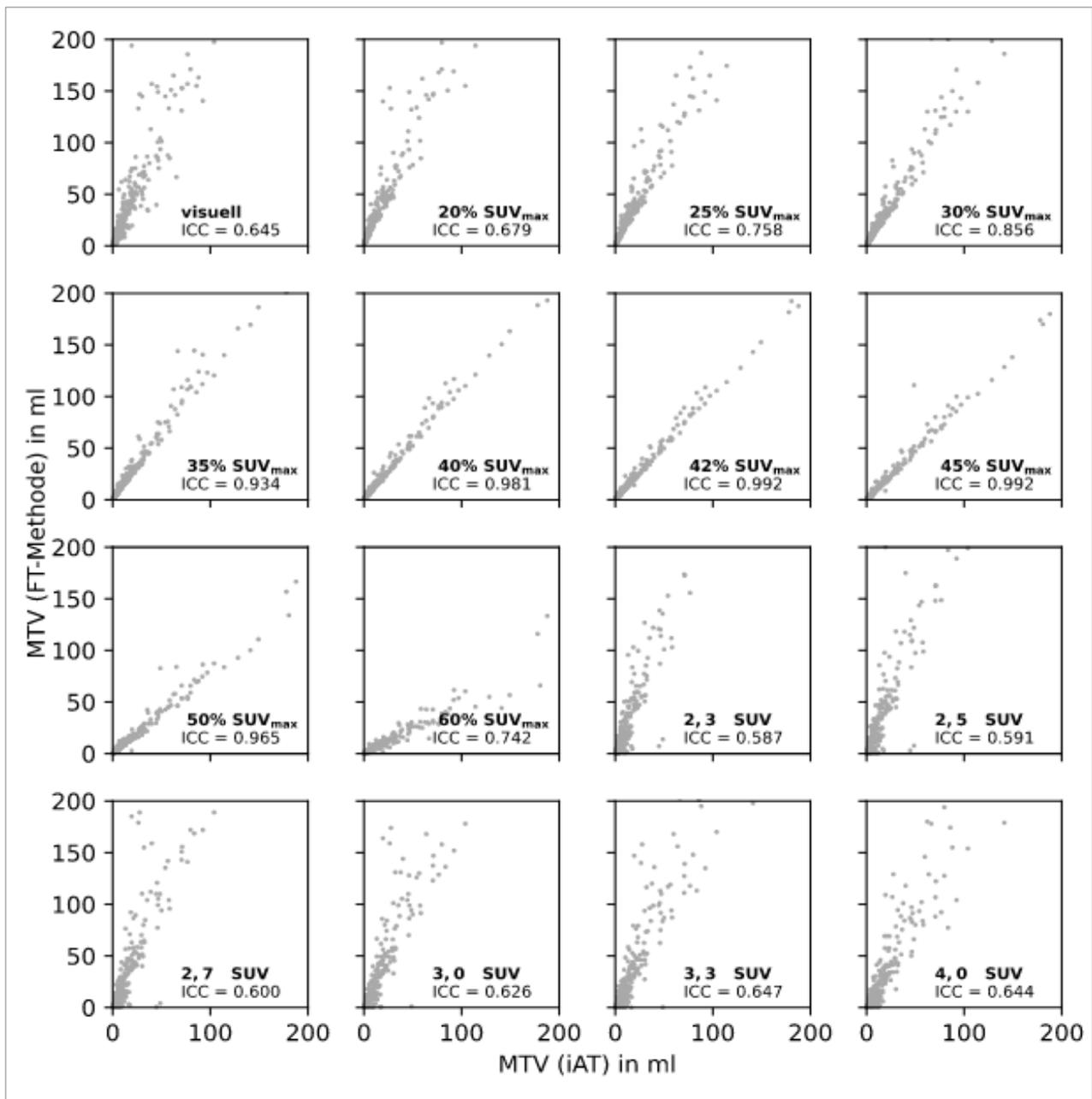


Abbildung 15: **Übereinstimmung des MTV zwischen iAT und verschiedenen FT-Methoden.** Die mit Absolutwerten arbeitenden FT-Algorithmen führen zu systematisch höheren MTV-Werten. Gleiches gilt für relative FT mit Schwellenwerten $\leq 35\%$ des SUV_{max} . Das Gegenteil ist für Schwellen $> 50\%$ des SUV_{max} der Fall. Die größte Übereinstimmung der MTV liegt für relative Schwellenwerte von 40%, 42% und 45% des SUV_{max} vor.

4.2.2. Erfolgsrate bei FT-Segmentierung

Nicht in jedem Fall konnte eine sinnvolle Segmentierung ohne größere Korrekturen erzielt werden. Abb. 16 zeigt die Erfolgsrate der einzelnen FT-Methoden.

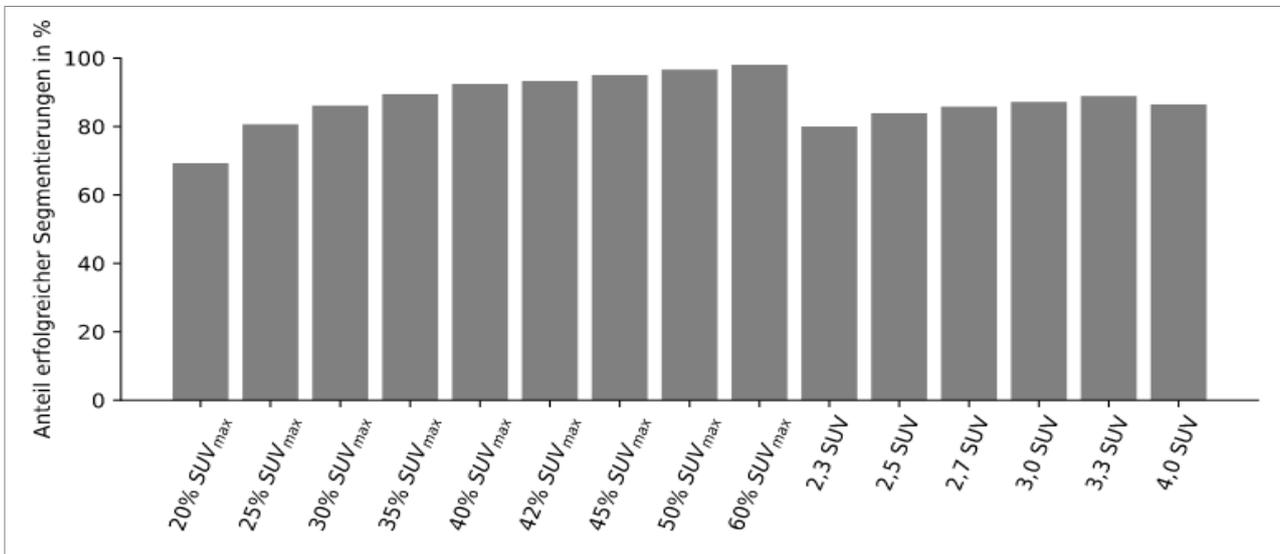


Abbildung 16: Erfolgsrate verschiedener Fixed Thresholding (FT) basierter Segmentierungsalgorithmen. Während beim relativen FT (links) die Erfolgsrate mit zunehmendem Schwellenwert stetig ansteigt, kann sie beim absoluten FT grundsätzlich wieder abfallen, da der Schwellenwert bei weniger [¹⁸F]FDG-aviden Tumoren den SUV_{max} auch übersteigen kann.

In den meisten Fällen scheiterte die Segmentierung einer Läsion an einer zu engen Nachbarschaft des Tumors zu physiologisch oder pathologisch FDG-avidem Gewebe. So führte die Nähe zu sehr stoffwechselaktiven Organen wie Herz oder Leber (s. Abb. 17) aber auch z.B. eine an den Tumor reichende, poststenotische Pneumonie (s. Abb. 18) zu ausgeprägten Überstrahlungseffekten, die eine Differenzierung der Tumorgrenzen in der PET erschwerte.

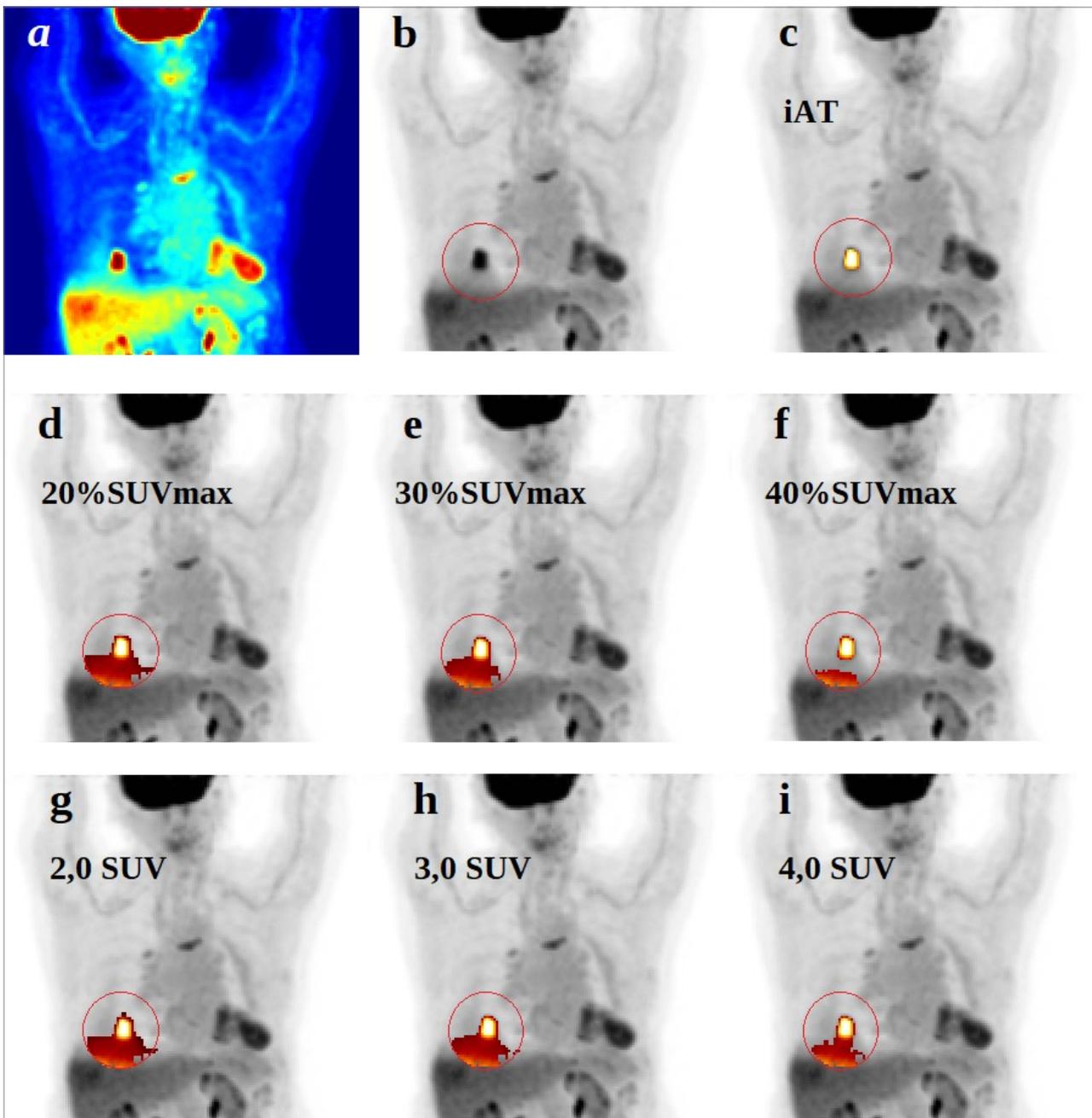


Abbildung 17: Überstrahlungseffekte durch Weichteilnähe. Maximum Intensity Projection (MIP) mit Darstellung eines NSCLC im rechten Unterlappen. Durch die Nähe zur Leber und deren Atembeweglichkeit kommt es zu Überstrahlungseffekten, welche gut in der farbskalierten Darstellung (a) zu erkennen sind. In (b) ist die Maske gezeigt, innerhalb der eine ROI segmentiert werden soll. Bei sehr niedrigschwelliger Segmentierung (d,e) kommt es zur Miterfassung von umgebendem Gewebe, sodass eine manuelle Korrektur auf Grundlage einer visuell vermuteten Tumorgrenze erforderlich wird. Erst ab einem relativen Schwellenwert von 40% (f) ist der Tumor klar abzugrenzen. Das hier durch die sehr groß gewählte Maske miterfasste Lebergewebe hat keine Verbindung zum Tumor und kann problemlos korrigiert werden. Für absolutes Thresholding im Bereich 2 - 4 SUV gelingt keine Darstellung der Tumorgrenzen (g-i). In (c) ist die Segmentierung bei Verwendung der iAT-Methode gezeigt.

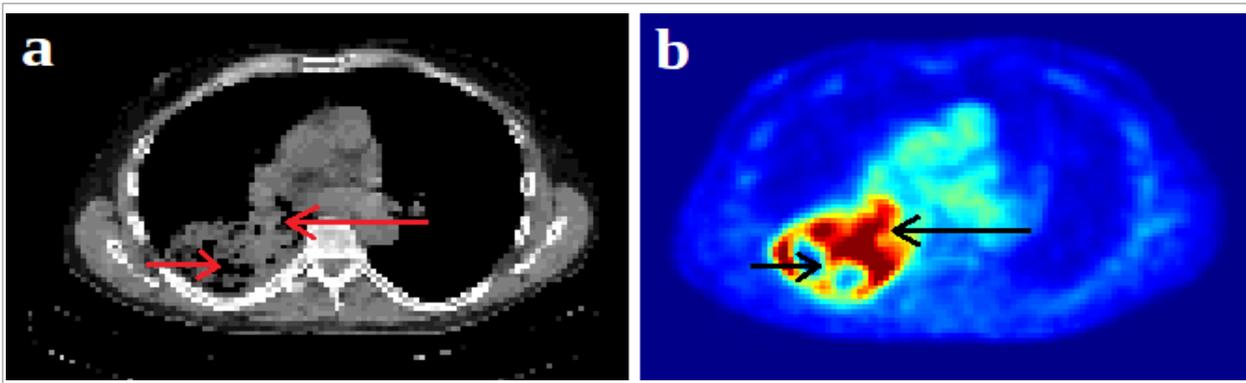


Abbildung 18: **Entzündlicher Begleiterguss bei NSCLC im rechten Unterlappen.** Low-Dose CT (a) und MIP PET (b) mit Darstellung eines Lungentumors im rechten Unterlappen mit poststenotischer Pneumonie. Eine exakte Konturierung des Tumors (langer Pfeil) auf Grundlage der PET ist in Folge einer Überstrahlung durch die umliegende Entzündung (kurzer Pfeil) nicht sicher möglich.

Je geringer die Intensität bzw. der Kontrast des Tumors zur Umgebung, desto häufiger versagen niedrighschwelligere Segmentierungsmethoden. Diese Problematik ist in Abb. 19 an einem Beispiel illustriert.

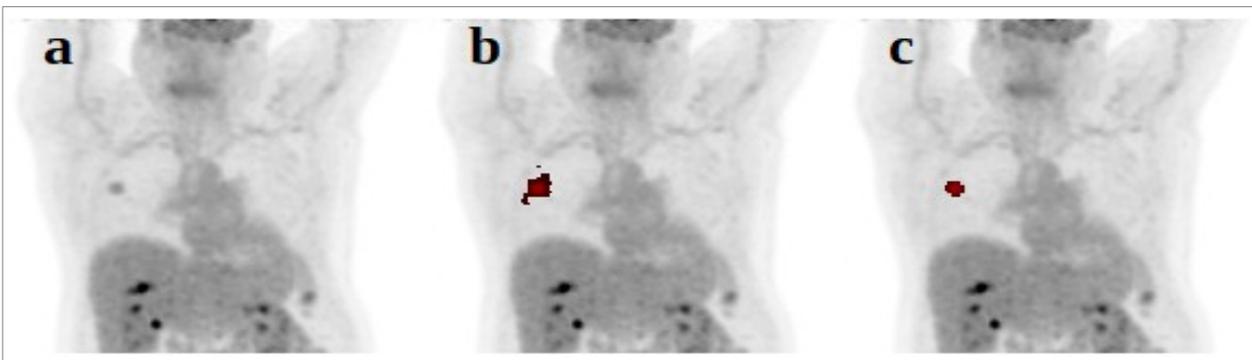


Abbildung 19: **MIP eines Patienten mit NSCLC im Stadium UICC III (Primärtumor im rechten Oberlappen).** Mit einem SUV_{max} von 3,57 SUV war die Aktivität des Primärtumors vergleichsweise gering (mittlere Leberaktivität: 3,06 SUV) (a). Eine Tumorsegmentierung mit dem niedrigen Schwellenwert von 20% des SUV_{max} führt zu einer fälschlichen Miterfassung von umgebendem Lungengewebe sowie Thoraxwand (b). Eine deutlich genauere Segmentierung wird mit dem Schwellenwert von 50% des SUV_{max} erreicht (c).

Ein quantitativer Zusammenhang zwischen dem Kontrast des Tumors (gemessen als Verhältnis der maximalen Tumorintensität zur mittleren Leberintensität) und dem minimal erforderlichen FT-Schwellenwert ist in Abb. 20 dargestellt. Unabhängig vom Kontrast ist nur selten ein absoluter Schwellenwert größer als 3,5 SUV für eine erfolgreiche Segmentierung erforderlich (Abb. 20

Resultate und Auswertung

links). Dieser Wert ist mit der mittleren Leberintensität im vorliegenden Datensatz vergleichbar (Median: 2,9 SUV, IQR: 2,5 - 3,5). Mit abnehmendem Tumorkontrast steigt der zur erfolgreichen Segmentierung erforderliche FT-Schwellenwert (Abb. 20 rechts). Liegt der SUV_{max} des Tumors im Bereich der mittleren Leberintensität (Tumor:Leber Ratio ~ 1), so ist im Trend ein relativer FT von $> 40\%$ des SUV_{max} erforderlich.

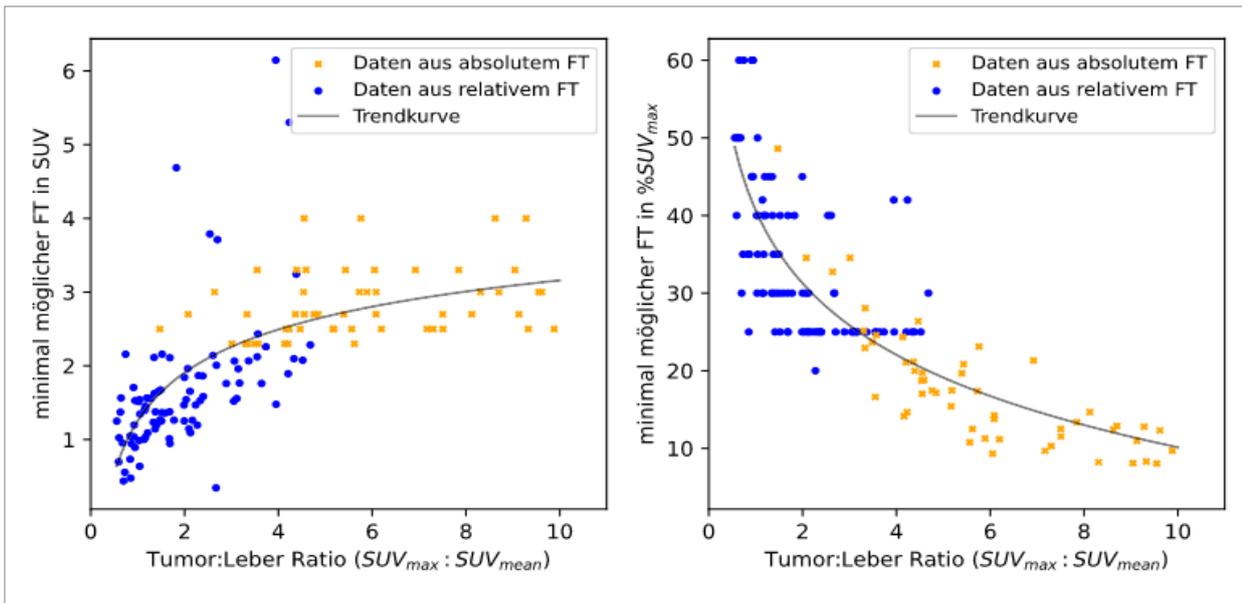


Abbildung 20: Minimal möglicher FT in Abhängigkeit des Verhältnisses von Tumorintensität (SUV_{max}) zu mittlerer Leberintensität (SUV_{mean}). Das Tumor-zu-Leber Verhältnis kann als Maß für den Kontrast des Tumors zur Umgebung interpretiert werden. Abgebildet sind 143 Fälle, für die der Übergang von nicht-erfolgreicher Segmentierung zu erfolgreicher Segmentierung, d.h. der erste adäquate Schwellenwert bekannt war. Dabei kommt es zur Diskretisierung eines Anteils der Daten entlang der Ordinatenachse, da nur eine begrenzte Zahl an Schwellenwerten (9 relative und 6 absolute) zur Verfügung standen um diesen Übergang zu identifizieren. Die linke Abbildung zeigt den Übergang in Absolutwerten (SUV), die rechte Abbildung zeigt ihn in Relation zum SUV_{max} .

4.2.3. Prognostische Validität - FT vs. iAT

Für jede Segmentierungsmethode erfolgte ein direkter Vergleich der adjustierten HR mit der iAT (angewendet auf diejenigen Fälle, für die eine erfolgreiche Abgrenzung möglich war). Abb. 21 und 22 zeigen die graphische Zusammenfassung für das DFS bzw. OS. Wie zuvor dargestellt, war eine erfolgreiche Segmentierung nicht für alle Methoden möglich. Dennoch sind die Ergebnisse aller Segmentierungsmethoden konsistent und legen ein höheres Risiko mit steigender ASP nahe.

Resultate und Auswertung

Es zeigt sich der Trend, dass das iAT Verfahren im Allgemeinen mit höher geschätzten Effektgrößen einhergeht. Dies ist insbesondere im Vergleich mit den absoluten FT-Methoden zu erkennen. Angewendet auf die für absolutes FT zugänglichen Fälle (z.B. N = 289 für 2,3 SUV) liegen die HR-Schätzer beim iAT sogar deutlich über den für die Gesamtkohorte bestimmten Werten. Der Effekt relativiert sich mit steigendem absoluten Schwellenwert und erreicht schließlich bei 4,0 SUV ein vergleichbares Niveau.

Die HR der relativen FT-Methoden in dem Bereich 30-50% SUV_{max} decken sich sehr gut mit den jeweiligen HR der iAT-Methode. Außerhalb dieses Bereichs, und ebenso für das visuelle Thresholding, nehmen die Abweichungen der Effektgrößen zu.

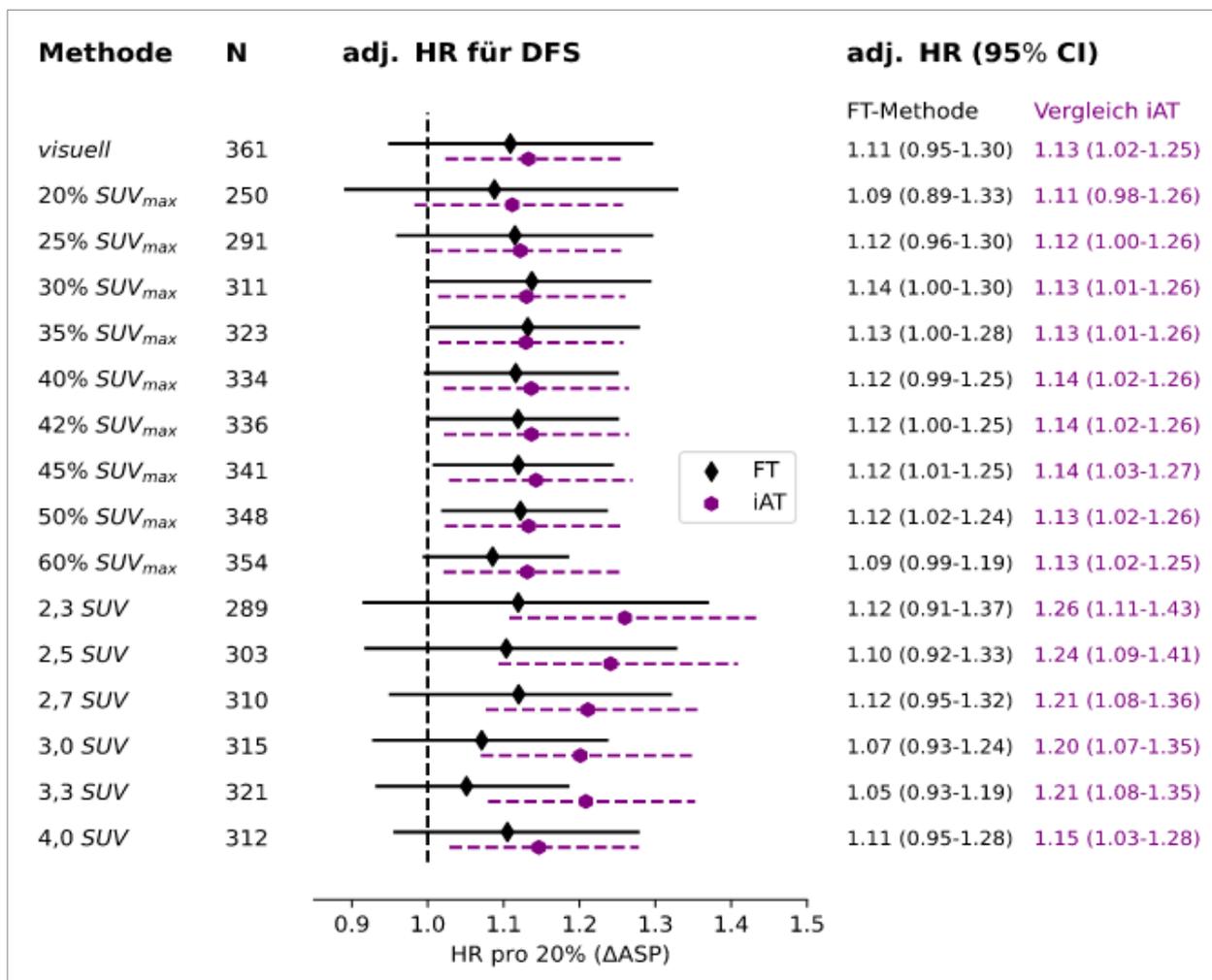


Abbildung 21: **HR für DFS - verschiedene Segmentierungsmethoden.** Dargestellt sind die Anzahl von Fällen (N), für die eine Segmentierung mit der jeweiligen FT-Methode möglich war. Zum Vergleich wurde auf diese Subgruppe auch das iAT angewendet (gestrichelt).

Resultate und Auswertung

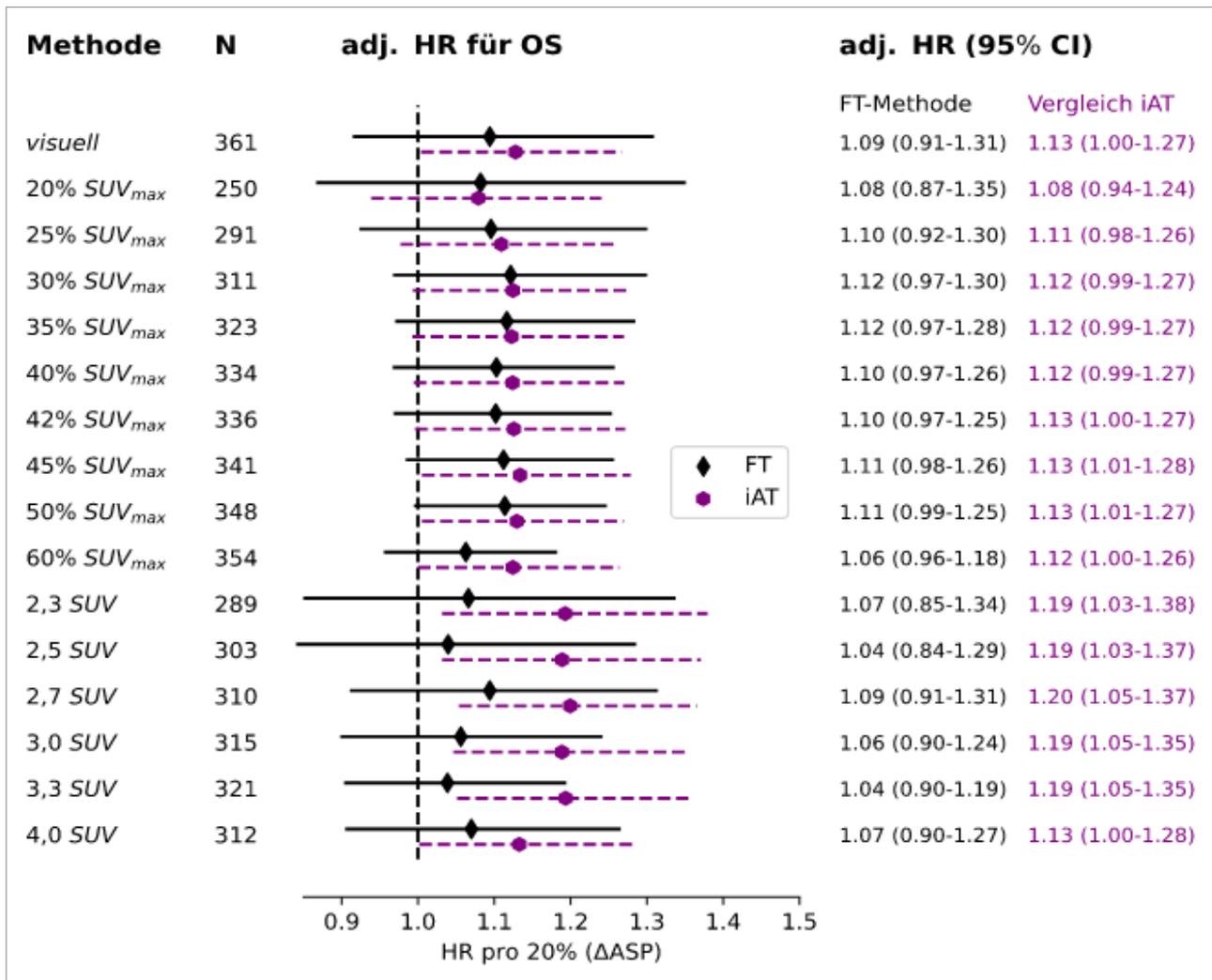


Abbildung 22: **HR für OS - verschiedene Segmentierungsmethoden.** In Analogie zu Abb. 21.

Aufgrund der auffälligen Unterschiede in den Effektgrößen für iAT zwischen der Gesamtkohorte und den Subgruppen, in denen sich der Tumor mit absolutem Thresholding abgrenzen ließ, erfolgte post-hoc eine Gegenüberstellung der Gruppen, in denen die Segmentierung mit FT = 2,3 SUV erfolgreich war (Succ. bzw. Successful) und die Gruppe aller anderen Patient*innen, in denen dies nicht der Fall war (Fail bzw. Failed). Das Niveau 2,3 SUV lag deutlich unterhalb der mittleren Leberintensität von 2,8 SUV. Wie Abb. 23 zeigt, wies die Gruppe der gescheiterten Segmentierungen signifikant höhere N-Deskriptoren, einen höheren R-Status, sowie höhere ASP, MTV und SUV_{max} auf.

Resultate und Auswertung

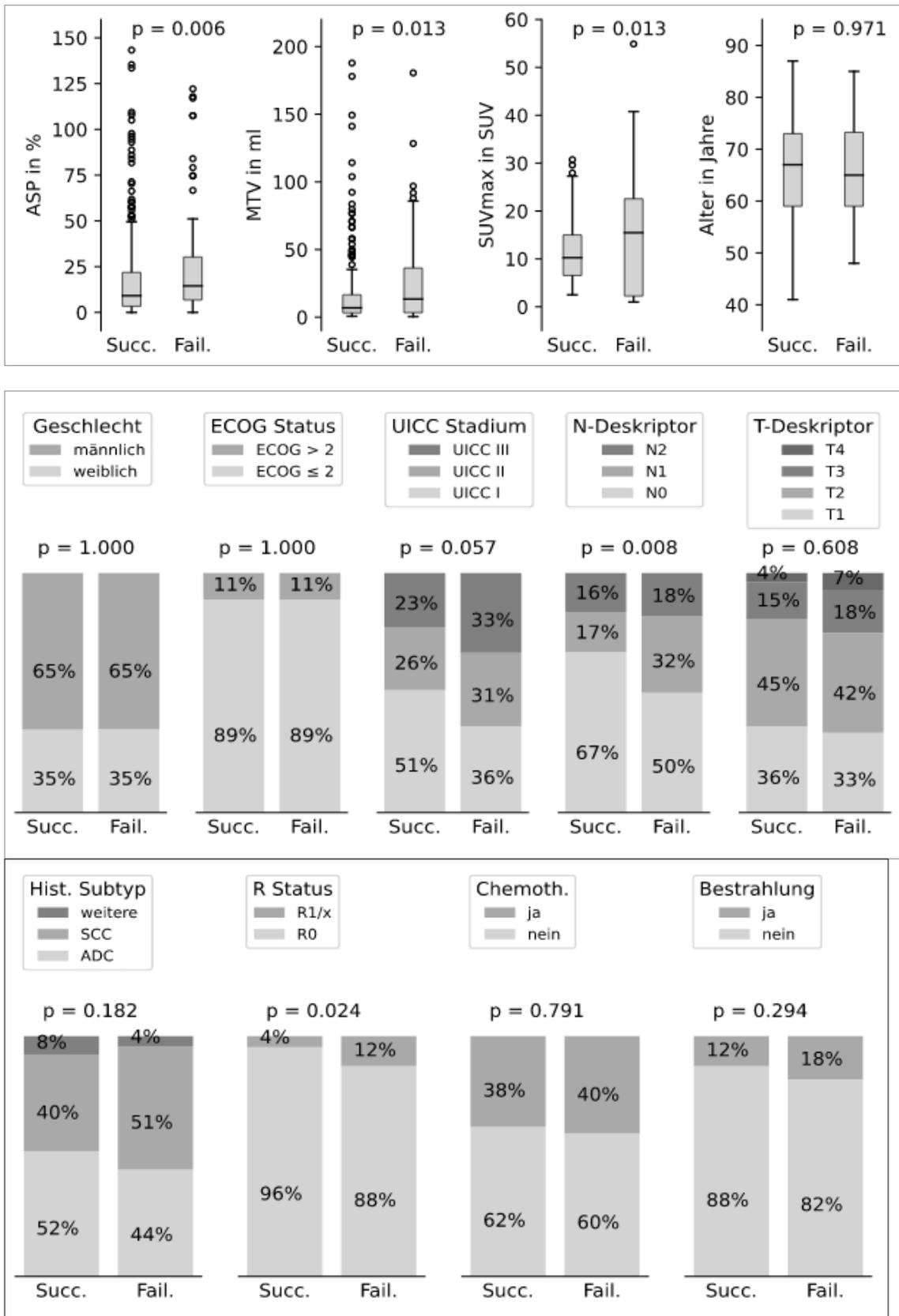


Abbildung 23 **Segmentierbarkeit und Assoziation zu untersuchten Variablen.** Vergleich der Gruppe, in denen die Segmentierung mit FT (Schwelle 2,3 SUV) erfolgreich (Succ. = Successful) vs. nicht erfolgreich war (Fail = Failed).

5. Diskussion

Formunregelmäßigkeiten eines Lungentumors wie z.B. Spikulae und zentrale Nekrosen sind Zeichen für Malignität und Aggressivität der Läsion und gehen im Allgemeinen mit einer ungünstigen Prognose einher (4,7,107,108). Die ASP des metabolischen Tumolvolumens stellt einen Marker für die Formunregelmäßigkeit des Tracer-Uptakes in der PET-Bildgebung dar. In dieser Arbeit konnte gezeigt werden, dass die ASP einen unabhängigen prognostischen Marker für das Überleben von Patient*innen mit NSCLC unter primär kurativer Behandlung darstellt. Hierzu erfolgte eine retrospektive Betrachtung von 361 Patient*innen in den Stadien UICC I-IIIa, die an der Charité-Universitätsmedizin Berlin in den Jahren 2009 bis 2017 in Behandlung waren. Sekundäres Ziel war es, die Stabilität der ASP und ihrer prognostischen Bedeutung in Bezug auf die Verwendung unterschiedlicher Segmentierungsalgorithmen zu untersuchen.

Ein Teil der Daten, die dieser Arbeit zugrunde liegen, wurde durch den Autor bereits im Rahmen einer Ko-Autorenschaft veröffentlicht (109). Hier wurde die prognostische Bedeutung der ASP insbesondere als binarisierte Variable gezeigt. Die Untersuchung schloss sich einer vorangegangenen Studie von Apostolova et al. aus dem Jahr 2014 an, in der eine unabhängige prognostische Aussagekraft der binarisierten ASP für das NSCLC gezeigt wurde. Dort stellte sich bei einer Fallzahl von 60 Patient*innen unter den PET-Parametern die ASP sowie der R-Status als prognostisch signifikant heraus (100). Der Einfluss verschiedener Segmentierungsalgorithmen wurde in beiden Publikationen nicht untersucht.

Neben der Betrachtung der Segmentierungsalgorithmen müssen vorab folgende methodische Unterschiede zu o.g. beiden Arbeiten aufgezeigt werden:

- i. In die Analyse in dieser Dissertationsschrift wurden auch Patient*innen mit kürzerem Follow-up als 12 Monate eingeschlossen, um einen Survivorship-Bias zu vermeiden.
- ii. Patient*innen im UICC-Stadium IIIB wurden aufgrund vergleichsweise niedriger Fallzahl bei gleichzeitig hoher therapeutischer Heterogenität ausgeschlossen.
- iii. Die Variablenselektion für die multivariable Analyse wurde nicht von den Ergebnissen der univariablen Cox-Regression abhängig gemacht, stattdessen erfolgte die Auswahl auf Grundlage der AIC sowie aus klinischen Überlegungen, da dieses Vorgehen im Allgemeinen ein robusteres Verfahren darstellt (110,111). Ferner wurde in dieser Arbeit auch in den Subgruppenanalysen eine Adjustierung vorgenommen.

- iv. Die ASP wurde als kontinuierliche Variable betrachtet und nicht binarisiert. Somit wird einerseits ein Informationsverlust vermieden, andererseits bleiben die Ergebnisse unabhängig von etwaig ausgewählten Schwellenwerten für die Binarisierung.

5.1. ASP als binarisierte vs. kontinuierliche Variable

Für die Binarisierung in Gruppen hoher und niedriger ASP fand sich in der Studie von Apostolova et al. ein optimaler Schwellenwert von 50,2% für das OS und 46,6% für das DFS. Diese Schwellenwerte wurden auf Grundlage eines Hazard-Ratio Plots abgeleitet, aus welchem der signifikanteste Cut-off identifiziert wurde. Wird diese Methode auf den dieser Arbeit zu Grunde liegenden Datensatz angewendet, dann liegt der Bereich prognostisch signifikanter Schwellenwerte bei 18% - offen (OS) bzw. 9-59% (DFS) und die optimalen Schwellenwerte betragen 33,3,% (OS) und 21,2% (DFS) (109). Die großen Spannweiten signifikanter Cutoffs unterstreicht die Robustheit der ASP als prognostischer Marker und ist konsistent mit dem in der Publikation von Apostolova et al. ermittelten, optimalen Cutoff. Dass sich allerdings die optimalen Cutoffs zwischen beiden Studien so deutlich unterscheiden (33,3% vs. 50,2% bzw. 21,2% vs. 46,6%) ist darin begründet, dass zu seiner Bestimmung aus einem Kontinuum an ASP-Werten derjenige Cutoff ermittelt wurde, der zur deutlichsten Trennung beider Gruppen im Sinne einer Minimierung des p-Werts im Log Rank Test führt. Kritikpunkte sind bei diesem Verfahren eine Alpha-Fehler-Kumulierung sowie auch der Umstand, dass die optimierten Werte sehr von der zugrunde liegenden Studienpopulation abhängen. So sind in der Studie von Apostolova et al. höhere Tumorstadien deutlich häufiger vertreten, wodurch, wie die vorliegende Arbeit gezeigt hat, eine Verschiebung der ASP-Verteilung in höhere Bereiche und damit auch höhere optimale Schwellenwerte zu erwarten sind. Darüber hinaus wurden in der aktuellen Studie zum Zweck der Homogenisierung ausschließlich Patient*innen mit resezierten Tumoren eingeschlossen, während dies in der Studie von Apostolova et al. kein Einschlusskriterium darstellte. Insgesamt besteht hier also das Risiko, dass ein optimierter Cutoff vielmehr die Risikoverteilung der zugrunde liegenden Kohorte widerspiegelt (im Sinne eines „Overfitting“), als die Existenz eines „echten“, biologisch legitimierbaren Cutoffs. Ein für eine Kohorte im frühen Stadium optimierter Cutoff wird sich allein aufgrund der unterschiedlichen ASP-Verteilung vermutlich nicht in gleicher Weise auf eine Kohorte übertragen lassen, in der ein späteres Tumorstadium dominiert. Darüber hinaus wird bei der Entscheidung zur Binarisierung bereits die möglicherweise ungerechtfertigte Annahme getroffen, dass sich die

Kohorte durch einen einzelnen Cutoff in zwei homogene Gruppen einteilen lässt. Dieses Vorgehen wird in der Literatur kontrovers diskutiert (112–114). In dieser Arbeit wurde daher auf eine Binarisierung der ASP verzichtet und die ASP stattdessen als kontinuierliche Variable beibehalten, wobei eine Schrittgröße (Δ ASP) von 20% verwendet wurde, um eine leichte Interpretation der berechneten Effektgrößen zu ermöglichen.

5.2. Prognostische Validität der ASP in der Gesamtkohorte

Über die gesamte Studienpopulation hinweg zeigte sich die ASP in der univariablen Cox-Regression als prognostisch signifikanter Parameter sowohl für das DFS als auch das OS, wobei höhere ASP-Werte mit einem jeweils höheren Risiko für ein Ereignis assoziiert waren. In der multivariablen Regression behielt die ASP nach Adjustierung für demographische Faktoren (Alter, Geschlecht, ECOG-Status), UICC-Stadium und R-Status ihren prognostischen Wert für das DFS und OS.

Für die Gesamtkohorte wurde dabei eine adjustierte HR von 1,13 (95% CI: 1,02-1,25; $p = 0,017$) für das DFS bzw. 1,13 (95% CI: 1,00-1,27; $p = 0,043$) für das OS pro Δ ASP von 20% ermittelt. Eine ASP von 0% entspricht der eines kompakten, runden Tumors. Eine ASP von 20% ist vereinbar mit einer deutlichen Spikulierung des Tumors (s. Abb. 2). Eine ASP von 50% liegt beispielsweise vor bei einem kugelförmigen Tumor mit zentraler, konzentrischer Nekrose, welche 1/6 des Tumolvolumens ausmacht. Nimmt die Nekrose bereits 1/3 des Tumors ein, so wächst die ASP auf $\sim 100\%$ an. Verglichen mit einem runden, kompakten Tumor (ASP = 0%) bedeutet eine ASP von $\sim 20\%$ somit rechnerisch eine HR von 1,13, eine ASP $\sim 50\%$ eine HR von 1,36, eine ASP $\sim 100\%$ eine HR von 1,84 und extreme ASP von $\sim 150\%$ eine HR von 2,50 in Bezug auf DFS und OS, bei sonst konstant gehaltenen demographischen Parametern, UICC-Stadium und R-Status. Dabei ist allerdings zu beachten, dass sich im vorliegenden Datensatz ein Großteil der Tumoren (70%) im niedrigeren ASP Bereich von 0-20% befindet. Nur 10% der Tumoren wiesen eine ASP $> 50\%$ auf, und nur 3% eine ASP $> 100\%$. Im Median zeigte sich in der Gesamtkohorte eine ASP von 10%. Somit erweist sich die ASP, abhängig davon, ab welcher Höhe ein Risiko als klinisch relevant eingestuft wird, nur für einen kleineren Anteil der Fälle als prognostisch bedeutsam. Dies schränkt die Sensitivität der ASP als prognostischen Marker ein.

Diskussion

Die bekannte prognostische Bedeutung anderer etablierter klinischer Parameter konnte in der aktuellen Arbeit reproduziert werden. Hiernach ist ein höheres Patient*innenalter, männliches Geschlecht, ein höherer ECOG-Status, höheres UICC-Stadium sowie ein R1/x-Status mit einem kürzeren Überleben assoziiert (115,116). Der histologische Subtyp zeigte sich in der univariablen Analyse signifikant prognostisch, wobei das Plattenepithelkarzinom im Vergleich zum Adenokarzinom mit einem signifikant höheren Risiko zu versterben einherging (HR 1,55; 95%CI: 1,09 - 2,19; $p = 0,014$). In der vorliegenden Arbeit war der histologische Subtyp allerdings nicht Teil der multivariablen Analyse. In der Literatur wird der Einfluss des histologischen Subtyps auf das Überleben kontrovers diskutiert, unter anderem auch bedingt durch das unterschiedliche Ansprechen auf Chemotherapeutika. (117,118).

5.3. Assoziation der ASP mit anderen Variablen

In dieser Arbeit zeigte sich eine Assoziation der ASP mit dem UICC-Stadium mit höheren medianen ASP-Werten in höheren Tumorstadien, wobei ein signifikanter Unterschied zwischen den Stadien UICC I und UICC II bestand ($p < 0,001$), hingegen nur ein geringfügiger Unterschied zwischen UICC II und UICC III.

Innerhalb der einzelnen UICC-Stadien gingen höhere ASP-Werte mit einem höheren T-Stadium einher. Hingegen war ein höheres N-Stadium mit niedrigerer ASP assoziiert. Ersteres hat v.a. einen technischen Grund. Wäre ein unlimitiertes räumliches Auflösungsvermögen realisierbar, so würden zwei gleich geformte Tumoren (d.h. im mathematischen Sinn: „ähnliche“ Körper) stets die gleiche ASP aufweisen, völlig unabhängig von ihrer Größe. Tatsächlich ist das theoretisch maximal erreichbare räumliche Auflösungsvermögen der PET aus physikalischen und praktischen Gründen auf ca. 2,5mm FWHM beschränkt (119) und gängige Systeme liegen derzeit mit ca. 4-8mm FWHM noch deutlich darüber. So auch das in dieser Studie verwendete PET/CT System (Philips Gemini TF System, FWHM: ca. 7mm). Dies führt bei kleinen Tumoren systematisch zu einer Unterschätzung der ASP, da kleine Formdetails schlicht nicht aufgelöst werden können bzw. durch den Partialvolumeneffekt verwaschen. Die Abhängigkeit eines gemessenen Umfangs (bzw. Oberfläche) eines Objekts von der gewählten Messgenauigkeit ist ein bereits aus dem Themenfeld der Fraktale, einem Teilgebiet der Geometrie, bekanntes sehr interessantes Phänomen und bereits ausführlich beschrieben worden (120). Allerdings ist auch unabhängig von der räumlichen Auflösung von einer Korrelation der ASP mit der Tumorgröße auszugehen, da mit zunehmender Tumorgröße (höheres T-

Diskussion

Stadium) Formunregelmäßigkeiten durch z.B. Asymmetrien im Wachstum, Bildung zentraler Nekrosen und Infiltration von Nachbarstrukturen wahrscheinlicher werden (121). Gestützt werden diese Überlegungen durch die beobachtete hochsignifikante, positive Korrelation der ASP mit dem MTV. Zusammenfassend lassen sich somit Informationen, die die Größe und die Form einer Läsion beschreiben, in der Realität nicht vollständig voneinander entkoppeln. In einer Phantomstudie von Limkin et al. 2018 konnte allerdings gezeigt werden, dass die Sphärizität (das mathematische Pendant zur ASP) im Vergleich zu anderen Form-beschreibenden Parametern eine relativ hohe Stabilität gegenüber Volumenveränderungen aufweist (122).

Die negative Assoziation von ASP und N-Stadium lässt sich hingegen nicht physikalisch begründen. Sie ist ein Artefakt der UICC-Klassifizierung: Eine Zuordnung in das Stadium UICC II setzt eine Klassifizierung in T2N0, T3N0, T1N1 oder T2N1 voraus (TNM Edition 7). Ein höheres T-Stadium (und damit einhergehend höhere ASP) erfordert also einen niedrigen N-Status und umgekehrt, um weiterhin dem UICC II zugeordnet werden zu können. Bei der vom UICC-Stadium unabhängigen Betrachtung von ASP und N-Stadium zeigt sich auch hier eine Zunahme der ASP mit steigendem N-Stadium. Die Resultate sind konsistent mit den Ergebnissen einer Untersuchung von Apostolova et al. 2016, in der eine ähnliche Assoziation der ASP mit dem T- und N-Stadium gesehen wurde (102). Hier wurde auch eine Assoziation der ASP mit tumorbiologischen Markern beim NSCLC gezeigt. Eine hohe ASP war mit einem höheren Ki-67 Proliferationsindex und einer höheren EGFR-Expression assoziiert und die Arbeitsgruppe schloss daraus einen möglicherweise sogar kausalen Zusammenhang zwischen Formunregelmäßigkeit und Proliferationsaktivität des Tumors. Der vorliegenden Arbeit standen keine immunohistochemischen Befunde zur Auswertung zur Verfügung, allerdings konnten signifikante Unterschiede in der Verteilung der ASP zwischen den histologischen Subtypen gezeigt werden, mit höheren Werten für das SCC im Vergleich zum ADC. SCC gehen bekanntermaßen mit einem höheren Ki-67 Proliferationsindex einher und weisen in der Bildgebung häufiger Nekrosen auf (123,124). Insgesamt ist die Evidenz für einen Zusammenhang zwischen Radiomics-Features und tumorbiologischen Markern noch sehr schwach und unzureichend validiert, wie La Greca Saint-Estevan et al. in einem systematischen Review 2021 darlegen konnten (125). Diese Unklarheiten betreffen auch die ASP. So findet sich entgegen der von Apostolova et al. beobachteten positiven Korrelation von ASP und EGFR-Expression in der Literatur auch eine Arbeit, in der im Falle des ADC das Vorliegen einer EGFR-Mutation mit einer *niedrigeren* ASP assoziiert war (126).

Diskussion

Patient*innen mit R1/x Status wiesen verglichen mit R0 resezierten Patient*innen ebenfalls signifikant höhere ASP-Werte auf ($p < 0,001$). Dies lässt sich eventuell dadurch erklären, dass eine höhere ASP mit größeren Tumolvolumina assoziiert ist, was eine vollständige Resektion erschweren könnte. Auf der anderen Seite wiesen alle Patient*innen mit einer ASP $< 5\%$ einen R0-Status auf.

Zusammenfassend besteht somit eine positive Assoziation der ASP mit Variablen, für die eine prognoseverschlechternde Bedeutung bereits bekannt ist. Dies erklärt die sehr deutlichen Unterschiede zwischen den geschätzten HR der multivariablen Analyse im Vergleich zur univariablen Analyse (HR 1,13 vs. 1,21 für das DFS bzw. 1,13 vs. 1,23 für das OS pro 20% Δ ASP) und unterstreicht die Notwendigkeit einer differenzierten Interpretation der ASP für spezifische Patient*innengruppen, um den tatsächlichen zusätzlichen prognostischen Wert zu extrahieren.

5.4. Prognostischer Wert der ASP nach Tumorstadium

Bei Stratifizierung der Studienpopulation nach UICC-Stadium zeigten sich deutliche Unterschiede der prognostischen Bedeutung der ASP innerhalb der einzelnen Gruppen. In den frühen Stadien wurden mit 1,19 (UICC I) bzw. 1,21 (UICC II) vergleichsweise hohe adjustierte HR für das DFS berechnet. In der Gruppe UICC IIIA fällt das HR mit 1,11 dagegen deutlich geringer aus. Für das OS zeigt sich ein stetiger Abfall des geschätzten HR von 1,29 (UICC I) über 1,13 (UICC II) auf schließlich 1,07 (UICC IIIA).

Bei Stratifizierung nach T/N-Stadien zeigten sich noch ausgeprägtere Unterschiede mit deutlich überdurchschnittlichen adjustierten HR in den frühen Stadien, d.h. in den Stadien ohne Lymphknotenbefall, aber auch im gepoolten Stadium T1N1/2. Das Pooling von T1N1- und T1N2 Fällen (d.h. Pooling von UICC IIB und UICC IIIA) zur Erhöhung der Fallzahl führte allerdings zu größerer Heterogenität mit resultierend großer Unsicherheit in der Schätzung, was an dieser Stelle angemerkt werden soll.

Das Stadium T1N0 entspricht gemäß UICC 7. Edition einem Tumordurchmesser < 3 cm und somit einem Tumolvolumen von etwa 10 ml. Im vorliegenden Datensatz wurde für diese Gruppe ein medianes MTV von 3 ml berechnet. Wie bereits zuvor erläutert, sind kleine MTV in der PET auch mit geringeren ASP-Werten assoziiert. Für T1N0 Tumore wurde eine mediane ASP von 6 % (IQR: 1- 10 %; Range: 0 - 86%) bestimmt, welches deutlich unterhalb dem Median der

Diskussion

Gesamtkohorte von 10% liegt. Obgleich die ASP-Werte dieser Tumoren nur auf ein kleines Spektrum begrenzt sind, wurden auch hier hohe Effektgrößen für die ASP geschätzt. Die Konfidenzintervalle zeigten sich zwar vergleichsweise groß, allerdings ist dies vielmehr durch die geringe Anzahl an Ereignissen und dem (für dieses prognostisch günstige Stadium) eher kurzen Follow-up Zeitraum zu erklären.

Zusammenfassend scheint der prognostische Wert der ASP besonders für die Stadien ohne Lymphknotenbefall sowie für das Stadium T1N1 relevant zu sein. Aus biologischer Sicht erscheint dies plausibel, da die ASP i.d.R. nur für den Primärtumor berechnet wurde (wenige Ausnahmen bestehen bei Primärtumoren, die in der PET mit Lymphknoten „verschmolzenen“ erschienen). Bei weiter fortgeschrittenem Tumorleiden mit ausgeprägtem Lymphknotenbefall werden letztere Umstände das Überleben vermutlich deutlich stärker beeinflussen als etwaige Formheterogenitäten des Primärtumors. Die vom Tumorstadium unabhängige prognostische Bedeutung des Lymphknotenbefalls bei R0-resezierten Patient*innen ist in der Literatur bereits gezeigt worden (127). Inwiefern die ASP des Primärtumors für fortgeschrittene Tumorstadien prognostische Information trägt, kann anhand des vorliegenden Datensatzes nicht beurteilt werden. Hierzu wäre eine gezielte Betrachtung der Stadien IIB, IIIA und IIIB erforderlich. Dies setzt allerdings eine hohe Fallzahl voraus, um für die ausgeprägte Heterogenität im Hinblick auf die T/N-Stadien (T1-4, N0-3) und Behandlungsstrategien, die mit diesen Stadien verbunden sind, adjustieren zu können (29).

5.5. Subgruppenanalyse: Chemotherapie

Etwa jede*r siebte Patient*in im Stadium UICC I (n=25), die Hälfte der Patient*innen im Stadium UICC II (n=49) und etwa zwei Drittel der Patient*innen im Stadium UICC IIIA (n=64) erhielt eine adjuvante oder neoadjuvante Chemotherapie. Bei isolierter Betrachtung dieser Gruppe zeigte die ASP im Vergleich zur Gesamtkohorte eine deutlich stärkere Assoziation mit dem Überleben mit adjustierten HR von 1,24 für das DFS und 1,23 für das OS pro 20 % Δ ASP. Die Subgruppe unterschied sich vom Rest der Gesamtkohorte durch größere Tumorumfänge (MTV im Median 12 ml vs. 6 ml), höhere Avidität (SUV_{max} im Median 13 SUV vs. 10 SUV), höhere Formunregelmäßigkeit (ASP im Median 13 % vs. 8 %) und fortgeschrittenere Tumorstadien (Anteil UICC IIIA: 46% vs. 12%; Anteil N_{1/2}-Stadium: 65% vs. 18%). Letzteres ist eine direkte Folge der Wahl dieser Subgruppe, da die klinische Indikation für eine Chemotherapie gerade in höheren Tumorstadien und Lymphknotenbefall gegeben ist. Demographische Unterschiede bestanden in

niedrigerem Alter (Median 62 vs. 70 Jahre) und besserem Allgemeinzustand (ECOG \geq 2: 3 % vs. 16 %) der Chemotherapiegruppe im Vergleich zum Rest der Gesamtkohorte.

Es ist vorstellbar, dass Patient*innen mit hoher ASP in dieser Subgruppe von einer intensivierten Therapie oder intensiveren Nachfolgeuntersuchungen profitieren könnten. Der Umkehrschluss, dass eine niedrige ASP bei Patient*innen im Stadium T2bN0, in denen eine Chemotherapie optional ist, die Entscheidung gegen eine solche erleichtern könnte, ist grundsätzlich denkbar, allerdings muss an dieser Stelle die relativ geringe Sensitivität der ASP angesprochen werden, die im vorangegangenen Abschnitt bereits diskutiert wurde. In jedem Fall ist ein solcher Nutzen der ASP für die klinische Entscheidungsfindung allein auf Basis dieser und vorangegangener retrospektiver Studien rein hypothetisch und müsste im Rahmen einer prospektiven Untersuchung bestätigt werden.

5.6. Assoziation der ASP mit anderen PET-Parametern

In der Literatur sind bereits seit längerem die SUV_{max} sowie das MTV als PET-basierte Prognosemarker für das NSCLC beschrieben. So konnten Liu et al. im Jahr 2016 in einer Metaanalyse über 36 Studien einen prognostischen Nutzen beider PET Parameter für das DFS und OS zeigen (62). Sie sind bisher allerdings nicht als unabhängige Faktoren neben den bekannten Prädiktoren in der Routineversorgung etabliert. Für die ASP liegen in der Literatur bisher nur wenige Daten zum Vergleich mit anderen prognostischen PET-Parametern vor. Während Apostolova et al. in ihrer Untersuchung einen klaren prognostischen Wert der ASP und keinen signifikanten Wert des MTV fanden, erwies sich das MTV in einer Studie von Hatt et al. 2017 der ASP überlegen. Beide Untersuchungen wiesen allerdings eine deutlich geringere Fallzahl mit 60 bzw. 87 Patient*innen und teilweise erhebliche Unterschiede zwischen den Studienpopulationen, insbesondere im Hinblick auf vertretene Tumorstadien und primäre Therapien, auf (100,128).

Tatsächlich konnte ein prognostischer Wert des MTV und des SUV_{max} auch in der vorliegenden Arbeit in Teilen reproduziert werden, und die Effekte erreichen zumindest in der univariablen Cox-Regression das formale Signifikanzniveau. Für das MTV ergab sich hierbei ein Risiko für DFS von 1,13 (95%CI: 1,03 - 1,23; $p = 0,007$) bei Wahl eines 30ml Inkrements. Für das OS konnte keine prognostische Bedeutung gezeigt werden (HR 1,10; 95%CI: 0,99 - 1,23; $p = 0,081$). Für den SUV_{max} berechnen sich mit der ASP vergleichbare adjustierte HR für DFS und OS bei Wahl eines

Diskussion

Inkrementen von 10 SUV, allerdings gehen die Schätzungen mit sehr hohen Konfidenzintervallen einher (HR 1,24 [95%CI: 1,02 - 1,50; $p = 0,031$] für das DFS bzw. HR 1,25 [95%CI: 1,01 - 1,55; $p = 0,043$] für das OS).

Es ist sehr wichtig anzumerken, dass die Parameter ASP, MTV und SUV_{max} zu hohem Grad paarweise miteinander positiv assoziiert sind. Dies wurde nicht nur in dieser Arbeit sondern bereits in anderen Publikationen zur quantitativen PET-Bildgebung beschrieben. Insbesondere die ASP und das MTV weisen eine hoch signifikante Korrelation miteinander auf (100,126,128). Eine Erklärung dafür wurde bereits in einem vorangegangenen Abschnitt versucht. Der Effekt nivelliert jedoch mit strengeren Segmentierungsmethoden, die zu einer Unterschätzung des MTV neigen (128). Die positive Korrelation von MTV und SUV_{max} begründet sich u.a. durch den abnehmenden Einfluss von Partialvolumeneffekt und Atembewegungen bei größeren Läsionen. Zudem ist eine positive Korrelation dieser beiden Größen mit der Aggressivität des Tumorwachstums anzunehmen (62).

Um Multikollinearität zu reduzieren und mögliche Effekt nicht zu unterschätzen, wurden das MTV und der SUV_{max} daher zur Beantwortung der zentralen Frage dieser Arbeit in der multivariablen Analyse ausgeschlossen. Dennoch erfolgte zusätzlich eine Untersuchung aller Kombinationen der PET-Parameter in der multivariablen Analyse, um ein Verständnis ihrer Interaktion miteinander zu gewinnen (vgl. Abb. 12). Dabei zeigte sich die ASP als äußerst robust für alle Kombinationen mit dem MTV und SUV_{max}. Im Gegensatz dazu zeigten sich die geschätzten Effektgrößen für das MTV sehr sensibel für Kombinationen mit der ASP. Bezüglich des OS kam es rechnerisch zu einer leichten Umkehr des Risikos. Dies lässt schlussfolgern, dass der prognostische Wert der ASP im vorliegenden Cox-Modell robuster nachweisbar und weniger anfällig für *Overfitting* ist. Das MTV weist im Vergleich zur ASP zudem eine stärkere Korrelation mit dem UICC-Stadium auf (Spearman $\rho = 0,377$ vs. $\rho = 0,332$, jeweils $p < 0,001$). Basierend auf den vorliegenden Daten dieser Arbeit erscheint die ASP als robusterer Marker, der im Vergleich zum MTV und SUV_{max} einen deutlicheren unabhängigen Wert in Verbindung mit einem klinischen Prognosemodell bietet.

Die Relevanz einer solchen Gegenüberstellung prognostischer Parameter ist einerseits für die klinische Praxis bedeutsam, da eine pragmatische Entscheidungsfindung auf Basis weniger, aber effektstarker und robuster Variablen im Gegensatz zu multiplen, unter Umständen redundanten, Variablen vorzuziehen ist. Andererseits führt auch die automatisierte, Machine-Learning basierte Verarbeitung von Bilddaten, etwa in Form einer Radiomics-Signatur, zum Problem des *Overfittings*, wenn die untersuchten Parameter untereinander höhergradig korreliert sind (44,59).

5.7. Segmentierungsmethode – Einfluss auf MTV und ASP

Grundlage der Berechnung der ASP ist die Segmentierung des MTV. In den vorangegangenen Untersuchungen dieser Arbeitsgruppe zur ASP wurde dafür ein von Hofheinz et al. entwickelter, iterativer, adaptiver Thresholding Algorithmus (iAT) verwendet (77).

Auch wenn das iAT ein exzellentes Interreader-Agreement aufweist (109), ist es nur eines unter einer Vielzahl verschiedener in der Literatur beschriebener und verwendeter Segmentierungsverfahren. Es ist zudem ein softwarespezifisches Verfahren. Demgegenüber sind Fixed Threshold (FT) Verfahren eine technisch simpel realisierbare und weit verfügbare Alternative, weshalb sie die am häufigsten in der Literatur verwendeten Verfahren darstellen. Allerdings besteht auch hier eine große Varianz im Hinblick auf die Wahl des optimalen Schwellenwerts. Die z.T. sehr hohe Variabilität quantitativer PET-Parameter in Abhängigkeit vom verwendeten Segmentierungsalgorithmus wurde bereits ausführlich in der Literatur thematisiert (58,61,74,78,86) und wird als ein großes Hindernis für die zentrumsübergreifende klinische Anwendung von PET-Radiomics wahrgenommen (61,129). Hatt et al. verglich drei fortgeschrittenere Segmentierungsalgorithmen sowie Fixed Thresholding mit Schwellen von 40% und 50% vom SUV_{max} miteinander. Zu ersteren gehören die Statistik-basierten Methoden *Fuzzy Locally Adaptive Bayesian* (FLAB), *Ant Colony Optimization* (ACO) sowie *Gradient-aided region-based active contour* (GARAC). Während ACO, FLAB und GARAC zu akkurateren Segmentierungen als Fixed Thresholding führten, ergab sich für die berechneten Sphärizitäten kein signifikanter Unterschied (128). Dediziertere Daten zur Variabilität der ASP im Hinblick auf verschiedene FT-Segmentierungen lagen bislang nicht vor.

Diese Arbeit konnte zeigen, dass FT mit niedrigem Schwellenwert im Allgemeinen zu einer Zunahme des MTV und einer Abnahme der Formunregelmäßigkeit der segmentierten Läsion, also zu einer Verschiebung der ASP hin zu kleineren Werten und vice versa führt. Für FT-Methoden mit Schwellen von 40-45% des SUV_{max} fand sich eine exzellente Übereinstimmung von MTV und ASP mit den Werten der iAT-Methode (ICC: $\geq 0,981$ für MTV bzw. ICC: $\geq 0,980$ für ASP). Demgegenüber weichen die Werte für absolutes FT deutlich ab. Die beste Übereinstimmung mit der iAT-Methode fand sich für 4,0 SUV (ICC: 0,644 für MTV bzw. ICC: 0,779 für ASP). Die Diskrepanz zwischen absolutem und relativem Thresholding lässt sich durch das Prinzip der iAT-Methode erklären, welche mathematisch näher mit dem relativen FT verwandt ist (77). Innerhalb der Gruppe der absoluten FT selbst zeigt sich wiederum eine gute Übereinstimmung der MTV und ASP (ICC zwischen 2,3 SUV und 4,0 SUV: 0,825 für MTV bzw. 0,851 für ASP).

Diskussion

Analog zu den Beobachtungen von Hatt et al. wurde auch in dieser Arbeit eine hohe Übereinstimmung der ASP über ein großes Spektrum verschiedener Segmentierungsmethoden hinweg gesehen, während dies für das MTV nicht der Fall war. Die systematische Überschätzung des MTV für niederschwelliges FT bzw. Unterschätzung des MTV für höherschwelliges FT ist erwartbar, da ein höherer Schwellenwert zu einem Ausschluss von Voxeln aus der ROI führt. Dagegen ist die Veränderung der ASP in Abhängigkeit von der Segmentierungsmethode deutlich weniger vorhersagbar. Zwar gehen auch hier höhere FT-Schwellen mit tendenziell höheren ASP einher, allerdings ist der Effekt deutlich schwächer ausgeprägt. Je nach Beschaffenheit der intraläsionalen Aktivitätsverteilung kann sich die ASP für ein und dieselbe Läsion in beide Richtungen verändern, wie Abb. 24 illustriert.

Für die Praxis relevant ist die Beobachtung, dass niedrigschwelliges absolutes und relatives Thresholding z.T. hohe Ausfallraten aufweisen. In jedem fünften und fast jedem dritten Fall scheiterte die Segmentierung mit Schwellenwerten von 2,3 SUV bzw. 20% SUV_{max} v.a. durch Miterfassung von umliegendem Gewebe. Unter den absoluten FT-Methoden zeigte sich als zuverlässigste Methode der Schwellenwert von 3,3 SUV, bei der es nur in 11% der Fälle zu inadäquaten Segmentierungen kam, die umfangreichere manuelle Nachkorrekturen erforderten. Mit dem relativen FT waren inadäquate Abgrenzungen nochmals seltener mit 7% der Fälle (für 40% SUV_{max}) bzw. < 4 % (für 50% und 60% SUV_{max}). In der Literatur wird über das Versagen einer Segmentierungsmethode nur selten berichtet.

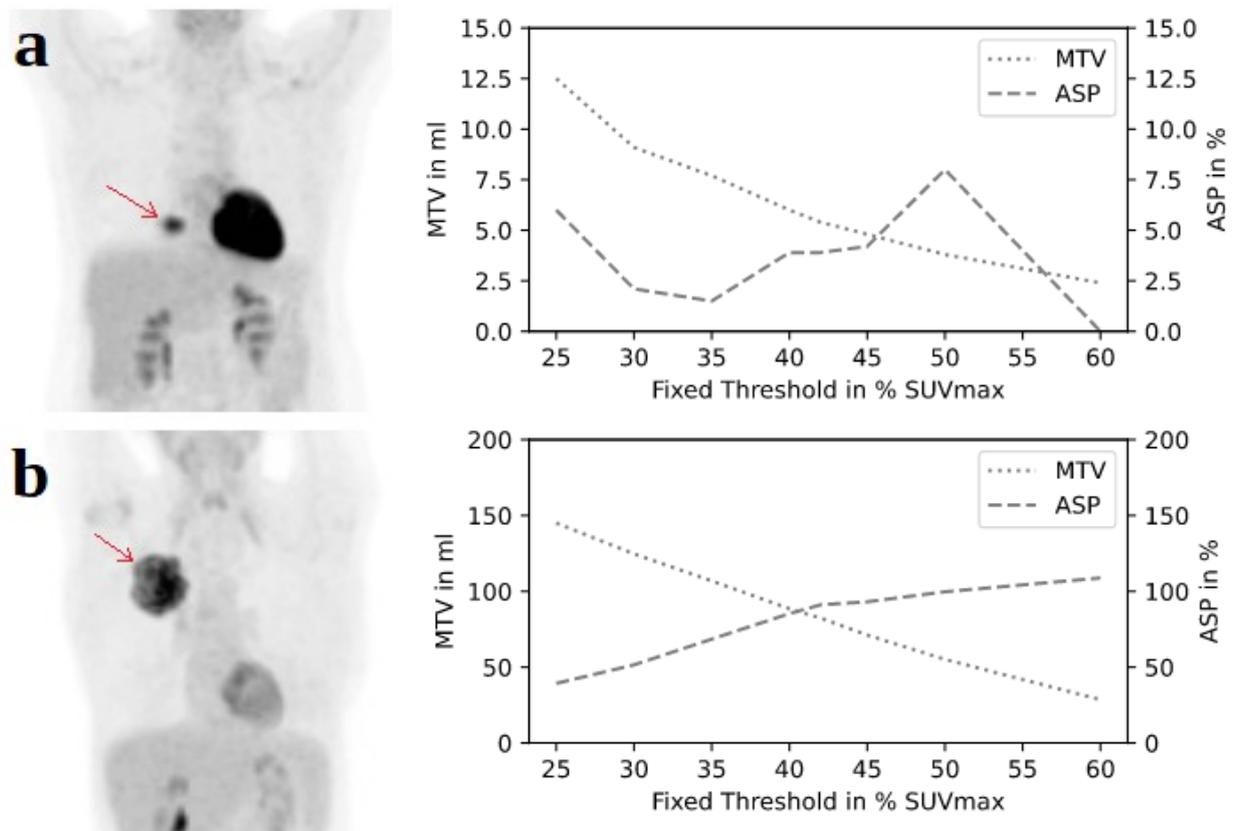


Abbildung 24: Unterschiedliche Ausprägung der ASP in Abhängigkeit vom FT-Schwellenwert. Illustriert am Beispiel eines kleinen homogenen (a) und eines großen heterogenen (b) Tumors (rote Pfeile). Während das MTV monoton mit zunehmendem FT-Schwellenwert abnimmt, ist die Veränderung der ASP abhängig von der Aktivitätsverteilung des Tumors. Bei sehr kleinen, homogenen Tumoren kommt es i.d.R. zu unvorhersehbaren, allerdings kleineren (s. Skala) Fluktuationen. Bei heterogenen Tumoren nimmt die ASP im Allgemeinen mit zunehmendem FT-Schwellenwert zu, kann allerdings bei sehr hohen FT-Werten mit dadurch unvollständiger Tumorerfassung auch wieder sinken (hier nicht abgebildet).

5.8. Segmentierungsmethode – Einfluss auf prognostischen Wert der ASP

Ogleich die ASP mit der verwendeten Segmentierungsmethode leicht variierte, zeigte sie sich im Hinblick auf ihre prognostische Bedeutung für alle untersuchten Thresholding-basierten Segmentierungsmethoden konsistent, wenn auch nicht in allen Fällen signifikant. Die höchste Übereinstimmung zum iAT fand sich für relatives Fixed Thresholding im Bereich 30 - 50 % SUV_{max} . Höhere oder niedrigere Schwellen, sowie auch absolutes Thresholding mit Schwellen unterhalb von 4,0 SUV führten zu größerer Variation der geschätzten HR. Für das OS

Diskussion

zeigten sich die HR bei absolutem Thresholding systematisch geringer als bei den Methoden 30 - 50% SUV_{max}.

Demgegenüber wurde von Hatt et al. (128) ein prognostischer Wert der ASP nur für zwei der fünf betrachteten Segmentierungsalgorithmen gesehen. Dabei handelte es sich um die komplexeren Verfahren FLAB und ACO. Hingegen wurde für GARAC sowie für die Fixed Thresholding Verfahren 40% und 50% SUV_{max} kein prognostischer Wert beobachtet. Die Autoren vermuteten, dass fortgeschrittenere Methoden eine präzisere Segmentierung von Tumoren mit komplexerer Form und heterogenem Uptake erlauben (128). Demgegenüber konnte in der aktuellen Arbeit gezeigt werden, dass die ASP sowohl für das iAT als auch für das FT mit Schwellen von 42 %, 45 %, und 50 % SUV_{max} sowie 3,3 SUV eine signifikante prognostische Wertigkeit für das DFS und OS aufweist. Die Diskrepanz könnte durch folgende Faktoren bedingt sein:

- i. Hatt et al. wählten ein strenges Signifikanzniveau von $\alpha = 0,01$ bei gleichzeitig geringer Fallzahl, wodurch ein Fehler 2. Art wahrscheinlicher wird (unzureichende Power). Tatsächlich berichteten die Autoren für die Verfahren 40% SUV_{max} und 50% SUV_{max} einen p-Wert von 0,029 bzw. 0,038 im Log-Rank-Test nach Binarisierung in hohe und niedrige ASP.
- ii. Es lag eine sehr unterschiedliche Verteilung der Tumorstadien vor. In der Studie von Hatt et al. waren ausschließlich Patient*innen im Stadium UICC II und III vertreten, wobei Letztere mit 66% die Population klar dominierten. Ein prognostischer Wert der ASP war gemäß der aktuellen Analyse hingegen insbesondere in den frühen Stadien relevant.
- iii. Während in der vorliegenden Arbeit nur Patient*innen mit kurativer Therapieintention einschließlich OP eingeschlossen wurden, lag der Anteil dieser Patient*innen in der Studie von Hatt et. al. bei nur 39%. Etwa jede*r siebte Patient*in wurde einer primär palliativen Chemotherapie zugeführt. Dies macht einen Vergleich der prognostischen Aussagekraft der beobachteten Parameter schwierig.
- iv. Die Methodik für die Berechnung der Tumoroberfläche unterlag einem wesentlichen Unterschied: Hatt et al. nutzten zur Abschätzung der Tumoroberfläche ein dem Radiomics-Standard (IBSI) folgendes Verfahren basierend auf einer Triangulation der MTV-Oberfläche (130). Dahingegen erfolgte in der aktuellen Untersuchung die Oberflächenschätzung durch Summation der Voxeloberflächen, die die innere und äußere Oberfläche darstellen, multipliziert mit dem Korrekturfaktor 2/3, wie durch Mullikin et al. 1993 beschrieben (131).

Diskussion

Die Verfahren können nicht als gleichwertig betrachtet werden, sodass grundsätzlich von relevanten Unterschieden in den berechneten Parametern auszugehen ist (132).

Allgemein liegt die Vermutung nahe, dass komplexere Segmentierungsalgorithmen, die eine präzisere Erfassung der metabolischen Tumorgrenzen und damit eine genauere ASP-Bestimmung ermöglichen, in der Folge auch das prognostische Potential der ASP besser ausschöpfen könnten. Nichtsdestotrotz eignet sich gemäß aktuellen Ergebnissen zur Segmentierung auch das relative FT mit Schwellenwerten von 30 - 50 %. Diese lassen eine HR von 1,12 – 1,14 pro 20 % Δ ASP für das DFS und 1,10 -1,12 für das OS erwarten.

Als interessante, wenn auch indirekte, Beobachtung zeigte sich in der vorliegenden Arbeit ein Anwachsen der Effektgrößen (HR), wenn die ASP_{iAT} ausschließlich auf „leicht segmentierbare“ (d.h. auch mit niederschwelligem FT sicher abgrenzbare) Tumoren angewendet wird. Betrachtet man ausschließlich die Fälle, die mit Fixed Thresholding von 2,3 SUV (liegt unterhalb der mittleren Leberintensität von 2,8 SUV) erfolgreich segmentierbar waren, so wies die Gruppe der gescheiterten Segmentierung im Vergleich zur Gruppe erfolgreicher Segmentierungen signifikant höhere N-Deskriptoren auf und gingen mit einem signifikant höheren Anteil des R1/x-Resektionsstatus einher. Vereinbar damit gestaltete sich in der Praxis die Segmentierung von Tumoren mit unmittelbarer Nähe zu physiologisch FDG-avidem Gewebe wie Brustwand, Zwerchfell, Perikard oder Mediastinum und zentralen Lymphknoten deutlich komplexer. Eine Abgrenzung mit niedrigen Schwellenwerten, die in den Bereich von physiologischem Gewebe fallen, ist dann ohne große und oft subjektive Korrekturen nicht möglich. Die Befunde sind vereinbar mit der zuvor gestellten Vermutung, dass sich der Effekt der ASP vornehmlich auf die frühen Stadien I-II beschränkt, und dass darüber hinaus der Effekt von Lymphknotenbefall und R-Status dominieren. Es soll jedoch festgehalten werden, dass diese Ergebnisse aus einer post-hoc Analyse resultieren und die vorliegende Arbeit nicht primär darauf abzielte, diesen Sachverhalt zu untersuchen.

5.9. Limitationen

Bei der vorliegenden Arbeit handelt es sich um eine retrospektive, monozentrische Studie mit einer aufgrund unterschiedlicher Therapien und Tumorstadien relativ heterogen zusammengesetzten Stichprobe. Aufgrund des Fehlens einer Kontrollgruppe sind alle beobachteten Effekte ausschließlich prognostischer Natur, hingegen können keine prädiktiven Aussagen getroffen werden, wodurch die therapeutische Konsequenz des Prognosemodells mit dieser Arbeit nicht beantwortet werden kann.

Es konnte gezeigt werden, dass signifikante Assoziationen der ASP mit etablierten Prognosemarkern bestehen. Allen voran steht hier das UICC Stadium, aber auch der R-Status und histologische Subtyp. Aufgrund dieser Multikollinearität ist grundsätzlich eine hohe Zahl von Fällen und Ereignissen erforderlich, um den adjustierten prognostischen Effekt im multivariablen Cox-Modell zu beweisen. Für die Subgruppen war dies nicht in jedem Fall gegeben. In der Gruppe der UICC I Patient*innen wurde das mediane OS nicht erreicht. Da die ASP gerade in den frühen Stadien UICC I und II ihren prognostischen Wert zu haben scheint, limitiert dieser Umstand die Power dieser Studie. Da es sich bei der vorliegenden Untersuchung allerdings um eine retrospektive Arbeit ohne formale Fallzahlplanung handelt, ist das Signifikanzniveau alleine nur eingeschränkt interpretierbar. Von Interesse sind unter diesem Gesichtspunkt vielmehr die beobachteten Effektgrößen und Konfidenzintervalle.

Die Effektgröße (HR) variiert erheblich innerhalb der einzelnen Subgruppen (UICC-Stadien, TNM-Stadien). So ist der Effekt pro 20% Δ ASP für kleine Tumoren (T₁-Stadium) besonders hoch, dies wird allerdings dadurch relativiert, dass nur ein sehr kleiner Anteil der Tumoren in diesem Stadium tatsächlich auch eine ASP von 20% oder höher aufweist. Für höhere Stadien nimmt dieser Anteil zu, gleichzeitig sinkt aber auch die geschätzte HR. Die mediane ASP der Kohorte liegt bei 10%, und mehr als 2/3 der Tumoren wiesen eine ASP im Bereich von 0-20% auf. Für die Praxis bedeutet dies eine vergleichsweise geringe Sensitivität der ASP als eigenständiger Prognosemarker.

Es muss betont werden, dass es sich bei der vorliegenden Arbeit um eine monozentrische Studie handelt. Viele quantitative Messungen in der PET sind stark von dem verwendeten Scanner, Untersuchungsprotokoll und der Bild-Prozessierung abhängig. In einer Studie, in der die ASP in einem simulierten multizentrischen Setting untersucht wurde, konnte gezeigt werden, dass die ASP deutlich von der rekonstruierten räumlichen Auflösung der PET-Bilddaten abhängt. Systematische Unterschiede zwischen Scannern bzw. Rekonstruktionsverfahren können zu diskordanten

Diskussion

Zuordnungen von Patient*innen in hohes vs. niedriges Risiko führen (57). Trotz dieser systematischen Unterschiede korrelierte die ASP zwischen den unterschiedlichen Bildrekonstruktionsverfahren sehr gut miteinander. Es kann somit davon ausgegangen werden, dass die ASP als kontinuierliche Variable ihren prognostischen Wert auch zwischen Zentren bewahrt. So könnte beispielsweise ein Effekt bei einer Schrittgröße Δ ASP von 20% in einem Zentrum äquivalent zu einer Schrittgröße von Δ ASP 15% oder 30% in einem anderen Zentrum sein. Allerdings ist für eine direkte Vergleichbarkeit der ASP zwischen Scannern eine stringente Harmonisierung hinsichtlich der rekonstruierten räumlichen Auflösung erforderlich.

Die untersuchten Algorithmen zur Segmentierung einer ROI beschränkten sich in dieser Arbeit auf das iAT und die vergleichsweise simplen FT-Methoden. Diese Verfahren wurden gewählt, da sie am weitesten verbreitet sind. Rand- oder regionsbasierte Verfahren könnten für die Segmentierung komplexer, heterogener Tumoren von zusätzlichem Nutzen sein. Nichtsdestotrotz ist der Umstand, dass das FT gleichermaßen wie das iAT zu ähnlichen Effektgrößen führte, ein Hinweis darauf, dass die ASP ein sehr robuster Parameter ist.

6. Schlussfolgerung

In dieser Arbeit konnte die ASP des Primärtumors in der [¹⁸F]FDG-PET als unabhängiger prognostischer Marker für das DFS und OS in einer Kohorte von 361 Patient*innen mit primär kurativ therapiertem NSCLC in den Stadien UICC I bis IIIA validiert werden. Demnach geht eine höhere ASP des Primärtumors mit einem signifikant höheren Risiko in Bezug auf Rezidiv, Progress oder Tod einher. Die Stratifizierung nach UICC-Stadien bzw. TNM-Stadien weist auf einen Effekt vor allem in den frühen Tumorstadien (UICC I-II bzw. N0-Status oder T1N1) hin. Wie diese Arbeit zeigt, kann eine undifferenzierte Anwendung der ASP über verschiedene Stadien hinweg zu hoher Variabilität in der prognostischen Wertigkeit der ASP führen. Eine besondere Subgruppe stellten diejenigen Patient*innen dar, die eine adjuvante oder neoadjuvante Chemotherapie erhielten. In dieser Gruppe zeigte sich der prognostische Wert der ASP besonders deutlich. Der Nutzen der ASP für die Entscheidung über eine Intensivierung von Therapie und Nachkontrolle ist hier im Rahmen prospektiver Studien zu untersuchen. Zusammenfassend unterstreichen die Ergebnisse dieser Arbeit die Notwendigkeit einer differenzierten Anwendung der ASP sowohl im klinischen als auch im wissenschaftlichen Kontext.

Für die Segmentierung des Primärtumors in der PET wurde in dieser Arbeit ein iterativer, Hintergrund-adaptiver Thresholding Algorithmus (iAT) verwendet, der bereits in den vorangegangenen Veröffentlichungen zur ASP Anwendung gefunden hat. Die Ergebnisse des zweiten Teils dieser Arbeit legen nahe, dass sich der prognostische Wert der ASP allerdings auch für weitere, gängigere Segmentierungsmethoden basierend auf v.a. relativem FT reproduzieren lässt. Zwar wurde im Gegensatz zum iAT für die meisten Segmentierungsmethoden das formale Signifikanzniveau nicht erreicht, die geschätzten Effektgrößen zeigten sich in ihrer Richtung allerdings durchweg konsistent. Hörschwelliges relatives Fixed Thresholding zeigte sich dabei besonders robust und in sehr guter Übereinstimmung mit den Ergebnissen des iAT. Dieser Umstand sowie auch Beobachtungen aus vorherigen Studien lassen die zu prüfende Hypothese aufstellen, dass eine strengere Wahl der Tumorgrenzen den prognostischen Wert der ASP verstärkt.

Die Wahl der Segmentierungsmethode beeinflusst, wie diese Arbeit zeigen konnte, maßgeblich das MTV und kann zu deutlichen Unterschieden in den resultierenden ASP-Verteilungen führen. Aus diesem Grund sollten ASP-Werte und daraus abgeleitete prognostische Informationen nicht ohne Kenntnis des Segmentierungsalgorithmus beurteilt werden. Für eine zentrumsübergreifende,

Schlussfolgerung

quantitative Betrachtung der ASP muss zunächst die Vergleichbarkeit der berechneten ASP-Werte sichergestellt werden, für die mitunter eine Standardisierung des Segmentierungsverfahrens erforderlich ist.

Literaturverzeichnis

1. Robert Koch-Institut (Hrsg) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. Krebs in Deutschland für 2015/2016. 12. Ausgabe. Berlin, 2019
2. Barta JA, Powell CA, Wisnivesky JP. Global Epidemiology of Lung Cancer. 2019;85(1):1-16.
3. Kraywinkel K, Schönfeld I. Epidemiologie des nichtkleinzelligen Lungenkarzinoms in Deutschland. Der Onkologe. 2018 Dec 6;24(12):946-51.
4. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF): Prävention, Diagnostik, Therapie und Nachsorge des Lungenkarzinoms, Langversion 1.0, 2018, AWMF-Registernummer: 020/007OL
5. Pallis AG, Syrigos KN. Lung cancer in never smokers: Disease characteristics and risk factors. Crit Rev Oncol Hematol. 2013 Dec;88(3):494–503.
6. Kanaji N, Watanabe N, Kita N, Bandoh S, Tadokoro A, Ishii T, Dobashi H , Matsunaga T. Paraneoplastic syndromes associated with lung cancer. World J Clin Oncol. 2014;5(3):197–223.
7. Hollings N, Shaw P. Diagnostic imaging of lung cancer. Eur Respir J. 2002 Apr;19(4):722–42. Available from: <http://erj.ersjournals.com/lookup/doi/10.1183/09031936.02.00280002>
8. van Tinteren H, Hoekstra OS, Smit EF, van den Bergh JH, Schreurs AJ, Stallaert RA, van Velthoven PC, Comans EF, Diepenhorst FW, Verboom P, van Mourik JC, Postmus PE, Boers M, Teule GJ. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. Lancet. 2002 Apr 20;359(9315):1388-93. doi: 10.1016/s0140-6736(02)08352-6.
9. Herder GJ, Kramer H, Hoekstra OS, Smit EF, Pruijm J, van Tinteren H, Comans EF, Verboom P, Uyl-de Groot CA, Welling A, Paul MA, Boers M, Postmus PE, Teule GJ, Groen HJ; POORT Study Group. Traditional versus up-front [18F] fluorodeoxyglucose-positron emission tomography staging of non-small-cell lung cancer: a Dutch cooperative randomized study. J Clin Oncol. 2006 Apr 20;24(12):1800-6. doi: 10.1200/JCO.2005.02.4695.
10. Brierley J, Gospodarowicz M, Wittekind C. UICC TNM classification of malignant tumours. Eighth edition. Chichester: Wiley; 2017.

Literaturverzeichnis

11. Greene FL, Sobin LH. The Staging of Cancer: A Retrospective and Prospective Appraisal. *CA Cancer J Clin.* 2008;58(3):180–90.
12. Gerbaudo VH, Julius B. Anatomic-metabolic characteristics of atelectasis in F-18 FDG-PET/CT imaging. *Eur J Radiol.* 2007 Dec;64(3):401–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17870274>
13. de Langen AJ, Raijmakers P, Riphagen I, Paul MA, Hoekstra OS. The size of mediastinal lymph nodes and its relation with metastatic involvement: a meta-analysis. *Eur J Cardio-Thoracic Surg.* 2006 Jan;29(1):26–9. Available from: <https://academic.oup.com/ejcts/article-lookup/doi/10.1016/j.ejcts.2005.10.002>
14. Pieterman RM, van Putten JW, Meuzelaar JJ, Mooyaart EL, Vaalburg W, Koëter GH, Fidler V, Pruijm J, Groen HJ. Preoperative staging of non-small-cell lung cancer with positron-emission tomography. *N Engl J Med.* 2000 Jul 27;343(4):254-61. doi: 10.1056/NEJM200007273430404.
15. Toloza EM, Harpole L, McCrory DC. Noninvasive staging of non-small cell lung cancer: a review of the current evidence. *Chest.* 2003 Jan;123(1 Suppl):137S-146S. doi: 10.1378/chest.123.1_suppl.137s.
16. Shim SS, Lee KS, Kim BT, Chung MJ, Lee EJ, Han J, Choi JY, Kwon OJ, Shim YM, Kim S. Non-small cell lung cancer: prospective comparison of integrated FDG PET/CT and CT alone for preoperative staging. *Radiology.* 2005 Sep;236(3):1011-9. doi: 10.1148/radiol.2363041310. Epub 2005 Jul 12.
17. Billé A, Pelosi E, Skanjeti A, Arena V, Errico L, Borasio P, Mancini M, Ardisson F. Preoperative intrathoracic lymph node staging in patients with non-small-cell lung cancer: accuracy of integrated positron emission tomography and computed tomography. *Eur J Cardiothorac Surg.* 2009 Sep;36(3):440-5. doi: 10.1016/j.ejcts.2009.04.003. Epub 2009 May 22.
18. Qu X, Huang X, Yan W, Wu L, Dai K. A meta-analysis of ¹⁸FDG-PET-CT, ¹⁸FDG-PET, MRI and bone scintigraphy for diagnosis of bone metastases in patients with lung cancer. *Eur J Radiol.* 2012 May;81(5):1007-15. doi: 10.1016/j.ejrad.2011.01.126. Epub 2011 Feb 26.
19. Wu Q, Luo W, Zhao Y, Xu F, Zhou Q. The utility of ¹⁸F-FDG PET/CT for the diagnosis of adrenal metastasis in lung cancer. *Nucl Med Commun.* 2017 Dec;38(12):1117–24. Available from: <https://journals.lww.com/00006231-201712000-00015>

Literaturverzeichnis

20. Nia ES, Garland LL, Eshghi N, Nia BB, Avery RJ, Kuo PH. Incidence of Brain Metastases on Follow-up 18 F-FDG PET/CT Scans of Non-Small Cell Lung Cancer Patients: Should We Include the Brain? *J Nucl Med Technol.* 2017 Sep;45(3):193–7. Available from: <http://tech.snmjournals.org/lookup/doi/10.2967/jnmt.117.194571>
21. Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The Eighth Edition Lung Cancer Stage Classification. *Chest.* 2017;151(1):193–203.
22. Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, Escriu C, Peters S; ESMO Guidelines Committee. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2017 Jul 1;28(suppl_4):iv1-iv21. doi: 10.1093/annonc/mdx222. PMID: 28881918.
23. Jones DR, Daniel TM, Denlinger CE, Rundall BK, Smolkin ME, Wick MR. Stage IB Nonsmall Cell Lung Cancers: Are They All the Same? *Ann Thorac Surg.* 2006 Jun;81(6):1958–62. < <https://linkinghub.elsevier.com/retrieve/pii/S0003497505023581>>
24. Strauss GM, Herndon JE 2nd, Maddaus MA, Johnstone DW, Johnson EA, Harpole DH, Gillenwater HH, Watson DM, Sugarbaker DJ, Schilsky RL, Vokes EE, Green MR. Adjuvant paclitaxel plus carboplatin compared with observation in stage IB non-small-cell lung cancer: CALGB 9633 with the Cancer and Leukemia Group B, Radiation Therapy Oncology Group, and North Central Cancer Treatment Group Study Groups. *J Clin Oncol.* 2008 Nov 1;26(31):5043-51. doi: 10.1200/JCO.2008.16.4855.
25. Mizuno T, Ishii G, Nagai K, Yoshida J, Nishimura M, Mochizuki T, Kawai O, Hasebe T, Ochiai A. Identification of a low risk subgroup of stage IB lung adenocarcinoma patients. *Lung Cancer.* 2008 Dec;62(3):302-8. doi: 10.1016/j.lungcan.2008.03.032.
26. Tandberg DJ, Tong BC, Ackerson BG, Kelsey CR. Surgery versus stereotactic body radiation therapy for stage I non-small cell lung cancer: A comprehensive review. *Cancer.* 2018 Feb 15;124(4):667–78. Available from: <http://doi.wiley.com/10.1002/cncr.31196>
27. Douillard JY, Tribodet H, Aubert D, Shepherd FA, Rosell R, Ding K, Veillard AS, Seymour L, Le Chevalier T, Spiro S, Stephens R, Pignon JP; LACE Collaborative Group. Adjuvant cisplatin and vinorelbine for completely resected non-small cell lung cancer: subgroup analysis of the Lung Adjuvant Cisplatin Evaluation. *J Thorac Oncol.* 2010 Feb;5(2):220-8. doi: 10.1097/JTO.0b013e3181c814e7.

Literaturverzeichnis

28. Lim E, Harris G, Patel A, Adachi I, Edmonds L, Song F. Preoperative versus Postoperative Chemotherapy in Patients with Resectable Non-small Cell Lung Cancer: Systematic Review and Indirect Comparison Meta-Analysis of Randomized Trials. *J Thorac Oncol*. 2009 Nov;4(11):1380–8.
29. Tanner NT, Gomez M, Rainwater C, Nietert PJ, Simon GR, Green MR, Silvestri GA. Physician preferences for management of patients with stage IIIA NSCLC: impact of bulk of nodal disease on therapy selection. *J Thorac Oncol*. 2012 Feb;7(2):365-9. doi: 10.1097/JTO.0b013e31823a385f.
30. Chansky K, Sculier J-P, Crowley JJ, Giroux D, Van Meerbeeck J, Goldstraw P. The International Association for the Study of Lung Cancer Staging Project: Prognostic Factors and Pathologic TNM Stage in Surgically Managed Non-small Cell Lung Cancer. *J Thorac Oncol* . 2009 Jul;4(7):792–801.
31. Woodard GA, Jones KD, Jablons DM. Lung Cancer Staging and Prognosis. 2016;47–75. Available from: http://link.springer.com/10.1007/978-3-319-40389-2_3
32. Saffarzadeh AG, Blasberg JD. Beyond TNM: Searching for New Patient-Centric Prognostic Indicators in NSCLC. *Ann Surg Oncol*. 2018 Nov 28;25(12):3425–6. Available from: <http://link.springer.com/10.1245/s10434-018-6714-0>
33. Yang S-R, Schultheis AM, Yu H, Mandelker D, Ladanyi M, Büttner R. Precision medicine in non-small cell lung cancer: Current applications and future directions. *Semin Cancer Biol*. 2020 Jul; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1044579X20301644>
34. Shui L, Ren H, Yang X, Li J, Chen Z, Yi C, Zhu H, Shui P. The Era of Radiogenomics in Precision Medicine: An Emerging Approach to Support Diagnosis, Treatment Decisions, and Prognostication in Oncology. *Front Oncol*. 2021 Jan 26;10:570465. doi: 10.3389/fonc.2020.570465.
35. Gester F, Paulus A, Sibille A, Corhay JL, Duysinx B, Louis R. Facteurs pronostiques du cancer pulmonaire non à petites cellules. *Rev Med Liege*. 2016;71(1):34–9.
36. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, Geisinger K, Hirsch FR, Ishikawa Y, Kerr KM, Noguchi M, Pelosi G, Powell CA, Tsao MS, Wistuba I; WHO Panel. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since

Literaturverzeichnis

- the 2004 Classification. *J Thorac Oncol.* 2015 Sep;10(9):1243-1260. doi: 10.1097/JTO.0000000000000630. PMID: 26291008.
37. Thakur M, Gadgeel S. Predictive and Prognostic Biomarkers in Non-Small Cell Lung Cancer. *Semin Respir Crit Care Med* [Internet]. 2016 Oct 12;37(05):760–70. Available from: <http://www.thieme-connect.de/DOI/DOI?10.1055/s-0036-1592337>
38. Shen H, Li Y, Liao Y, Zhang T, Liu Q, Du J. Lower Blood Calcium Associates with Unfavorable Prognosis and Predicts for Bone Metastasis in NSCLC. Yang P-C, editor. *PLoS One* . 2012 Mar 30;7(3):e34264.
39. Zhang SM, Zhu QG, Ding XX, Lin S, Zhao J, Guan L, Li T, He B, Zhang HQ. Prognostic value of EGFR and KRAS in resected non-small cell lung cancer: a systematic review and meta-analysis. *Cancer Manag Res.* 2018 Sep 10;10:3393-3404. doi: 10.2147/CMAR.S167578.
40. Bianconi F, Fravolini ML, Bello-Cerezo R, Minestrini M, Scialpi M, Palumbo B. Evaluation of Shape and Textural Features from CT as Prognostic Biomarkers in Non-small Cell Lung Cancer. *Anticancer Res.* 2018 Apr;38(4):2155-2160. doi: 10.21873/anticancerres.12456.
41. Im HJ, Pak K, Cheon GJ, Kang KW, Kim SJ, Kim IJ, Chung JK, Kim EE, Lee DS. Prognostic value of volumetric parameters of (18)F-FDG PET in non-small-cell lung cancer: a meta-analysis. *Eur J Nucl Med Mol Imaging.* 2015 Feb;42(2):241-51. doi: 10.1007/s00259-014-2903-7.
42. Baek S, He Y, Allen BG, Buatti JM, Smith BJ, Tong L, Sun Z, Wu J, Diehn M, Loo BW, Plichta KA, Seyedin SN, Gannon M, Cabel KR, Kim Y, Wu X. Deep segmentation networks predict survival of non-small cell lung cancer. *Sci Rep.* 2019 Nov 21;9(1):17286. doi: 10.1038/s41598-019-53461-2.
43. Kourou K, Exarchos TP, Exarchos KP, Karamouzis M V., Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J.* 2015;13:8–17. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2001037014000464>
44. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, Bellomi M. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp.* 2018 Nov 14;2(1):36. doi: 10.1186/s41747-018-0068-z.

Literaturverzeichnis

45. Lee G, Lee HY, Park H, Schiebler ML, van Beek EJR, Ohno Y, Seo JB, Leung A. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. *Eur J Radiol*. 2017 Jan;86:297-307. doi: 10.1016/j.ejrad.2016.09.005.
46. Shi L, He Y, Yuan Z, Benedict S, Valicenti R, Qiu J, Rong Y. Radiomics for Response and Outcome Assessment for Non-Small Cell Lung Cancer. *Technol Cancer Res Treat*. 2018 Jan 1;17:1533033818782788. doi: 10.1177/1533033818782788.
47. Ancey P, Contat C, Meylan E. Glucose transporters in cancer – from tumor cells to the tumor microenvironment. *FEBS J*. 2018 Aug 25;285(16):2926–43. Available from: <https://onlinelibrary.wiley.com/doi/10.1111/febs.14577>
48. Basu S, Kwee TC, Surti S, Akin EA, Yoo D, Alavi A. Fundamentals of PET and PET/CT imaging. *Ann N Y Acad Sci*. 2011;1228(1):1–18.
49. Basu S, Hess S, Nielsen Braad P-E, Olsen BB, Inglev S, Høilund-Carlsen PF. The Basic Principles of FDG-PET/CT Imaging. *PET Clin* 2014 Oct;9(4):355–70. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1556859814000716>
50. Cherry SR, Sorenson J, Phelps ME, Methé BM. Physics in Nuclear Medicine. *Med Phys*. 2004 Aug;31(8):2370–1. Available from: <http://doi.wiley.com/10.1118/1.1776595>
51. Lee TC, Alessio AM, Miyaoka RM, Kinahan PE. Morphology supporting function: attenuation correction for SPECT/CT, PET/CT, and PET/MR imaging. *Q J Nucl Med Mol Imaging* . 2016 Mar;60(1):25–39.
52. Schaart DR. Physics and technology of time-of-flight PET detectors. *Phys Med Biol*. 2021 Apr 21;66(9). doi: 10.1088/1361-6560/abee56.
53. Townsend DW. Physical principles and technology of clinical PET imaging. *Ann Acad Med Singapore* . 2004 Mar;33(2):133–45.
54. Akamatsu G, Ishikawa K, Mitsumoto K, Taniguchi T, Ohya N, Baba S, Abe K, Sasaki M. Improvement in PET/CT image quality with a combination of point-spread function and time-of-flight in relation to reconstruction parameters. *J Nucl Med*. 2012 Nov;53(11):1716-22. doi: 10.2967/jnumed.112.103861.
55. Rogasch JM, Steffen IG, Hofheinz F, Großer OS, Furth C, Mohnike K, Hass P, Walke M, Apostolova I, Amthauer H. The association of tumor-to-background ratios and SUVmax deviations related to point spread function and time-of-flight F18-FDG-PET/CT

Literaturverzeichnis

- reconstruction in colorectal liver metastases. *EJNMMI Res.* 2015 May 6;5:31. doi: 10.1186/s13550-015-0111-5.
56. Huang K, Feng Y, Liang W, Li L. Impact of time of flight and point spread function on quantitative parameters of lung lesions in 18F-FDG PET/CT. *BMC Med Imaging.* 2021 Dec 13;21(1):169.
 57. Rogasch JMM, Furth C, Bluemel S, Radojewski P, Amthauer H, Hofheinz F. Asphericity of tumor FDG uptake in non-small cell lung cancer: reproducibility and implications for harmonization in multicenter studies. *EJNMMI Res.* 2020 Dec 2;10(1):134.
 58. Cheebsumon P, Yaqub M, Van Velden FHP, Hoekstra OS, Lammertsma AA, Boellaard R. Impact of [18F]FDG PET imaging parameters on automatic tumour delineation: Need for improved tumour delineation methodology. *Eur J Nucl Med Mol Imaging.* 2011;38(12):2136–44.
 59. Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present... any future? *Eur J Nucl Med Mol Imaging .* 2017 Jan;44(1):151–65.
 60. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J Nucl Med.* 2004 Sep;45(9):1519-27.
 61. Foster B, Bagci U, Mansoor A, Xu Z, Mollura DJ. A review on segmentation of positron emission tomography images. *Comput Biol Med.* 2014 Jul;50(1):76–96.
 62. Liu J, Dong M, Sun X, Li W, Xing L, Yu J. Prognostic value of 18F-FDG PET/CT in surgical non-small cell lung cancer: A meta-analysis. *PLoS One.* 2016;11(1):1–19.
 63. Yamaguchi O, Kaira K, Hashimoto K, Mouri A, Shiono A, Miura Y, Murayama Y, Kobayashi K, Kagamu H, Kuji I. Tumor metabolic volume by 18F-FDG-PET as a prognostic predictor of first-line pembrolizumab for NSCLC patients with PD-L1 \geq 50. *Sci Rep.* 2020 Sep 14;10(1):14990. doi: 10.1038/s41598-020-71735-y.
 64. Li X, Wang D, Yu L. Prognostic and Predictive Values of Metabolic Parameters of 18F-FDG PET/CT in Patients With Non-Small Cell Lung Cancer Treated With Chemotherapy. *Mol Imaging.* 2019;18(6):1–12.

Literaturverzeichnis

65. Vorwerk H, Beckmann G, Bremer M, Degen M, Dietl B, Fietkau R, Gsänger T, Hermann RM, Alfred Herrmann MK, Höller U, van Kampen M, Körber W, Maier B, Martin T, Metz M, Richter R, Siekmeyer B, Steder M, Wagner D, Hess CF, Weiss E, Christiansen H. The delineation of target volumes for radiotherapy of lung cancer patients. *Radiother Oncol.* 2009 Jun;91(3):455-60. doi: 10.1016/j.radonc.2009.03.014.
66. Pfaehler E, Burggraaff C, Kramer G, Zijlstra J, Hoekstra OS, Jalving M, Noordzij W, Brouwers AH, Stevenson MG, de Jong J, Boellaard R. PET segmentation of bulky tumors: Strategies and workflows to improve inter-observer variability. *PLoS One.* 2020 Mar 30;15(3):e0230901. doi: 10.1371/journal.pone.0230901.
67. Konert T, Vogel W, MacManus MP, Nestle U, Belderbos J, Grégoire V, Thorwarth D, Fidarova E, Paez D, Chiti A, Hanna GG. PET/CT imaging for target volume delineation in curative intent radiotherapy of non-small cell lung cancer: IAEA consensus report 2014. *Radiother Oncol.* 2015 Jul;116(1):27-34. doi: 10.1016/j.radonc.2015.03.014.
68. Nestle U, De Ruyscher D, Ricardi U, Geets X, Belderbos J, Pöttgen C, Dziaduszek R, Peeters S, Lievens Y, Hurkmans C, Slotman B, Ramella S, Faivre-Finn C, McDonald F, Manapov F, Putora PM, LePéchoux C, Van Houtte P. ESTRO ACROP guidelines for target volume definition in the treatment of locally advanced non-small cell lung cancer. *Radiother Oncol.* 2018 Apr;127(1):1-5. doi: 10.1016/j.radonc.2018.02.023.
69. Konert T, Vogel WV, Everitt S, MacManus MP, Thorwarth D, Fidarova E, Paez D, Sonke JJ, Hanna GG. Multiple training interventions significantly improve reproducibility of PET/CT-based lung cancer radiotherapy target volume delineation using an IAEA study protocol. *Radiother Oncol.* 2016 Oct;121(1):39-45. doi: 10.1016/j.radonc.2016.09.002.
70. Paulino AC, Johnstone PAS. FDG-PET in radiotherapy treatment planning: Pandora's box? *Int J Radiat Oncol Biol Phys.* 2004 May 1;59(1):4-5.
71. Hellwig D, Graeter TP, Ukena D, Groeschel A, Sybrecht GW, Schaefers HJ, Kirsch CM. 18F-FDG PET for mediastinal staging of lung cancer: which SUV threshold makes sense? *J Nucl Med.* 2007 Nov;48(11):1761-6. doi: 10.2967/jnumed.107.044362.

Literaturverzeichnis

72. Oven Ustaalioglu BB, Gumus M, Bilici A, Seker M, Ustaalioglu R, Kefeli U, Salepci T, Canpolat N, Mayadagli A, Dogusoy I. Is there a cut-off value for standardized uptake values in positron emission tomography for predicting response to treatment and survival in patients with advanced non-small cell lung cancer? Single center experience. *J BUON*. 2010 Jul-Sep;15(3):529-36. PMID: 20941823.
73. Hong R, Halama J, Bova D, Sethi A, Emami B. Correlation of PET standard uptake value and CT window-level thresholds for target delineation in CT-based radiation treatment planning. *Int J Radiat Oncol Biol Phys* [Internet]. 2007 Mar 1;67(3):720–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17293230>
74. Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, Kirsch CM. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-Small cell lung cancer. *J Nucl Med*. 2005 Aug;46(8):1342-8. PMID: 16085592.
75. Biehl KJ, Kong FM, Dehdashti F, Jin JY, Mutic S, El Naqa I, Siegel BA, Bradley JD. 18F-FDG PET definition of gross tumor volume for radiotherapy of non-small cell lung cancer: is a single standardized uptake value threshold approach appropriate? *J Nucl Med*. 2006 Nov;47(11):1808-12. PMID: 17079814.
76. Deniaud-Alexandre E, Touboul E, Lerouge D, Grahek D, Foulquier JN, Petegnief Y, Grès B, El Balaa H, Keraudy K, Kerrou K, Montravers F, Milleron B, Lebeau B, Talbot JN. Impact of computed tomography and 18F-deoxyglucose coincidence detection emission tomography image fusion for optimization of conformal radiotherapy in non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2005 Dec 1;63(5):1432-41. doi: 10.1016/j.ijrobp.2005.05.016.
77. Hofheinz F, Langner J, Petr J, Beuthien-Baumann B, Steinbach J, Kotzerke J, van den Hoff J. An automatic method for accurate volume delineation of heterogeneous tumors in PET. *Med Phys*. 2013 Aug;40(8):082503. doi: 10.1118/1.4812892.
78. Drever L, Roa W, McEwan A, Robinson D. Comparison of three image segmentation techniques for target volume delineation in positron emission tomography. *J Appl Clin Med Phys*. 2007;8(2):93–109.
79. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern*. 1980 Apr;36(4):193–202. Available from: <http://link.springer.com/10.1007/BF00344251>

Literaturverzeichnis

80. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide. *Radiology*. 2019;290(3):590–606.
81. Sibille L, Seifert R, Avramovic N, Vehren T, Spottiswoode B, Zuehlsdorff S, Schäfers M. 18F-FDG PET/CT Uptake Classification in Lymphoma and Lung Cancer by Using Deep Convolutional Neural Networks. *Radiology*. 2020 Feb;294(2):445-452. doi: 10.1148/radiol.2019191114.
82. Zadeh Shirazi A, Fornaciari E, McDonnell MD, Yaghoobi M, Cevallos Y, Tello-Oquendo L, Inca D, Gomez GA. The Application of Deep Convolutional Neural Networks to Brain Cancer Images: A Survey. *J Pers Med*. 2020 Nov 12;10(4):224. doi: 10.3390/jpm10040224.
83. Arif M, Schoots IG, Castillo Tovar J, Bangma CH, Krestin GP, Roobol MJ, Niessen W, Veenland JF. Clinically significant prostate cancer detection and segmentation in low-risk patients using a convolutional neural network on multi-parametric MRI. *Eur Radiol*. 2020 Dec;30(12):6582-6592. doi: 10.1007/s00330-020-07008-z. Epub 2020 Jun 27. PMID: 32594208; PMCID: PMC7599141.
84. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Med Phys*. 2017 Feb;44(2):547–57.
85. Mahmood SA, Ahmed HA. An improved CNN-based architecture for automatic lung nodule classification. *Med Biol Eng Comput* [Internet]. 2022 Jul 7;60(7):1977–86. Available from: <https://link.springer.com/10.1007/s11517-022-02578-0>
86. Wang XY, Zhao YF, Liu Y, Yang YK, Zhu Z, Wu N. Comparison of different automated lesion delineation methods for metabolic tumor volume of 18 F-FDG PET/CT in patients with stage I lung adenocarcinoma. *Med (United States)*. 2017;96(51):1–9.
87. Warfield SK, Zou KH, Wells WM. Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation. *IEEE Trans Med Imaging*. 2004 Jul;23(7):903–21.
88. Nestle U, Kremp S, Grosu A-L. Practical integration of [18F]-FDG-PET and PET-CT in the planning of radiotherapy for non-small cell lung cancer (NSCLC): The technical basis, ICRU-target volumes, problems, perspectives. *Radiother Oncol*. 2006 Nov;81(2):209–25.
89. Caldwell CB, Mah K, Skinner M, Danjoux CE. Can PET provide the 3D extent of tumor motion for individualized internal target volumes? A phantom study of the limitations of CT

Literaturverzeichnis

and the promise of PET. *Int J Radiat Oncol*. 2003 Apr;55(5):1381–93. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0360301602046096>

90. Bradley J, Thorstad WL, Mutic S, Miller TR, Dehdashti F, Siegel BA, Bosch W, Bertrand RJ. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2004 May 1;59(1):78-86. doi: 10.1016/j.ijrobp.2003.10.044.
91. Paesmans M, Berghmans T, Dusart M, Garcia C, Hossein-Foucher C, Lafitte JJ, Mascaux C, Meert AP, Roelandts M, Scherpereel A, Terrones Munoz V, Sculier JP; European Lung Cancer Working Party, and on behalf of the IASLC Lung Cancer Staging Project. Primary tumor standardized uptake value measured on fluorodeoxyglucose positron emission tomography is of prognostic value for survival in non-small cell lung cancer: update of a systematic review and meta-analysis by the European Lung Cancer Working Party for the International Association for the Study of Lung Cancer Staging Project. *J Thorac Oncol*. 2010 May;5(5):612-9. doi: 10.1097/JTO.0b013e3181d0a4f5.
92. Lodge MA, Chaudhry MA, Wahl RL. Noise considerations for PET quantification using maximum and peak standardized uptake value. *J Nucl Med*. 2012 Jul;53(7):1041-7. doi: 10.2967/jnumed.111.101733. Epub 2012 May 24.
93. Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of Positron Emission Tomography for Diagnosis of Pulmonary Nodules and Mass Lesions. *JAMA*. 2001 Feb 21;285(7):914.
94. Hashimoto Y, Tsujikawa T, Kondo C, Maki M, Momose M, Nagai A, Ohnuki T, Nishikawa T, Kusakabe K. Accuracy of PET for diagnosis of solid pulmonary lesions with ¹⁸F-FDG uptake below the standardized uptake value of 2.5. *J Nucl Med*. 2006 Mar;47(3):426-31.
95. Büsing KA, Schönberg SO, Brade J, Wasser K. Impact of blood glucose, diabetes, insulin, and obesity on standardized uptake values in tumors and healthy organs on ¹⁸F-FDG PET/CT. *Nucl Med Biol*. 2013 Feb;40(2):206–13.
96. Sprinz C, Altmayer S, Zanon M, Watte G, Irion K, Marchiori E, Hochegger B. Effects of blood glucose level on ¹⁸F-FDG uptake for PET/CT in normal organs: A systematic review. *PLoS One*. 2018 Feb 27;13(2):e0193140. doi: 10.1371/journal.pone.0193140.

Literaturverzeichnis

97. Vanhove K, Mesotten L, Heylen M, Derwael R, Louis E, Adriaensens P, Thomeer M, Boellaard R. Prognostic value of total lesion glycolysis and metabolic active tumor volume in non-small cell lung cancer. *Cancer Treat Res Commun.* 2018;15:7-12. doi: 10.1016/j.ctarc.2017.11.005.
98. Apostolova I, Steffen IG, Wedel F, Lougovski A, Marnitz S, Derlin T, Amthauer H, Buchert R, Hofheinz F, Brenner W. Asphericity of pretherapeutic tumour FDG uptake provides independent prognostic value in head-and-neck cancer. *Eur Radiol.* 2014 Sep;24(9):2077-87. doi: 10.1007/s00330-014-3269-8.
99. Wadell H. Volume, Shape, and Roundness of Quartz Particles. *J Geol.* 1935 Apr;43(3):250–80.
100. Apostolova I, Rogasch J, Buchert R, Wertz H, Achenbach HJ, Schreiber J, Riedel S, Furth C, Lougovski A, Schramm G, Hofheinz F, Amthauer H, Steffen IG. Quantitative assessment of the asphericity of pretherapeutic FDG uptake as an independent predictor of outcome in NSCLC. *BMC Cancer.* 2014 Dec 1;14:896. doi: 10.1186/1471-2407-14-896.
101. Hofheinz F, Lougovski A, Zöphel K, Hentschel M, Steffen IG, Apostolova I, Wedel F, Buchert R, Baumann M, Brenner W, Kotzerke J, van den Hoff J. Increased evidence for the prognostic value of primary tumor asphericity in pretherapeutic FDG PET for risk stratification in patients with head and neck cancer. *Eur J Nucl Med Mol Imaging.* 2015 Mar;42(3):429-37. doi: 10.1007/s00259-014-2953-x.
102. Apostolova I, Ego K, Steffen IG, Buchert R, Wertz H, Achenbach HJ, Riedel S, Schreiber J, Schultz M, Furth C, Derlin T, Amthauer H, Hofheinz F, Kalinski T. The asphericity of the metabolic tumour volume in NSCLC: correlation with histopathology and molecular markers. *Eur J Nucl Med Mol Imaging.* 2016 Dec;43(13):2360-2373. doi: 10.1007/s00259-016-3452-z.
103. Meißner S, Janssen JC, Prasad V, Brenner W, Diederichs G, Hamm B, Hofheinz F, Makowski MR. Potential of asphericity as a novel diagnostic parameter in the evaluation of patients with 68Ga-PSMA-HBED-CC PET-positive prostate cancer lesions. *EJNMMI Res.* 2017 Oct 23;7(1):85. doi: 10.1186/s13550-017-0333-9.
104. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials.* 1996 Aug;17(4):343–6.

Literaturverzeichnis

105. Vrieze SI. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol Methods* . 2012;17(2):228–43.
106. Burnham KP, Anderson DR. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*. 2004;33(2):261-304. doi:10.1177/0049124104268644
107. Onn A, Choe DH, Herbst RS, Correa AM, Munden RF, Truong MT, Vaporciyan AA, Isobe T, Gilcrease MZ, Marom EM. Tumor cavitation in stage I non-small cell lung cancer: epidermal growth factor receptor expression and prediction of poor outcome. *Radiology*. 2005 Oct;237(1):342-7. doi: 10.1148/radiol.2371041650.
108. Yip SSF, Liu Y, Parmar C, Li Q, Liu S, Qu F, Ye Z, Gillies RJ, Aerts HJWL. Associations between radiologist-defined semantic and automatically computed radiomic features in non-small cell lung cancer. *Sci Rep*. 2017 Jun 14;7(1):3519. doi: 10.1038/s41598-017-02425-5.
109. Rogasch JMM, Furth C, Chibolela C, Hofheinz F, Ochsenreither S, Rückert JC, Neudecker J, Böhmer D, von Laffert M, Amthauer H, Frost N. Validation of Independent Prognostic Value of Asphericity of 18F-Fluorodeoxyglucose Uptake in Non-Small-Cell Lung Cancer Patients Undergoing Treatment With Curative Intent. *Clin Lung Cancer*. 2020 May;21(3):264-272.e6. doi: 10.1016/j.clcc.2019.10.001.
110. Sun G-W, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol*. 1996 Aug;49(8):907–16. Available from: <https://linkinghub.elsevier.com/retrieve/pii/089543569600025X>
111. Heinze G, Dunkler D. Five myths about variable selection. *Transpl Int*. 2017 Jan;30(1):6–10. Available from: <https://onlinelibrary.wiley.com/doi/10.1111/tri.12895>
112. Abdoell M, LeBlanc M, Stephens D, Harrison R V. Binary partitioning for continuous longitudinal data: categorizing a prognostic variable. *Stat Med*. 2002 Nov 30;21(22):3395–409. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/sim.1266>
113. Chang C, Hsieh M-K, Chang W-Y, Chiang AJ, Chen J. Determining the optimal number and location of cutoff points with application to data of cervical cancer. Sung S-Y, editor. *PLoS One* . 2017 Apr 27;12(4):e0176231.

Literaturverzeichnis

114. Schulgen G, Lausen B, Olsen JH, Schumacher M. Outcome-oriented Cutpoints in Analysis of Quantitative Exposures. *Am J Epidemiol.* 1994 Jul 15;140(2):172–84. Available from: <https://academic.oup.com/aje/article/93866/Outcome-oriented>
115. Schild SE, Tan AD, Wampfler JA, Ross HJ, Yang P, Sloan JA. A new scoring system for predicting survival in patients with non-small cell lung cancer. *Cancer Med.* 2015 Sep 23;4(9):1334–43. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/cam4.479>
116. Strand T-E. Survival after resection for primary lung cancer: a population based study of 3211 resected patients. *Thorax.* 2006 Aug 1;61(8):710–5. Available from: <https://thorax.bmj.com/lookup/doi/10.1136/thx.2005.056481>
117. Selvaggi G, Scagliotti G V. Histologic subtype in NSCLC: does it matter? *Oncology (Williston Park).* 2009 Nov 30;23(13):1133–40. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20043461>
118. Okamoto T. Prognostic value of the histological subtype in completely resected non-small cell lung cancer. *Interact Cardiovasc Thorac Surg.* 2006 May 24;5(4):362–6. Available from: <https://academic.oup.com/icvts/article-lookup/doi/10.1510/icvts.2005.125989>
119. Moses WW. Fundamental limits of spatial resolution in PET. *Nucl Instruments Methods Phys Res Sect A Accel Spectrometers, Detect Assoc Equip.* 2011 Aug;648:S236–40. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0168900210026276>
120. Mandelbrot B. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science.* 1967 May 5;156(3775):636-8. doi: 10.1126/science.156.3775.636.
121. Hanahan D, Weinberg RA. The Hallmarks of Cancer. *Cell* 2000 Jan;100(1):57–70. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867400816839>
122. Limkin EJ, Reuzé S, Carré A, Sun R, Schernberg A, Alexis A, Deutsch E, Ferté C, Robert C. The complexity of tumor shape, spiculatedness, correlates with tumor radiomic shape features. *Sci Rep.* 2019 Mar 13;9(1):4329. doi: 10.1038/s41598-019-40437-5.
123. Warth A, Cortis J, Soltermann A, Meister M, Budczies J, Stenzinger A, Goeppert B, Thomas M, Herth FJ, Schirmacher P, Schnabel PA, Hoffmann H, Dienemann H, Muley T, Weichert W. Tumour cell proliferation (Ki-67) in non-small cell lung cancer: a critical reappraisal of its prognostic role. *Br J Cancer.* 2014 Sep 9;111(6):1222-9. doi: 10.1038/bjc.2014.402.

Literaturverzeichnis

124. Gharraf HS, Mehana SM, ElNagar MA. Role of CT in differentiation between subtypes of lung cancer; is it possible? *Egypt J Bronchol.* 2020 Dec 17;14(1):28. Available from: <https://ejb.springeropen.com/articles/10.1186/s43168-020-00027-w>
125. La Greca Saint-Estevan A, Vuong D, Tschanz F, van Timmeren JE, Dal Bello R, Waller V, Pruschy M, Guckenberger M, Tanadini-Lang S. Systematic Review on the Association of Radiomics with Tumor Biological Endpoints. *Cancers (Basel).* 2021 Jun 16;13(12):3015. doi: 10.3390/cancers13123015.
126. Whi W, Ha S, Bae S, Choi H, Paeng JC, Cheon GJ, Kang KW, Lee DS. Relationship of EGFR Mutation to Glucose Metabolic Activity and Asphericity of Metabolic Tumor Volume in Lung Adenocarcinoma. *Nucl Med Mol Imaging.* 2020 Aug;54(4):175-182. doi: 10.1007/s13139-020-00646-7.
127. Taylor MD, LaPar DJ, Thomas CJ, Persinger M, Stelow EB, Kozower BD, Lau CL, Jones DR. Lymph node ratio predicts recurrence and survival after R0 resection for non-small cell lung cancer. *Ann Thorac Surg.* 2013 Oct;96(4):1163-1170. doi: 10.1016/j.athoracsur.2013.04.031.
128. Hatt M, Laurent B, Fayad H, Jaouen V, Visvikis D, Le Rest CC. Tumour functional sphericity from PET images: prognostic value in NSCLC and impact of delineation method. *Eur J Nucl Med Mol Imaging.* 2018 Apr;45(4):630-641. doi: 10.1007/s00259-017-3865-3.
129. Yang F, Simpson G, Young L, Ford J, Dogan N, Wang L. Impact of contouring variability on oncological PET radiomics features in the lung. *Sci Rep.* 2020;1–10. Available from: <http://dx.doi.org/10.1038/s41598-019-57171-7>
130. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, Bogowicz M, Boldrini L, Buvat I, Cook GJR, Davatzikos C, Depeursinge A, Desseroit MC, Dinapoli N, Dinh CV, Echegaray S, El Naqa I, Fedorov AY, Gatta R, Gillies RJ, Goh V, Götz M, Guckenberger M, Ha SM, Hatt M, Isensee F, Lambin P, Leger S, Leijenaar RTH, Lenkowicz J, Lippert F, Losnegård A, Maier-Hein KH, Morin O, Müller H, Napel S, Nioche C, Orhac F, Pati S, Pfaehler EAG, Rahmim A, Rao AUK, Scherer J, Siddique MM, Sijtsema NM, Socarras Fernandez J, Spezi E, Steenbakkers RJHM, Tanadini-Lang S, Thorwarth D, Troost EGC, Upadhaya T, Valentini V, van Dijk LV, van Griethuysen J, van Velden FHP, Whybra P, Richter C, Löck S. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology.* 2020 May;295(2):328-338. doi: 10.1148/radiol.2020191145.

Literaturverzeichnis

131. Mullikin JC, Verbeek PW. Surface area estimation of digitized planes. *Bioimaging*. 1993;1(1):6–16.
132. Fornacon-Wood I, Mistry H, Ackermann CJ, Blackhall F, McPartlin A, Faivre-Finn C, Price GJ, O'Connor JPB. Reliability and prognostic value of radiomic features are highly dependent on choice of feature extraction platform. *Eur Radiol*. 2020 Nov;30(11):6241-6250. doi: 10.1007/s00330-020-06957-9. Epub 2020 Jun 1.

Eidesstattliche Versicherung

„Ich, Christoph Chibolela, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema:

Die Asphärizität des präoperativen metabolischen Tumorumens in der FDG-PET: Prognostische Bedeutung beim NSCLC und Einfluss der Segmentierungsmethod

(Asphericity of preoperative metabolic tumour volume in FDG-PET: prognostic value in NSCLC and dependence on segmentation method)

selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere Datenakquisition und statistische Aufarbeitung) und Resultate (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem Erstbetreuer, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.org) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte.

Eidesstattliche Versicherung

Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Datum

Unterschrift

Anteilserklärung an etwaigen erfolgten Publikationen

Christoph Chibolela hatte folgenden Anteil an der folgenden Publikation:

Rogasch, J.M.M., Furth, C., Chibolela, C., Hofheinz, F., Ochsenreither, S., Rückert, J.C., Neudecker, J., Böhmer, D., von Laffert, M., Amthauer, H., Frost, N.: Validation of Independent Prognostic Value of Asphericity of 18 F-Fluorodeoxyglucose Uptake in Non-Small-Cell Lung Cancer Patients Undergoing Treatment With Curative Intent. Clin Lung Cancer. 2020;21(3):264-272.e6.

Beiträge im Einzelnen:

- **Studiendesign:** Diskussion und Festlegung der zugrunde liegenden Studienpopulation, Definition der Endpunkte und zu verwendende Datenquellen gemeinsam mit dem Erstautor.
- **Patient*inneneinschluss:** Sichtung des Klinischen Krebsregisters der Charité und Auswahl der Patient*innen nach zuvor definierten Ein- und Ausschlusskriterien.
- **Erhebung klinischer Daten:** Recherche der onkologischen Verläufe und Ermittlung von Gesamtüberleben, krankheitsfreien Überleben und weiterer klinischer Metadaten. Einholen der Follow-up Daten durch Kontaktaufnahme zu Patient*innen und/oder Behandler*innen.
- **Auswertung der Bild-Daten:** Einlesen der PET-Daten in die ROVER-Software. Identifikation und Segmentierung der tumorösen Läsionen unter fachärztlicher Supervision. Extraktion der zu untersuchenden PET-Parameter (ASP, MTV, SUV_{max}) sowie weiterer, technischer Metadaten.

Eidesstattliche Versicherung

- **Statistische Analyse:** Zusammenführung aller erhobenen Daten. Durchführung der Lebenszeitanalysen sowie Subgruppenanalysen in *SPSS* unter Anleitung durch den Erstautor.
- **Literaturrecherche:** Sichtung und Studium relevanter Literatur zur ASP bei NSCLC in *PubMed*®. Kritische Prüfung, Diskussion und Einpflegen in die Publikation.
- **Review-Verfahren:** Durchsicht und Aufbereitung der Daten in Rücksprache mit dem Erstautor zur Beantwortung spezifischer Fragestellungen der Reviewer.

Unterschrift, Datum und Stempel des/der erstbetreuenden Hochschullehrers/in

Unterschrift des Doktoranden/der Doktorandin

Lebenslauf

"Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht."

Publikationsliste

- Rogasch, J.M.M., Furth, C., **Chibolela, C.**, Hofheinz, F., Ochsenreither, S., Rückert, J.C., Neudecker, J., Böhmer, D., von Laffert, M., Amthauer, H., Frost, N.: Validation of Independent Prognostic Value of Asphericity of 18 F-Fluorodeoxyglucose Uptake in Non-Small-Cell Lung Cancer Patients Undergoing Treatment With Curative Intent. *Clin Lung Cancer*. 2020;21(3):264-272.e6.

Danksagung

In erster Linie möchte ich mich bei PD Dr. Julian Rogasch bedanken. Er stand mir von Beginn bis zum Ende dieser Arbeit geduldig und mit großem Engagement zur Seite, er hat mich in die Methoden eingeführt und mir stets eine großartige fachliche Unterstützung sowie strukturiertes Feedback entgegengebracht.

Ein besonderer Dank gilt Prof. Dr. Holger Amthauer, der es mir überhaupt möglich gemacht hat, in seinem Arbeitsbereich klinische Nuklearmedizin ein Forschungsprojekt zu beginnen. Für seine freundliche und stets motivierende Zusprache und Unterstützung möchte ich mich sehr bedanken.

Großer Dank an Dr. Nokolaj Frost, Leiter des Charité Lungenkrebszentrums CVK, für die klinische Zusammenarbeit und Bereitstellung der Daten aus dem Krebsregister der Charité.

Vielen Dank auch an Dr. Frank Hofheinz vom Helmholtz-Zentrum Dresden-Rossendorf, der für technische Fragen stets zur Verfügung stand und die Software zur Analyse der Bilddaten zur Verfügung stellte.

Danke an alle Patient*innen, behandelnden Hausarzt*innen und Onkolog*innen, die im Rahmen des Follow-ups bereitwillig mitgewirkt haben.

Ein großer Dank an all jene, die sich der Open Source Software Bewegung verschreiben und ihr Wissen und ihre Ressourcen bereitwillig allen Menschen frei im Internet zur Verfügung stellen. Diese Arbeit hat stark auf die Verwendung der Software Python und verschiedensten assoziierten Bibliotheken gefußt, welche die Auswertung und Darstellung der Daten erheblich erleichtert hat.

Abschließend möchte ich ein herzliches Danke meiner Familie und meinen Freunden aussprechen, die stets an den Erfolg dieser Arbeit geglaubt haben. Vielen Dank an dieser Stelle auch an Markus Galler MSc., Kommilitone, Freund und Kollege, der mich seit unserer ersten Begegnung immer wieder aufs Neue mit seiner wissenschaftlichen Neugierde im Bereich der Physik und Medizin angesteckt hat.

Bescheinigung Statistik



CharitéCentrum für Human- und Gesundheitswissenschaften

Charité | Campus Charité Mitte | 10117 Berlin

Name, Vorname: Chibolela, Christoph
Emailadresse: christoph.chibolela@charite.de
Matrikelnummer: 217287
PromotionsbetreuerIn: Prof. Holger Amthauer
Promotionsinstitution/ Klinik: Nuklearmedizin

Institut für Biometrie und Klinische Epidemiologie (iBikE)

Direktor: Prof. Dr. Geraldine Rauch

Postanschrift:
Charitéplatz 1 | 10117 Berlin
Besucheranschrift:
Reinhardtstr. 58 | 10117 Berlin

Tel. +49 (0)30 450 562171
geraldine.rauch@charite.de
<https://biometrie.charite.de/>



Bescheinigung

Hiermit bescheinige ich, dass Herr Christoph Chibolela innerhalb der Service Unit Biometrie des Instituts für Biometrie und Klinische Epidemiologie (iBikE) bei mir eine statistische Beratung zu einem Promotionsvorhaben wahrgenommen hat. Folgende Beratungstermine wurden wahrgenommen:

- 29.11.2021 & 25.1.2022

Folgende wesentliche Ratschläge hinsichtlich einer sinnvollen Auswertung und Interpretation der Daten wurden während der Beratung erteilt:

- Multivariable Cox-Regression sollte verwendet werden um den Zusammenhang zwischen Asphärizität und den beiden Endpunkten (disease-free survival, overall survival) zu schätzen, sodass für potenzielle Confounder adjustiert werden kann. Hierbei ist die Verwendung des kontinuierlichen Messwerts der Asphärizität einem arbiträren Cut-off des Messwerts vorzuziehen.
- Beobachtungen/Probanden sollten nicht von der Analyse ausgeschlossen werden, wenn diese keine minimale Follow-Up-Zeit haben, da dies die Gefahr eines Bias in der Schätzung birgt.
- Für die Darstellung der Ergebnisse aus der Cox-Analyse eignen sich Grafiken, die Hazard Ratio und 95% Konfidenzintervalle darstellen.
- Da es sich um eine Beobachtungsstudie ohne formale Fallzahlplanung handelt, sollte die Interpretation der berechneten Effektstärken, wie zB das Hazard Ratio, im Mittelpunkt stehen um die klinische Relevanz der Ergebnisse einzuschätzen und nicht die statistische Signifikanz.

Diese Bescheinigung garantiert weder die richtige Umsetzung der in der Beratung gemachten Vorschläge, die korrekte Durchführung der empfohlenen statistischen Verfahren noch die richtige Darstellung und Interpretation der Ergebnisse. Die Verantwortung hierfür obliegt allein dem Promovierenden. Das Institut für Biometrie und Klinische Epidemiologie übernimmt hierfür keine Haftung.

Datum: 28.1.2022

Name des Beraters/ der Beraterin: Annette Aigner

Unterschrift BeraterIn, Institutsstempel

