

Aus der Abteilung
für Nuklearmedizin
Prof. Dr. Dirk Hellwig
der Fakultät für Medizin
der Universität Regensburg

Metabolisches Tumervolumen in der initialen FDG-PET/CT bei aggressiven
Non-Hodgkin-Lymphomen in der OPTIMAL>60-Studie

Inaugural – Dissertation
zur Erlangung des Doktorgrades
der Medizin

der
Fakultät für Medizin
der Universität Regensburg

vorgelegt von
Lukas Pichler

2025

Aus der Abteilung
für Nuklearmedizin
Prof. Dr. Dirk Hellwig
der Fakultät für Medizin
der Universität Regensburg

Metabolisches Tumervolumen in der initialen FDG-PET/CT bei aggressiven
Non-Hodgkin-Lymphomen in der OPTIMAL>60-Studie

Inaugural – Dissertation
zur Erlangung des Doktorgrades
der Medizin

der
Fakultät für Medizin
der Universität Regensburg

vorgelegt von
Lukas Pichler

2025

Dekan:	Prof. Dr. Dirk Hellwig
1. Berichterstatter:	Prof. Dr. Dirk Hellwig
2. Berichterstatter:	Prof. Dr. Tobias Pukrop
Tag der mündlichen Prüfung:	13.10.2025

Inhaltsverzeichnis

1. Abstract	3
2. Einleitung	4
2.1 Motivation	4
2.2 Die Krankheitsgruppe der Lymphome	6
2.2.1 Definition – „Was ist ein malignes Lymphom?“	6
2.2.2 Epidemiologie	6
2.2.3 Ätiologie und Risikofaktoren	8
2.2.4 Klassifikation	8
2.2.5 Risikostratifizierung, Therapie und Prognose	9
2.2.5.1 Risikostratifizierung bei aggressiven Non-Hodgkin-Lymphomen	9
2.2.5.2 Standardtherapie bei DLBCL	10
2.2.5.3 Weitere Therapiemöglichkeiten bei aggressiven NHL verschiedener Entitäten	12
2.3 Die ¹⁸ F-FDG-PET/CT in der Lymphomdiagnostik	13
2.4 Metabolische Biomarker der FDG-PET/CT	15
2.4 Konturierung von Tumervolumina in der FDG-PET/CT	16
2.5 Definition der Forschungsfragen	17
3. Material und Methoden	18
3.1 Die OPTIMAL>60-Studie	18
3.1.1 Ein- und Ausschlusskriterien	18
3.2 Bereitstellung der initialen FDG-PET/CTs	19
3.3 PET-Befundungskriterien	21
3.4 Ablauf und Qualitätskontrolle der FDG-PET/CT-Untersuchung	23
3.4.1 Vorbereitung des Patienten	23
3.4.2 Standards der PET-Bildakquisition	24
3.5 Computergestützte Auswertung mithilfe des Programms syngo.via	24
3.6 Methodik der Interobserveranalysen	28
3.6.1 Methodik der ersten Interobserveranalyse mit dem Programm syngo.via	28
3.6.2 Methodik der zweiten Interobserver- und Interprogrammanalyse mit „syngo-via“ und „ROVER“	28
3.7 Statistische Analyse	29
4. Ergebnisse	31
4.1 Konturierungsqualität im visuellen Scoring	31
4.2 Ermittelte Ursachen der Unterschätzung des Tumervolumens mit scheinbarem MTV von 0 mL	32

4.3	SUV _{mean} der Leberreferenzregion und die PERCIST-Schwellwerte in der Gesamtanalyse.....	32
4.4	Einfluss von Körpergewicht, Radioaktivitätsmenge und Distributionszeit	34
4.5	Standortvergleich der SUV _{mean} und PERCIST-Schwellwerte	36
4.6	PET-Scannermodellvergleich der SUV _{mean} und PERCIST-Schwellwerte.....	38
4.7	Vergleich von SUV _{mean} mediastinal und dem SUV _{mean} der Leber.....	41
4.8	Interobserver-Analyse der repräsentativen Stichprobe in syngo.via.....	42
4.9	Interobserver-Analyse des MTV und der TLG bei Nutzung verschiedener Softwares und Konturierungsmethoden	45
4.10	MTV, TLG und SUV _{max} in den ausgewerteten Fällen.....	47
4.11	MTV, TLG und SUV _{max} je nach Risikostratifizierung	49
4.12	Statistische Ausreißer in der Bestimmung des MTV und der TLG	51
4.13	Einfluss der MTV-Last auf den SUV _{mean} der Leber und PERCIST-Schwellwert ...	51
5.	Diskussion	53
5.1	Evaluation der Ergebnisse der Interobserver-Analysen	53
5.2	Evaluation der Ergebnisse der PERCIST-Schwellwerte nach PET-Scannermodell, Studienzentrum und Patientengewicht	54
5.3	Evaluation der Ergebnisse zur Normalverteilung von SUV _{mean} , SUV _{SD} und dem PERCIST-Schwellwert.....	55
5.4	Evaluation des visuellen Scorings und der ermittelten Werte der Biomarker	56
5.5	Limitationen und potenzielle Fehlerquellen.....	58
6.	Zusammenfassung.....	63
7.	Literatur	65
8.	Abkürzungsverzeichnis	73
9.	Originaldokument des CRF.....	75
10.	Danksagung.....	76
11.	Lebenslauf.....	77

1. Abstract

Fragestellung: Internationale Leitlinien empfehlen die FDG-PET/CT zum Staging und Restaging von aggressiven Non-Hodgkin-Lymphomen (NHL). Bildbasierte Biomarker, allen voran das aus FDG-PET-Bildern durch Konturierung von Lymphommanifestationen abgeleitete metabolische Tumervolumen (MTV), ermöglichen eine Quantifizierung der Tumorlast und erweisen sich zunehmend als prognostische Faktoren. Das MTV und die damit verbundene Total Lesion Glykolyisis (TLG) sind allerdings je nach Messmethodik Schwankungen unterworfen. Die hier mit einer Schwellwert-basierten Konturierung gewonnenen Biomarker sollten einerseits bezüglich ihrer Reproduzierbarkeit und andererseits im Vergleich zu einer vorangegangenen Auswertung mit kontrastadaptiver Konturierung analysiert werden.

Methodik: Diese Arbeit nutzte die FDG-PET/CT-Bilddatensätze aus dem initialen Staging (n=661) der deutschlandweiten OPTIMAL>60-Studie zu aggressiven NHL mit 1152 Patienten zur Ermittlung von MTV und TLG. Hier wurden anhand ihrer FDG-Aufnahme im Vergleich zum Standardized Uptake Value (SUV) in einer Leberreferenzregion (PERCIST-Schwelle) selektiert und mit einer relativen Konturierung von 41% des tumoralen SUV-Maximums (SUV_{max}) vermessen (Programm: syngo.via). Die Qualität der Konturierung wurde visuell mit einem 5-stufigen Score beurteilt. Ein Subgruppe von n=60 Fällen diente zum Interobserver-Vergleich der Reproduzierbarkeit der Messung, bei n=651 Fällen konnte ein Vergleich mit der kontrastadaptiven Konturierung erfolgen (Programm: ROVER).

Ergebnisse: Die PERCIST-Schwellen und SUV_{mean} der Leber zeigten mit $4,23 \pm 1,00$ bzw. $2,34 \pm 0,53$ eine Normalverteilung, unterscheiden sich aber zwischen Standorten ($p < 0,023$) und PET-Scannermodell ($p < 0,001$). 56,8% der PERCIST-Schwellen lagen über der zunehmend empfohlenen SUV-Schwelle von 4,0. Visuell wurde das hier gemessene MTV in 84,3% der Fälle als adäquat konturiert eingestuft, eine subjektive Unterschätzung (12,8%) des MTV trat deutlich häufiger auf als eine Überschätzung (2,7%). Die Interobserver-Analyse zeigte eine hohe Reproduzierbarkeit für syngo.via-Messungen (MTV: $r = 0,999$; TLG: $r = 0,999$; SUV_{max} : $r = 0,997$, alle $p < 0,001$). Der Vergleich zu Messungen mittels ROVER zeigt bei ebenfalls gutem Ergebnis eine höhere Variabilität (MTV: $r = 0,95$; TLG: $r = 0,96$; SUV_{max} : $r = 0,97$, alle $p < 0,001$).

Schlussfolgerung: Mittels des zugelassenen Medizinproduktes syngo.via kann bei den allermeisten Fällen das MTV adäquat und bestens reproduzierbar konturiert werden, ebenso die Messung von TLG und SUV_{max} . Die experimentelle Auswertesoftware ROVER liefert gut vergleichbare Ergebnisse. Die hier gewonnenen Werte eignen sich für die weitergehende Analyse der OPTIMAL>60-Studie nach Aufhebung des Datenembargos Anfang 2025.

2. Einleitung

2.1 Motivation

Die apparative Untersuchung maligner Tumore mit Hilfe der PET/CT erfährt seit geraumer Zeit eine zunehmende nationale und internationale Bedeutung. Während nach Daten des „IMAGINE“-Projekts der Internationalen Atomenergie-Organisation (IAEA) besonders die USA und einzelne europäische Länder wie Finnland und Italien mit mehr als 3 PET-Scannern pro 1 000 000 Einwohner eine hohe PET-Scanner-Dichte besitzen, herrscht in Deutschland mit „zwischen 1 und 2 PET-Scanner pro 1 000 000 Einwohner“ eine vergleichsweise niedrige Anzahl an diesen [1, 2].

Trotz der noch eingeschränkten Verfügbarkeit und zusätzlichen Abrechnungshürden in Deutschland besitzt das PET/CT gesicherte Vorteile im Rahmen eines routinemäßigen „Stagings“ bzw. „Restagings“ von malignen Tumoren verschiedener Entitäten. Diese Vorteile basieren darauf, dass im Gegensatz zu einer Bildgebung mittels Magnetresonanztomographie (MRT) oder alleiniger Computertomographie (CT) die PET zusätzlich in der Lage ist, biologische Charakteristika des vorliegenden Tumors darzustellen. Die Grundlage hierfür ist der erhöhte Glukoseumsatz der malignen Zellen bei verschiedenen Tumorentitäten und analog dazu die erhöhte Verstoffwechslung des radioaktiven Tracers in eben diesen. Das instabile, positronenemittierende, radioaktive Glukoseanalogon F-18-Fluorodeoxyglukose (FDG) dient als eben solcher Tracer, welcher beim Zerfall Positronen emittiert, die durch Vernichtung mit einem Elektron als zugehörigem Antiteilchen messbare elektromagnetische Strahlung einer Energie von 511 keV aussenden. Bereiche mit erhöhten Glukoseumsatz, sowohl pathologischer als auch physiologischer Natur, erscheinen im FDG-PET/CT daher mit einer erhöhten Signalintensität. Das FDG-PET/CT wird als das genaueste nicht-invasive Verfahren zur Ausbreitungsdiagnostik bei aggressiven und den häufigsten indolenten Lymphomen gehandelt [3].

Mithilfe des FDG-PET/CT ist die Quantifizierung des metabolischen Tumervolumens (MTV) und dem daraus abgeleiteten Wert Total Lesion Glycolysis (TLG) möglich. Sowohl das MTV, als auch die TLG in der initialen-FDG-PET/CT werden zunehmend als prognostische Faktoren bei Hodgkin-Lymphomen [4] und Non-Hodgkin-Lymphomen (NHL), allen voran beim diffus großzelligen B-Zell-Lymphom (DLBCL) und weiteren hochmalignen NHL [3].

Das MTV bei Erstdiagnose wird als wertiger prognostischer Faktor, ähnlich bzw. unterstützend zum internationalen prognostischen Index (IPI), gehandelt. Aktuelle Studienergebnisse, wie die

von *Mikhaeel et. al (2022)*, bestätigen einer Bewertungsmethode, die das MTV in der initial-PET miteinbezieht, eine höhere Wertigkeit bei Patienten mit DLBCL in der Vorhersage der Gesamtüberlebenszeit (OS) und des progressionsfreien Überlebens (PFS), als dem IPI [5].

Ein vereinheitlichtes Verfahren zur Bestimmung der metabolischen Tumervolumina wurde bis dato noch nicht implementiert und daher existieren eine Reihe von relativen und absoluten Schwellwerten zur Bestimmung des MTV, welche die Ergebnisse nur bedingt untereinander vergleichbar machen.

Die OPTIMAL>60-Studie, auf deren PET-Bildgrundlagen sich diese Dissertation stützt, ist eine deutschlandweite, multizentrische, prospektive Studie an insgesamt 1152 Patienten mit aggressivem NHL. Die Zielsetzung der OPTIMAL>60-Studie liegt darin, im randomisierten Studiendesign die medikamentöse Therapie bildgebungsgesteuert zu optimieren. Das Vorgehen der OPTIMAL>60-Studie wurde von der zuständigen Ethikkommission genehmigt.

Das Ziel dieser Arbeit ist hierbei, aus den zur Verfügung gestellten initialen FDG-PET/CTs einen Selektions- und Konturierungsmechanismus zur Berechnung der MTV auf seine Validität und Reproduzierbarkeit in Interobserveranalysen zu prüfen und potenzielle Einflussfaktoren zu analysieren. Des Weiteren erlaubt die Einteilung der OPTIMAL>60-Studie der Patienten gemäß der IPI-Items in verschiedene Risikogruppen eine Analyse der metabolischen Biomarker nach vorangegangener Risikostratifizierung.

2.2 Die Krankheitsgruppe der Lymphome

2.2.1 Definition – „Was ist ein malignes Lymphom?“

Der Überbegriff „malignes Lymphom“ umfasst eine heterogene Erkrankungsgruppe bösartiger Tumore, welche ihren Ursprung im lymphatischen Gewebe haben. Die lymphatischen Zellen können sowohl in ihrer adulten Form, als auch in ihren Vorgängerformen tumorös proliferieren [6]. Im klinischen Alltag wird der Begriff „Lymphom“ mit dem Begriff „malignen Lymphom“ weitestgehend gleichgesetzt.

Bei M. Hodgkin/dem Hodgkin-Lymphom (HL) und dem Non-Hodgkin-Lymphom (NHL) handelt es sich um die zwei großen Gruppierungen der malignen Lymphome, welche darüber hinaus noch in eine Vielzahl weiterer zytologischer Subtypen eingeteilt werden. Als Grundlage hierfür dient die WHO-Klassifikation, welche zuletzt 2016 aktualisiert wurde [6, 7].

2.2.2 Epidemiologie

NHL gehören zu den häufigsten malignen Erkrankungen in Deutschland. Nach Daten des Robert Koch Instituts (RKI) betrug die altersstandardisierte Erkrankungsrate im Jahr 2020 allein in Deutschland 11,0 pro 100 000 Einwohner bei Frauen und 15,9 pro 100 000 Einwohner bei Männern. Im Mittel waren betroffene Frauen bei der Diagnosestellung 73 und Männer 71 Jahre alt. Demgegenüber tritt der M. Hodgkin deutlich seltener in Deutschland auf. Die altersstandardisierte Erkrankungsrate betrug hier im Jahr 2020 2,3 pro 100 000 Einwohner bei Frauen und 3,3 pro 100 000 Einwohner bei Männern. Patienten, welche ein Hodgkin-Lymphom ausbilden, sind typischerweise um einiges jünger, als Patienten mit der Diagnose eines NHL. Das mittlere Erkrankungsalter betrug 2016 bei Frauen 43 und für Männer 48 Jahre. Darüber hinaus ist das 5- und 10-Jahres-Überleben bei M. Hodgkin im Vergleich zu Non-Hodgkin-Lymphomen deutlich erhöht (Tab. 1) [8].

International zeigt sich ein ähnliches Bild. Anhand der Daten der WHO unterstellten „International Agency for Research on Cancer“ (IARC) wurden im Jahr 2020 weltweit 544 352 Fälle von NHL und 83 087 des M. Hodgkin gemeldet. Damit rangieren die NHL auf Rang 11 der häufigsten Tumorerkrankungen weltweit, während M. Hodgkin auch weltweit eine weitaus geringere Häufigkeit aufweist (Rang 26) [12].

	M. Hodgkin		Non-Hodgkin-Lymphome	
Geschlecht	♀	♂	♀	♂
Neuerkrankungen	990	1460	8230	10090
Mittleres Erkrankungsalter	43	48	73	71
Altersstandardisierte Erkrankungsrate	2,3	3,3	11,0	15,9
Altersstandardisierte Sterberate	0,1	0,3	2,9	5,4
Relative Überlebensrate 5J	87	87	72	71
Relative Überlebensrate 10J	84	82	68	64

Tabelle 1 - Epidemiologische Gesichtspunkte der HL und NHL nach Daten des RKI aus dem Jahr 2020 in Deutschland [8].

International zeigt sich ein ähnliches Bild. Anhand der Daten der WHO unterstellten „International Agency for Research on Cancer“ (IARC) wurden im Jahr 2020 weltweit 544 352 Fälle von NHL und 83 087 des M. Hodgkin gemeldet. Damit rangieren die NHL auf Rang 11 der häufigsten Tumorerkrankungen weltweit, während M. Hodgkin auch weltweit eine weitaus geringere Häufigkeit aufweist (Rang 26) [9].

Dennoch ist ein großer geographischer Unterschied in der Verbreitung einzelner Subtypen von NHL feststellbar. Während die adulte T-Zell-Leukämie, welche eine Sonderform der T-Zell-Leukämie darstellt und durch eine Infektion mit HTLV-1 (Humanes T-lymphotropes Virus 1) hervorgerufen wird, besonders in Ostasien zu finden ist, ist das follikuläre Lymphom vermehrt in Westeuropa und Nordamerika verbreitet. Das diffus-großzellige B-Zell-Lymphom stellt jedoch länderübergreifend die häufigste Form der NHL dar. [10]

Auf Basis der aktuellen WHO-Klassifikation rangiert das diffus großzellige B-Zell-Lymphom (DLBCL) mit einem Anteil von 31% aller NHL in der westlichen Welt epidemiologisch an erster Stelle, gefolgt von dem meist niedrigmalignen follikulären Lymphom (22%). Das Mantelzelllymphom (6%) und das Burkitt-Lymphom (2%) stellen weitere häufige Formen hochmaligner NHL dar, die CLL (chronisch lymphatischen Leukämie) kommt samt ihrer Unterform SLL (kleinzelliges lymphozytisches Lymphom) auf insgesamt 6% [11]. 80% der NHL basieren auf Neoplasien der B-Zellreihe, während Neoplasien der T- und NK-Zellreihe mit 10% in der westlichen Welt deutlich seltener zu finden sind [12].

2.2.3 Ätiologie und Risikofaktoren

Die Ätiologie der NHL bleibt gegenwärtig noch weitestgehend unklar und ist Gegenstand intensiver Forschungsarbeit. Genetische Ursachen sollen in Kombination mit einer gestörten Funktion des Immunsystems, wie dies z.B. bei Zöliakie, dem Sjögren-Syndrom und verschiedenen weiteren Erkrankungen des rheumatoiden Formenkreises der Fall ist, die Entwicklung begünstigen [13, 14].

Der Einfluss ausgewählter spezifischen Faktoren auf einzelne Subtypen der NHL ist jedoch mittlerweile besser erforscht. So gilt der Einfluss einer Infektion mit dem Epstein-Barr-Virus (EBV) auf die Genese des epidemischen Burkitt-Lymphoms mittlerweile als gesichert [15], genauso wie die Assoziation zwischen einer Besiedlung mit dem Bakterium *H. pylori* im Magen und der Entstehung eines MALT-Lymphom [16].

Bestimmte chromosomale Translokationen tragen zur Genese eines NHL bei und wurden besonders häufig bei DLBCL, dem Burkitt-Lymphom, der chronisch lymphatischen Leukämie (CLL) und dem Mantelzelllymphom (MCL) beschrieben [17–21]. Die in der Vergangenheit vieldiskutierte Frage, ob eine berufsabhängige Exposition gegenüber Insektiziden die Auftretenswahrscheinlichkeit eines NHL signifikant erhöht, konnte auf Basis einer aktuellen, gepoolten Fall-Kontroll-Studie nur in der Gruppe der Organophosphate und Carbamate bejaht werden [22].

2.2.4 Klassifikation

Wie zuvor erwähnt wurde, erfolgt die Identifizierung der verschiedenen Lymphome in der Praxis auf Basis des offiziellen Katalogs „classification of lymphoid, histiocytic and dendritic neoplasms“ der WHO [7]. Anhand dieser Einteilung ist es möglich, die vorhandene Neoplasie auf Basis von histologischen und zytologischen Eigenschaften zu identifizieren. Die Klassifikation und Stadieneinteilung eines Lymphoms anhand der zugrunde liegenden PET-Bildgebung wird hingegen erst in Kap. 2.3 besprochen.

Wie der Name bereits suggeriert, beinhaltet die WHO-Klassifikationen nicht nur maligne Lymphome, sondern ebenso weitere neoplastische Syndrome. Darunter fallen neben Lymphomen unter anderem auch Sarkome und Paraproteinämien wie die „ $\alpha/\gamma/\mu$ heavy-chain diseases“, letztere treten jedoch häufig in Kombination mit einem malignen Lymphom als eine paraneoplastisches Syndrom auf [23]. Die betroffene Zelllinie spielt zur Einordnung die zentrale Rolle und anhand dieser lassen sich einzelne Lymphome übergeordneten Gruppen zuordnen. Die Gruppe der erwachsenen B-Zell-Neoplasien beinhaltet die meisten

verschiedenen Neoplasien, gefolgt von der Gruppe der erwachsenen T- und NK-Zell-Neoplasien [7].

Man unterscheidet zudem „indolente“ bzw. „niedrigmaligne“ Lymphome von „aggressiven“ bzw. „hochmalignen“ Lymphomen voneinander. Im Gegensatz zu der Einteilung nach betroffener Zellreihe stehen hier die klinischen Parameter und das durchschnittliche Patienten-Outcome bei der jeweiligen Tumorentität im Fokus. Indolente Lymphome wachsen anhand langjähriger klinischer Forschung und Erfahrung langsamer als Lymphome, welche als aggressiv klassifiziert werden. In der folgenden Tabelle werden häufige maligne Lymphome gemäß dieser Einteilung aufgelistet (Tab. 2) [24–26].

niedrigmaligne Lymphome	hochmaligne Lymphome
Follikuläres Lymphom Typ I-IIIa	Diffus, großzelliges B-Zell-Lymphom (DLBCL)
Marginalzonenlymphom	Follikuläres Lymphom Typ IIIb
Kutanes B- und T-Zell-Lymphom	Burkitt-Lymphom
Chronisch lymphatische Leukämie (CLL)	Lymphoblastisches Lymphom
MALT-Lymphom	Multiples Myelom

Tabelle 2 - Einteilung häufiger maligner Lymphome nach Wachstumsgeschwindigkeit und Klinik [25, 26].

2.2.5 Risikostratifizierung, Therapie und Prognose

Die Therapie eines NHL kann je nach Klassifikation und Patientenverfassung eine Strahlentherapie, Chemotherapie, eine Therapie mit künstlich hergestellten Antikörpern, eine Therapie mit Signalinhibitoren und eine Knochenmarkstransplantation beinhalten. Diese Therapieansätze können fallgesteuert einzeln oder in Kombination durchgeführt werden. Die vorliegende Arbeit befasst sich in erster Linie mit der Gruppe der aggressiven Non-Hodgkin-Lymphome, weshalb an dieser Stelle auf eine genaue Analyse der verschiedenen Therapiemethoden aller NHL verzichtet wird. Stattdessen sollen die etabliertesten Therapieansätze der aggressiven NHL im Folgenden behandelt werden.

2.2.5.1 Risikostratifizierung bei aggressiven Non-Hodgkin-Lymphomen

Aufgrund der weltweiten Häufigkeit des DLBCL ist die Erprobung neuer Therapieansätze ein Forschungsthema mit Versorgungsrelevanz. Die Therapie eines DLBCL hängt wie bei den weiteren Formen der NHL entscheidend von dem mittels Bildgebung diagnostizierten Stadium (I – IV) und der Prognose ab. Die Prognose kann bei DLBCL, sowie bei den weiteren aggressiven Lymphomentitäten mittels dem Internationalen Prognostischen Indexes (IPI) in

vier Risikogruppen (0-1, 2, 3, 4-5) stratifiziert werden [27]. Die Therapiegrundlage des initial entwickelten IPI im Jahr 1993 basierte auf eine solitäre Therapie mit Doxorubicin, einem Anthracyclin, welches durch Hemmung der Topoisomerase II die Nukleinsäuresynthese blockiert [28]. Die Validität des IPI konnte jedoch in der modernen Ära der kombinierten Immunchemotherapie mit Rituximab nochmals bestätigt werden [29]. Des Weiteren konnte die Validität eines altersstandardisierter IPI (aaIPI) für Lymphom-Patienten < 60 Jahren bestätigt werden [28].

IPI	R-IPI
0, 1: „low risk profile“	0: „very-good prognosis“
2: „low-intermediate risk profile“	1, 2: „good prognosis“
3: „high-intermediate risk profile“	3, 4, 5: „poor prognosis“
4, 5: „high risk profile“	

Tabelle 3 - Vergleich des IPI mit dem R-IPI

Bei der Stratifizierung gemäß Risikoprofil und Prognose wird bei dem IPI bzw. R-IPI jeweils ein Punkt für folgende klinische Risikofaktoren vergeben:

- Lebensalter > 60 Jahre
- Stadium III/IV laut Bildgebung
- mindestens ein extranodaler Befall
- erhöhtes LDH in der Blutkontrolle
- erniedrigter Allgemeinzustand (ECOG-Score $\geq$ 2)

Der „revised International Prognostic Index“, R-IPI, beinhaltet dieselben Items wie der reguläre IPI. Während der reguläre IPI jedoch auf vier Unterscheidungen in der Prognose zurückgreift, basiert der R-IPI nur noch auf drei [30] (Tab. 3).

In der Arbeit von *Mikhaeel et. al (2022)* wird die Risikostratifizierung mittels eines „International metabolic prognostic Index“ (IMPI) eingeführt [5]. Diese Risikostratifizierung stützt sich auf eine Berechnung mit Hilfe des metabolischen Tumervolumens (MTV), auf welches in Kap. 2.4 näher eingegangen wird.

2.2.5.2 Standardtherapie bei DLBCL

Generell sollte nach Ansicht verschiedener Autoren unabhängig vom Lebensalter ein kurativer Ansatz erfolgen, falls die Komorbidität des Patienten eine medikamentöse Therapie zulässt. Es kann jedoch beobachtet werden, dass der Anteil an kompletten Remissionen bei Patienten über 65 Jahre geringer ausfällt, als bei jüngeren DLBCL-Patienten [27]. Als ein Grund neben der

erhöhten Auftretenswahrscheinlichkeit von therapierelevanten Komorbiditäten im höheren Lebensalter, wird die erhöhte genetische Komplexität, also die Auftretenswahrscheinlichkeit genetischer Variationen mit schlechtem Outcome, gehandelt [31]. Darunter zählen unter anderem die Entwicklung eines DLBCL aus den ABC (aktivierten B-Zellen), einem Rearrangement im Bereich des BCL-2, BCL-6 und MYC-Genlokus, sowie IRF-4-Translokationen [31–33]. Da das Auftreten dieser genetischen Variationen mit dem Lebensalter korreliert, findet sich indirekt eine Referenz auf diese in der IPI-Risikostratifizierung, direkt finden sie jedoch noch keinen Eintrag im IPI.

Der Grundstein zur Behandlung maligner Lymphome und daher auch zur Behandlung des DLBCL wurde mit der Anwendung des „CHOP-Schemas“ gelegt. Dieses Schema beinhaltet die Medikamente Cyclophosphamid, Doxorubicin (auch: „Hydroxydaunorubicin“), Vincristin (Handelsname: „Oncovin“) und Prednisolon. Bis dato bildet das CHOP-Schema den Grundstein der medikamentösen Therapie zur Behandlung von NHL. Das CHOP-Schema wird bei CD20-positiven B-Zell-Lymphomen um den monoklonalen Antikörper „Rituximab“ ergänzt und nun im „R-CHOP“-Schema verabreicht. Die Wirkung dieses Kombinationspräparats konnte unter anderem schon im Rahmen der „RICOVER-60“-Studie nachgewiesen werden, sowohl bei jüngeren, als auch bei älteren Patientenkohorten [34, 35]. Der positive Einfluss einer additiven Strahlentherapie auf das Patienten-Outcome mit der Diagnose eines DLBCL und einem zusätzlichen Nachweis einer „bulky-disease“ (Tumormassen mit einem Durchmesser >7,5 cm) konnte in einer prospektiven Studie von *Held et. al* (2014) ebenfalls nachgewiesen werden [36].

Während bei älteren Patienten und bei Patienten mit erhöhtem Risikoprofil ein dosisreduziertes R-CHOP-Schema derzeit noch nicht als Standardtherapie Frage kommt, stellt das dosisreduzierte Schema mit vier Zyklen R-CHOP-14 (Therapieintervall von 14 Tagen) + zwei Zyklen Rituximab den neuen Therapiestandard bei jüngeren Patienten laut S3-Leitlinie dar [37]. Im Vergleich zum Standard-Therapiearm mit sechs Zyklen R-CHOP-21 konnte die Dosisreduktion eine verringerte therapieassoziierte Toxizität bei vergleichbarer Effektivität aufzeigen [27, 38]. Bei älteren Patienten werden hingegen gute Ergebnisse mit sechs bis acht Zyklen R-CHOP-14 erzielt, obwohl die OPTIMAL>60-Studie die Effektivität einer Therapiereduktion bei älteren Patienten untersucht. Zudem haben Patienten, welche im interim-PET nach drei Zyklen R-CHOP-14 negativ erscheinen, im Vergleich zu PET-Positiven eine gute Prognose [39].

2.2.5.3 Weitere Therapiemöglichkeiten bei aggressiven NHL verschiedener Entitäten

Die Basis für die Therapie der weiteren aggressiven Lymphome ist auch hier meist das CHOP-Schema mit zusätzlicher Gabe von Rituximab (R-CHOP) bei CD20-positiven B-Zell-Lymphomen. Die Prognose äußert sich je nach Tumorentität und IPI-Score als unterschiedlich, jedoch zählt das Burkitt-Lymphom wegen der hohen Wahrscheinlichkeit zur Ausbildung von extranodalen Manifestationen und einer äußerst schnellen Tumorverdopplungszeit (TVDT) zu den aggressivsten [40]. Nach Daten von *Perkins et. al* (2008) zeigten 30-40 % der Patienten bei der Erstdiagnose mit einem Burkitt-Lymphom einen ausgeprägten Knochenmarksbefall und bis zu 40 % einen leptomeningealen Befall [41]. Durch das hohe Risiko auf befallene ZNS-Strukturen spielt die Verringerung der ZNS-Rezidivrate eine große Rolle in der Therapiesteuerung, genauso wie es durch einen massiven Befall des Knochenmarks zu einer „Burkitt-Leukämie“ kommen kann [42]. Eine intensive Polychemotherapie mit Rituximab ist trotz Kontroversen durch mehrere Studien untermauert [43, 44]. Ein Hochdosis-Schema basierend auf Methotrexat, einem Folsäure-Antagonist, ist eine wirksame Alternative bzw. kann auch in verschiedenen Schemata in Kombination mit Rituximab verabreicht werden [44, 45]. Das reguläre R-CHOP-Schema wird beim Burkitt-Lymphom im Vergleich zu DLBCL als unzureichend eingestuft.

Die CAR-T-Zell-Therapie stellt eine zukunftssträchtige und derzeit noch außerordentlich ressourcenbindende Strategie zur Behandlung von B-Zell-Lymphomen dar, welche jedoch in Deutschland derzeit nur als Zweit- oder Drittlinientherapie bzw. bei Rezidiven in Frage kommt [37]. Hierbei werden autologe T-Zellen des Patienten entnommen und diese ex-vivo mithilfe eines synthetischen „chimeric antigen receptor“ (kurz: CAR) modifiziert. Nach der Modifikation erfolgt die Retransfusion und anhand der antikörperähnlichen CAR-Domäne können beispielsweise Zellen, welche CD19 exprimieren, angegriffen werden [27]. CD19 wird, im Vergleich zu CD20, quantitativ besonders häufig an der Zelloberfläche von malignen B-Zell-Lymphomen exprimiert und eignet sich daher gut als Angriffspunkt [46]. Zudem konnte sich „B cell maturation antigen“ (BCMA) beim Multiplen Myelom als Angriffsziel von spezifischen CAR-T-Zellen (bb2121) in klinischen Studien beweisen. Laut Ergebnissen der Arbeitsgruppe von *Raje et. al* (2019) konnte in der CAR-T-Zelltherapie mit dem Rezeptormolekül bb2121 („Idecabtagene Vicleucel“) ein Ansprechen auf die Therapie von 85% erreicht werden, eine Komplettremission erfolgte zudem bei 45 % der Patienten im Kollektiv [47]. Allogene hämatopoetische Stammzellentransplantationen (allo-SCT) können

bei Patienten mit ≥ 2 Rezidiven und nach durchgeführter CAR-T-Zelltherapie bzw. nicht durchführbarer CAR-T-Zelltherapie in Erwägung gezogen werden [37].

Eine allo-SCT ist besonders bei Patienten, welche eine Richter-Transformation (hochmaligne Transformation) einer CLL in eine DLBCL erleiden, angezeigt [48]. Bei zusätzlicher Gabe von Rituximab als Graft-versus-Host-Prophylaxe (GvH-D) konnten nach Daten von *Glass et. al* (2014) bei Patienten mit aggressiven B – und T-Zell-Lymphomen jedoch keine Unterschiede in der Inzidenz von der GvH-D und dem Gesamtüberleben festgestellt werden [49]. Zudem kann eine autologe Stammzelltransplantation (auto-SCT) in Kombination mit einer Hochdosis-Chemotherapie bei therapierefraktären NHL-Patienten als ein kurativer Therapieansatz verfolgt werden [50]. Derselbe Ansatz kann gleichermaßen bei therapierefraktären Patienten mit M. Hodgkin verfolgt werden [51].

2.3 Die ^{18}F -FDG-PET/CT in der Lymphomdiagnostik

Die gegenwärtig angewandte Klassifikationsform von Hodgkin-Lymphomen und Non-Hodgkin-Lymphomen im Zuge des initialen Stagings stellt mittlerweile die Lugano-Klassifikation, eine Weiterentwicklung der Ann-Arbor-Klassifikation aus dem Jahre 1971, dar [52, 53]. Grundsätzlich unterscheidet die Lugano-Klassifikation (Tab. 4), wie auch die zugrundeliegende Ann-Arbor-Klassifikation, zwischen limitierten (I + II) und fortgeschrittenen Stadien (III + IV). Auf eine Zuordnung von A („absence of“) und B („presence of“) Symptomen wird bei der Bewertung von NHL im Gegensatz zu Hodgkin-Lymphomen verzichtet [52]. Ein Sonderfall in der Lugano-Klassifikation stellt ein malignes Lymphom im Stadium II gemeinsam mit einem positiven Nachweis einer „bulky-disease“ dar, bei aggressiven NHL typischerweise mit einem einzelnen Tumordurchmesser von $> 7,5$ cm. Eine Einteilung, ob sich das Lymphom im limitierten oder fortgeschrittenen Stadium befindet, kann hier nur in Einbeziehung der Histologie und des IPI-Scores vorgenommen werden [53].

Stadium	Nodaler Status	Extranodaler Status
I	(1) ein befallener Lymphknoten, oder (2) eine befallene Gruppe benachbarter Lymphknoten	Einzelner extranodaler Befall ohne nodalen Befall
II	Zwei oder mehr befallene Lymphknoten limitiert auf eine Zwerchfellseite	Stadium I oder II im nodalen Befallsmuster + limitierter extranodaler Befall
II + „bulky disease“	Stadium II + „bulky disease“	Nicht anwendbar
III	(1) Befallene Lymphknoten auf beiden Zwerchfellseiten, oder (2) nodaler Befall über dem Zwerchfell + Milzbefall	Nicht anwendbar
IV	Stadium III + weitere extralymphatische Manifestationen	Nicht anwendbar

Tabelle 4 - Stadieneinteilung gemäß der Lugano-Klassifikation (modifiziert nach [53]). Tonsillen, der Waldeyer-Rachenring und Milz zählen hier als nodales Gewebe.

Die Lugano-Klassifikation hob die ^{18}F -FDG-PET/CT als Mittel der Wahl für die Darstellung von nodalen und extranodalen Befällen „FDG-avider“ Tumoren hervor. Bei FDG handelt es sich um das Molekül F-18-Fluorodeoxyglukose, ein durch Positronenemission radioaktives Glukoseanalogon, welches annähernd gleichmäßig mittels der Blutzirkulation im Körper verteilt wird. In Zellverbänden mit höheren Glukoseverbrauch wird FDG vermehrt durch D-Glukosetransporter aufgenommen und mittels Phosphorylierung durch die Hexokinase örtlich fixiert [3, 54]. Etwa 93-99% aller histologischen Subtypen von Lymphomen sind FDG-avide und daher für das Staging mittels FDG-PET/CT geeignet, während eine Reihe von nicht-follikulären, niedrigmalignen Lymphomen keine relevante bzw. nur mäßig relevante FDG-Anreicherung zeigen [55].

Die FDG-PET/CT wird bei der Erkennung von Knochenmarksinfiltrationen als durchaus vorteilhaft und sensitiv angesehen und löste im Zuge der Etablierung der Lugano-Klassifikation ein durch Knochenmarkbiopsie gesteuertes Staging bei Patienten mit den Diagnosen Hodgkin-Lymphom und DLBCL zunehmend ab [53]. Die Arbeitsgruppe um *Kaddu-Mulindwa et al.* (2019) berichtete, dass die alleinige Auswertung der FDG-PET/CT-Bildgebung eine signifikant niedrigere falsch-negativ-Rate aufwies als eine auf Knochenmarksaspiration gestützte Diagnostik [56]. Darüber hinaus attestierten zahlreiche Arbeitsgruppen der FDG-PET/CT eine deutlich höhere Sensitivität bei der Erkennung von nodalen und extranodalen Manifestationen als CT-Diagnostik allein, besonders bei hochmalignen NHL [53, 55]. Eine Ausnahme in der Gruppe der extranodalen Manifestationen stellen nach Meinung einiger Autoren

gastrointestinale Läsionen, besonders gastrische Manifestation, dar. Über die Genauigkeit bei der Erkennung dieser Läsionen existieren kritische Stimmen, ein solitäres Staging mit der FDG-PET/CT ohne Gastroduodenoskopie wird von einigen Autoren in der Effektivität in Frage gestellt [57].

2.4 Metabolische Biomarker der FDG-PET/CT

Als Maß für den Glukoseumsatz lassen sich verschiedene Biomarker aus der FDG-PET-Bildgebung ableiten, allen voran der „Standardized Uptake Value“ (SUV), der das Verhältnis der gemessenen Konzentration des radioaktiven Tracers zu einer hypothetischen gleichmäßigen Verteilung in der Körpermasse darstellt [3]. Hierfür wird die gemessene Radioaktivitätskonzentration im kalibrierten PET-Bild, die verabreichte Radioaktivitätsmenge und das Gewicht des Patienten benötigt. Die auf das Körpergewicht adaptierte Formel errechnet sich nach dem folgenden mathematischen Zusammenhang:

$$SUV = \frac{\left(\text{Aktivitätskonzentration} \left[\frac{kBq}{mL} \right] * \text{Körpergewicht} [kg] \right)}{\text{applizierte Aktivität} [MBq]}$$

Diese FDG-Anreicherungen im Körper, welche mit dem SUV ein quantifizierbares Maß in der PET-Bildgebung erhalten, können nun entweder manuell oder maschinell [58] konturiert werden. Um das daraus konturierte Tumolvolumen zu beschreiben, wurde der Begriff des MTV (Metabolic Tumor Volume) bzw. des oft in Studien synonymhaft verwendeten TMTV (Total Metabolic Tumor Volume) geschaffen. Das metabolische Tumolvolumen beschreibt die Tumorlast im Körper des Patienten [3]. Die Angabe von metabolischen Tumorumfängen erfolgt standardmäßig in Milliliter (mL). Daraus abgeleitet ergibt sich das quantitative Maß der TLG (Total Lesion Glycolysis), welches neben dem MTV auch die über das MTV gemittelte FDG-Stoffwechselaktivität (SUV_{mean}) miteinbezieht und sich aus dem folgenden Produkt errechnen lässt [54]:

$$TLG = SUV_{mean} * MTV [mL]$$

Das metabolische Tumolvolumen gewinnt als quantitativer Parameter für die Abschätzung der Prognose von Krebserkrankungen immer mehr an Bedeutung und gilt bei follikulären und hochmalignen Lymphomen als prognostischer Faktor [3]. Einige Studien konnten MTV als Prognosefaktor dezidiert beim diffus großzelligen B-Zell-Lymphom (DLBCL) nachweisen [59], hohes totales MTV soll laut *Vercellino et al.* (2020) des Weiteren auf das progressionsfreie Überleben (PFS) und das Gesamtüberleben (OS) großen Einfluss haben, trotz adäquater

Ansprache auf das Chemoimmuntherapieschema mit „R-CHOP“. Dies zeigt den prognostischen Wert von MTV, unabhängig davon, ob die Therapie anspricht oder nicht. So können laut *Vercellino et al. (2020)* Hochrisikopatienten künftig früher identifiziert werden [60]. MTV etabliert sich zudem sowohl bei niedrigmalignen, als auch bei hochmalignen follikulären Lymphomen als valider prognostischer Faktor [61]. Ein Vergleich von FDG-PET/CT-Staging und CT-basierten Staging bei dieser Tumorentität führte zu einem post-PET/CT-Upstaging von 45/150 Patienten vom limitierten Stadium (I-II) zum fortgeschrittenen Stadium (III-IV) [62]. Auch bei Hodgkin-Lymphomen zeigte sich ein prognostischer Einfluss von MTV [4, 63].

In einem kürzlich erschienenen systematischen Review von verschiedenen Studienergebnissen, welche den Zusammenhang zwischen MTV, TLG und SUV_{max} untersuchen sollen, kamen *Frood et al. (2021)* zum Schluss, dass bei Studien mit DLBCL-Patienten SUV_{max} in 5/20 Studien eine prognostische Rolle aufwies, MTV in 17/19 Studien und TLG in 10/13 Studien. Der Review gab jedoch auch zu bedenken, dass die verfügbaren Studien zum prognostischen Wert dieser bildbasierten Biomarkern retrospektiv und äußerst heterogen in ihrer Methodik seien und sämtliche Studienergebnisse ein moderates bis starkes Risiko für Verzerrungen aufwiesen [64].

2.4 Konturierung von Tumervolumina in der FDG-PET/CT

Zur Abgrenzung von Tumervolumina dienen Bildverarbeitungsverfahren mit spezifischen Kriterien zur Selektion und Konturierung von Lymphommanifestationen.

Der PERCIST-Schwellwert stellt einen etablierten Ansatz zur Selektion von metabolischen Tumervolumina dar. Dieser Schwellwert wird aus dem Mittelwert und der Standardabweichung der FDG-Anreicherung in einer Leber-Referenzregion (SUV_{mean} und SUV_{SD}) generiert. Lebernormierte PERCIST-Schwellwerte im PET werden folgendermaßen errechnet [65]:

$$PERCIST = 1,5 * SUV_{mean}(Leber) + 2 * SUV_{SD}(Leber)$$

Die so selektierten Herde können anschließend mit Schwellwerten konturiert werden. Mögliche Konturierungsmethoden des MTV unterscheiden sich teils stark in den errechneten Volumina. Während absolute SUV-Grenzwerte wie $SUV > 2,5$, $SUV > 4$ eine hohe Interrater-Reliabilität besitzen, zeigen relative Methoden wie „41% des tumoralen- SUV_{max} “ weniger Übereinstimmung zwischen unterschiedlichen Observern [66]. Darüber hinaus resultieren bei absoluten Grenzwerten meist höhere Werte für metabolische Tumervolumina, als dies bei dem relativen SUV_{max} (41 %) Grenzwert beobachtet wird. Diese Überschätzung des metabolischen

Tumolvolumens bezeichnen *Martín-Saladich et al.* (2020) bezogen auf die Reproduzierbarkeit als vorteilhaft [67]. Die Studie von *Meignan et. al* (2014) konnte mittels Tumorsimulationsphantome die Konturierungsmethode mittels tumoralen-SUV_{max} (41%) evaluieren und hohe Korrelationswerte in der Messung aufzeigen [68].

2.5 Definition der Forschungsfragen

Da das ermittelte metabolische Tumolvolumen durch verschiedene Auswertungsmethoden großen Schwankungsbreiten unterworfen ist, sind Studienergebnisse aufgrund ihrer unterschiedlichen Methodik meist nur begrenzt miteinander vergleichbar. Ziel dieser Arbeit ist es daher, auf Basis der zur Verfügung gestellten Bilddatensätze der deutschlandweiten OPTIMAL>60-Studie die MTV-Werte auf Basis der etablierten Selektion über PERCIST-Schwellwert und Konturierung anhand von 41% des SUV_{max} zu erzeugen und diese im Vergleich zu bereits bestehenden Ergebnissen eines adaptiven Konturierungsverfahrens zu evaluieren. Folgende Forschungsfragen können formuliert werden:

- (1.) Zeigen die hier angewandten Konturierungs- und Selektionsmethoden eine hohe Korrelation und Reproduzierbarkeit?
- (2.) Zeigen die individuell kalkulierten PERCIST-Schwellwerte Standort- und/oder PET-Scannermodell-abhängige Unterschiede?

Da sich methodenübergreifend die Evidenzlage zu MTV als zukunftssträchtiger prognostischer Faktor beim NHL zunehmend erhärtet, soll speziell die Korrelation der ermittelten PET-Biomarker mit der Risikostratifizierung auf Basis der IPI-Items untersucht werden:

- (3.) Zeigen die kalkulierten metabolischen Tumolvolumina in der initialen PET je nach vorangegangener Einschätzung des Risikoprofils signifikante Unterschiede?

3. Material und Methoden

3.1 Die OPTIMAL>60-Studie

Die zugrundeliegenden Daten für diese Dissertation stammen aus der OPTIMAL>60-Studie (NCT01478542). Diese Studie ist eine deutschlandweite, multizentrische, prospektive Studie an insgesamt 1152 Patienten mit aggressivem NHL, bei der in einem randomisierten Studiendesign das Ansprechen auf die medikamentöse Therapie mittels der FDG-PET/CT bewertet und durch Therapiereduktion optimiert werden soll. Die Studie evaluiert die Effektivität des experimentellen Therapieschemas, welches besonders darauf abzielte, durch hohe Therapieindividualisierung für Patienten ab 60 Jahren eine Prognoseverbesserung zu erzielen [69]. Neben der Kohorte der „favourable“ NHL-Patienten (IPI=1) existiert zudem die Patientenkohorte der „less favourable“-Patienten, welche einen IPI>1 besitzen. Als Sponsor der OPTIMAL>60-Studie fungiert die Saarland Universität in Saarbrücken [70]. Zum Zeitpunkt der Auswertungen im Rahmen dieser Dissertation war das Datenembargo der OPTIMAL>60-Studie in Kraft, sodass hier keine zusätzliche Aussage über die Korrelation der metabolischen Tumormarker mit der OS/PFS möglich war.

3.1.1 Ein- und Ausschlusskriterien

Der histologische Nachweis eines CD20-positiven B-NHL stellt den Grundstein für eine Rekrutierung in die OPTIMAL>60-Studie dar. Die Probenentnahme sollte jeweils über eine Exzisionsbiopsie eines befallenen Lymphknotens erfolgen, oder ein anderer als adäquat deklarierten Probengewinn aus einem nodalen bzw. extranodalen Befall des B-NHL erfolgen. Des Weiteren waren ein Lebensalter von 61 bis 80 Jahren, ein gemessener ECOG-Score von 0 bis 2 nach Applikation der Vorphasentherapie (100 mg/d Prednison oder Prednisolon), eine schriftliche Einverständniserklärung des Patienten und einen von Studienzentrum und Sponsor unterschriebenen Teilnahmevertrag notwendig. Die in der OPTIMAL>60-Studie therapierbaren, hochmalignen Entitäten an B-NHL stellen die Basis für das FDG-PET/CT-Bildmaterial dieser Arbeit dar. Laut Studienprotokoll konnte ein Einschluss in die Studie auf Basis der WHO-Klassifikation von 2008 bei diesen histologischen Befunden erfolgen [70]:

- Follikuläres Lymphom IIIb
- DLBCL: nicht näher bezeichnet (NOS), „T-cell/histiocyte-rich BCL“, „primary cutaneous DLBCL, leg type“, „EBV-positive DLBCL of the elderly“, DLBCL assoziiert mit chronischer Inflammation
- Primär mediastinales großzelliges B-Zell-Lymphom (PMBL)

- Intravaskuläres großzelliges B-Zell-Lymphom (IVLBCL)
- ALK-positives großzelliges B-Zell-Lymphom (ALK LBCL)
- Plasmoblastisches Lymphom
- Primäres Effusionslymphom (PEL)
- Transformiertes indolentes Lymphom/simultanes hochmalignes B-Zell-Lymphom
- „B-cell lymphoma, unclassifiable, with features intermediate between diffuse large B-cell lymphoma and Burkitt lymphoma“
- “B-cell lymphoma, unclassifiable, with features intermediate between diffuse large B-cell lymphoma and Hodgkin lymphoma“

Die Ausschlusskriterien beschäftigen sich mit ernststen unerwünschten Nebenwirkungen der medikamentösen Immunchemotherapie, sodass ein Therapieabbruch für das Patientenwohl unabdingbar war und schwerwiegende internistischen Komorbiditäten. Inkludiert dabei sind Herz (Angina Pectoris CCS>2, Herzinsuffizienz NYHA>2 und/oder EF<50%/FS<25%), Lunge (FeV₁<50 % und/oder DLCO<50 %), Niere (Kreatinin>2x höher als der Referenzwert) und Leber (Bilirubin>2x höher als Referenzwert und/oder AST/ALT>3x höher als Referenzwert).

Des Weiteren sind Patienten mit primären oder sekundären tumorösen ZNS-Befall ausgeschlossen, sowie Patienten, welche bereits eine Chemotherapie erhalten bzw. erhalten haben, mit Ausnahme der in der OPTIMAL>60-Studie integrierten Vorphasentherapie [70].

3.2 Bereitstellung der initialen FDG-PET/CTs

Alle in dieser Arbeit verwendeten PET-Aufnahmen wurden im Rahmen der OPTIMAL>60-Studie angefertigt und zur protokollgemäßen Auswertung der hier untersuchten bildbasierten Biomarker bereitgestellt. Der Großteil besteht aus kombinierten PET/CT-Aufnahmen, während einzelne Bilddatensätze auch als solitäre PET bzw. als PET/MRT in die Auswertung eingeschlossen wurden. Als Tracer fungierte in der PET-Aufnahme ausnahmslos F-18-Fluorodeoxyglukose, welches gemäß den Stammdaten der ausgewerteten PET-Bilder im Durchschnitt mit einer Aktivität von 263 ± 72 MBq (n=661) intravenös appliziert wurde. Als Maßnahme zur Standardisierung der Zeitpunkte der PET-Aufnahmen wurde eine Frist von sechs Wochen vor Therapiebeginn festgelegt, in welcher die Durchführung der initialen FDG-PET-Aufnahme stattzufinden habe. Diese Frist inkludiert die Vorphasentherapie. PET-Studienzentren müssen des Weiteren die Empfehlungen der ACRIN (American College of Radiology Imaging Network) erfüllen. Die Rekrutierung von PET-Scannern ohne integrierter CT-Funktion wurde mit 01.07.2014 ausgesetzt. Zudem sollte laut OPTIMAL>60-

Studienprotokoll eine Weiterleitung der FDG-PET/CT-Datensätzen innerhalb von zwei Arbeitstagen nach Aufnahmetermin in einem gängigen Format (DICOM, ECAT) an das Referenzzentrum erfolgen [70].

Die hier durchgeführten Analysen sind Teil der protokollgemäßen Auswertung der Daten und durch die Ethikkommission der Ärztekammer des Saarlandes berufsrechtlich geprüft und positiv votiert worden. Insgesamt wurden im Rahmen dieser Dissertation 661 verschiedene PET-Aufnahmen von deutschlandweit 83 verschiedenen PET-Scanner-Standorten und 59 übergeordneten Therapiezentren ausgewertet. Eine Auflistung der Therapiezentren und die dazugehörigen Zentrums-IDs findet sich in nachfolgender Tabelle (Tab. 5). Auf die in der OPTIMAL>60-Studie verwendete Pseudonymisierung der Therapiezentren wurde zur Aufrechterhaltung des Datenschutzes verzichtet, da eine geographische Zuordnung der zugehörigen Klinika daraus möglich war. Stattdessen werden diese hier mit einer Alternativbezeichnung referenziert. Einen Überblick über die Datengrundlage dieser Dissertation bietet folgendes STARD-Diagramm (Abb. 1).

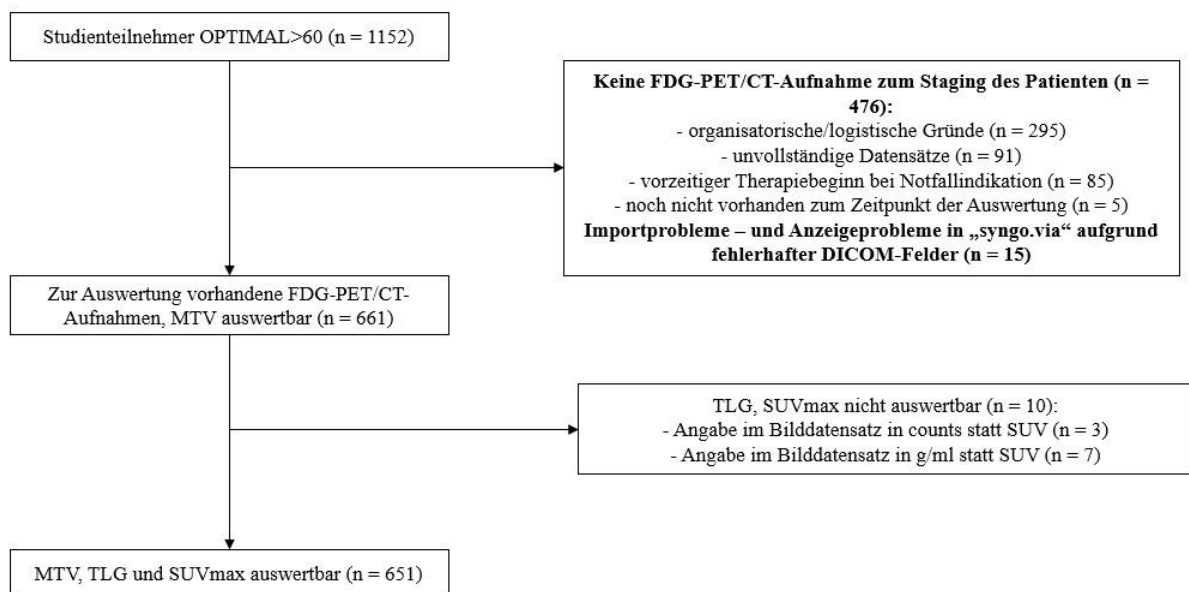


Abbildung 1 – STARD-Diagramm der ausgewerteten Bilddatensätze aus der OPTIMAL>60-Studie

Pseudonym	Zugehörige Zentrums-IDs	Pseudonym	Zugehörige Zentrums-IDs
1	DE-042	31	DE-024, DE-421
2	DE-072, DE-726, DE-752	32	DE-059, DE-381
3	DE-405	33	DE-362
4	DE-319	34	DE-453
5	DE-489, DE-724	35	DE-337
6	DE-090	36	DE-370
7	DE-278	37	DE-277, DE-372, DE-750
8	DE-141	38	DE-129
9	DE-701, DE-747	39	DE-125, DE-144
10	DE-106, DE-412	40	DE-012
11	DE-420, DE-507, DE-536	41	DE-107, DE-727
12	DE-440	42	DE-043
13	DE-369	43	DE-355
14	DE-403	44	DE-027
15	DE-347	45	DE-153, DE-230
16	DE-221	46	DE-697
17	DE-415, DE-422	47	DE-218
18	DE-498	48	DE-493
19	DE-061	49	DE-341
20	DE-034, DE-480	50	DE-434
21	DE-232	51	DE-510
22	DE-070	52	DE-096
23	DE-046	53	DE-509
24	DE-401, DE-739	54	DE-039, DE-052
25	DE-058	55	DE-505, DE-737
26	DE-045	56	DE-017
27	DE-073, DE-113, DE-353, DE-400, DE-482	57	DE-375
28	DE-223, DE-224	58	DE-314
29	DE-210	59	DE-089
30	DE-464		

Tabelle 5 - Übergeordnete Therapiezentren anhand der Nomenklatur des OPTIMAL>60-Studienprotokolls.

3.3 PET-Befundungskriterien

Unterschiedliche Kriterien zur Ermittlung der metabolischen Tumervolumina führen zu einer geringer werdenden Vergleichbarkeit einzelner Studienergebnisse untereinander. Der in dieser Dissertation implementierte Algorithmus zur Ermittlung von SUV-Schwellwerten stützt sich vorrangig auf die international erprobten PERCIST-Kriterien, welche sich an den jeweilig gemessenen SUV_{mean} -Werten der Leberreferenzregion im PET-Bild orientierten. Der Befund

einer Läsion wurde in der initialen FDG-PET/CT im Rahmen dieser Dissertation gestellt, falls alle der folgenden Kriterien erfüllt wurden:

- (1.) Der gemessene SUV_{max} der tumorsuspekten Region ist höher bzw. gleich hoch wie der ermittelte PERCIST-Schwellwert, welcher sich aus SUV_{mean} und SUV_{SD} der gesetzten Leberreferenzregion zusammensetzt.
- (2.) Die tumorsuspekte Region weist in der zumeist programmgesteuerten, relativen Konturierungsmethode SUV_{max} (41%) ein Volumen von zumindest 1 mL auf. Durch den Untersucher als hochsuspekte klassifizierte Läsionen mit einem Volumen unter 1 mL flossen nicht in die Kalkulation des MTV ein.
- (3.) Das FDG-anreichernde Areal muss vom jeweiligen Untersucher als höchstwahrscheinlich maligne bewertet werden.

Die Lokalisation der Leberreferenzregion wurde in dem Großteil der PET-Bilder automatisch zentral im rechten Leberlappen generiert. In Ausnahmefällen musste nach fehlerhafter Lokalisation durch die Analysesoftware (Kap. 3.5) die Leberreferenzregion jedoch manuell korrigiert werden. Das Volumen der Referenzregion beträgt als Kugel konstant 14,1 mL, während die mediastinale Referenzregion in der Aorta ein konstantes Volumen von 1,57 mL in Zylinderform besitzt.

Die Entscheidung, ob eine FDG-Anreicherung im PET als potenziell tumorös zu klassifizieren ist, wurde im Zuge dieser Dissertation dem Untersucher überlassen. Zusätzlich zu den oben definierten Kriterien existieren im DSHNHL-2009-1-Studienhandbuch zur OPTIMAL>60-Studie PET-Interpretationskriterien, welche auf den Empfehlungen zur PET-Befundung der Europäischen Vereinigung für Nuklearmedizin (EANM) basieren und auf diese sich bei etwaigen Interpretationsunsicherheiten gestützt wurde [71]. Diese unterstützen die Entscheidung empirisch häufig beobachtete FDG-Anreicherungen beim Gesunden sowie bei NHL-Patienten zu beurteilen. Generell gelte gemäß den EANM-Empfehlungen laut *Boellaard et. al* (2010) in allen PET-positiven Befunden, welche mittels einer visueller Beurteilung ermittelt wurden, dass eine sich vom Hintergrund abhebende, fokale bzw. diffuse FDG-Anreicherung existiert, welche nicht anatomisch/physiologisch erklärbar ist [71]. Zu den weiteren PET-Interpretationskriterien, welche in dieser Dissertation angewendet wurden, zählt, dass eine nodale FDG-Anreicherungen bei Existenz einer CT bzw. einer MRT ein Korrelat in eben dieser Bildgebung benötige, bzw. erst ab einem Durchmesser von 2 cm ohne sichtbares CT-Korrelat als maligne zu klassifizieren war. Zudem gäbe es die Empfehlung, etwaige

symmetrische bzw. asymmetrische FDG-Anreicherungen im Bereich des Mediastinums bzw. angrenzend an den Lungenhilus nur als PET-positiv zu werten, wenn in der restlichen Analyse eine weitere hochsuspekte Läsion entdeckt wurde. Visuell erkenntliche, fokale bzw. multifokale Mehranreicherung von FDG im Knochenmark führte ebenfalls zu einem PET-positiven Befund in der initialen PET [71].

Zusätzlich führten residuale Läsionen der Leber/Milz bei einem deutlich erhöhten SUV_{max} bei Existenz eines CT/MRT-Korrelats auch als einzige suspekte Region im PET-Bild zu einem positiven Befund.

Bei Unklarheiten im Bildbefund, ob eine FDG-Anreicherung als Tumor zu klassifizieren ist, wurden der Randomisierungsunterlagen des jeweiligen Falls konsultiert, auf eine standardmäßige Analyse der Randomisierungsunterlagen vor dem Befunden jedes PET-Bilddatensatzes wurde jedoch verzichtet.

3.4 Ablauf und Qualitätskontrolle der FDG-PET/CT-Untersuchung

Die initialen FDG-PET/CT-Aufnahmen im Rahmen der OPTIMAL>60-Studie wurden mit Verweis auf das DSHNHL-2009-1-Handbuch, welches sich eng an den Guidelines der Europäischen Vereinigung für Nuklearmedizin (EANM) orientiert, durchgeführt. Auf Basis dieser Guidelines unterlagen sie damit einer Reihe von standardgemäßen Qualitätskriterien und Minimalstandards [70, 71].

3.4.1 Vorbereitung des Patienten

Gemäß dem DSHNHL-2009-1-Handbuch zur FDG-PET/CT-Untersuchung ist die Vorbereitung des Stoffwechsels des Patienten essenziell, weshalb eine Fastenzeit von zumindest vier Stunden vor der FDG-PET/CT-Untersuchung notwendig ist, Wasser darf innerhalb dieser Fastenzeit jedoch getrunken werden. Zudem wird eine Hydratation von 1 L zwei Stunden vor Untersuchungsbeginn empfohlen. Eine etwaige notwendige intravenöse Gabe von Flüssigkeit während der Fastenperiode sollte lediglich durch eine Infusion von isotonen Infusionslösungen erfolgen. Nach intravenöser Gabe der FDG sollte der Patient 0,5 L Wasser trinken, oder falls das nicht möglich ist, 0,5 L einer isotonischen Infusionslösung erhalten. Die Bestimmung des Blutzuckers vor der Gabe von FDG ist Pflicht und der gemessene Wert der Blutglukose muss vor Gabe 150 mg/dl unterschreiten. Bei erhöhtem Blutzucker ist die Administration von kurzwirksamem Insulin während der Fastenperiode möglich, nach einer Insulingabe sollte die Gabe von FDG jedoch um weitere 45 Minuten verschoben werden. Die Größe und das Gewicht des Patienten sollten im Standardverfahren vor Untersuchungsbeginn

erfasst werden. Um die FDG-Anreicherung im braunen Fettgeweben zu reduzieren, welche besonders in der Zervikal – und Nackenregion zu FDG-Anreicherungen führen kann, sollte der Patient ausreichend warmgehalten werden. Eine Applikation von Betablockern ist bei Nichtvorhandensein von etwaigen Kontraindikationen zusätzlich zu erwägen [71].

Zusätzlich ist gemäß dem DSHNHL-2009-1-Handbuch die Gabe von bis zu 20 mg Furosemid möglich, um den FDG-Uptake im Urogenitaltrakt zu reduzieren, zudem sollte die Blase des Patienten vor der FDG-Gabe möglichst entleert worden sein. Die Gabe von milden Sedativa bei auftretender Klaustrophobie ohne Kontraindikationen ist möglich.

3.4.2 Standards der PET-Bildakquisition

Das DSHNHL-2009-1 gibt dezidierte Empfehlungen ab, wie die Bildakquisition bei PET bzw. bei PET/CT und PET/MRT-Aufnahmen zu erfolgen hat [71]. Diese Methodik stellte daher die standortübergreifende Grundlage für die Gewinnung der in dieser Dissertation ausgewerteten PET-Bilddatensätze dar. Attenuationskorrigierte PET bzw. PET/CT oder PET/MRT mussten für die anschließende Analyse akquiriert werden. Bei fehlendem Verdacht auf ZNS – und Knochenmarksbeteiligung war eine Aufnahme kranial beginnend unter der Orbita bis zu den Oberschenkeln kaudal akzeptierter Standard. Bei hinreichendem Verdacht auf eine ZNS bzw. Knochenmarksbeteiligung wurde das Abbildungsspektrum um weitere Areale, wie z.B. um ein gesondertes PET/CT bei einer Hautinfiltration des Non-Hodgkin-Lymphoms am Fuß, erweitert. Eine Bildgebungsverzögerung von 45 bis 120 Minuten nach intravenöser Gabe des FDG wurde angestrebt, wobei 60 Minuten nach der intravenösen FDG-Gabe als idealer Untersuchungszeitpunkt deklariert wurde.

3.5 Computergestützte Auswertung mithilfe des Programms syngo.via

Die Auswertung der zur Verfügung gestellten PET-Bilddatensätze erfolgte in dieser Dissertation ausschließlich mit dem von der Siemens Healthineers AG entwickelten Programm syngo.via. Bei dem Programm syngo.via handelt es sich um eine multimodale Befundungslösung auf einer Client-Server-Plattform, welches ein umfassendes Spektrum für die molekulare Bildgebung in der Onkologie, Kardiologie und Neurologie bietet. Mittels der Option den gewählten PET-Bilddatensatz mit der Darstellungsvariante „MM Onkologie“ zu öffnen, konnten die verschiedenen Bilddatensätze geöffnet und bearbeitet werden.

Innerhalb der nun geöffneten Bilddarstellung erhält man mit der Menüoption „Details zu Bildbefunden“ einen detaillierten Überblick über technische Details der PET-Bildaufnahme, welche von dem zugehörigen PET-Standort angegeben wurden. Diese umfassen bei

Vollständigkeit neben dem Hersteller bzw. dem Modellnamen des PET-Scanners und der angewandten Bildrekonstruktionsmethode unter anderem Aufnahmezeitpunkt, die Aktivität des Radiopharmakons in MBq und das Gewicht des Patienten. Ein großer Teil dieser Bildbefunddetails wurde zusätzlich zu den Ergebnissen der Auswertung routinemäßig in einem standardisierten CRF-Bogen bei jeder Auswertung mitdokumentiert (CRF-Bogen in Kap. 9) bzw. als CSV-Datei zusätzlich exportiert. Die typischerweise gewählte Darstellungsform der FDG-PET/CT in „MM Onkologie“ für die Auswertung wird zudem in folgender Abbildung dargestellt (Abb. 2). Alle Patientendaten wurden zuvor hinsichtlich Identität von der Studienkoordination pseudonymisiert, und mittels des Schemas „DE – PET-Standort – Patientenummer – 2009 – 1“ zugeordnet und auf die Systemdatenbank des Uniklinikum Regensburg hochgeladen.

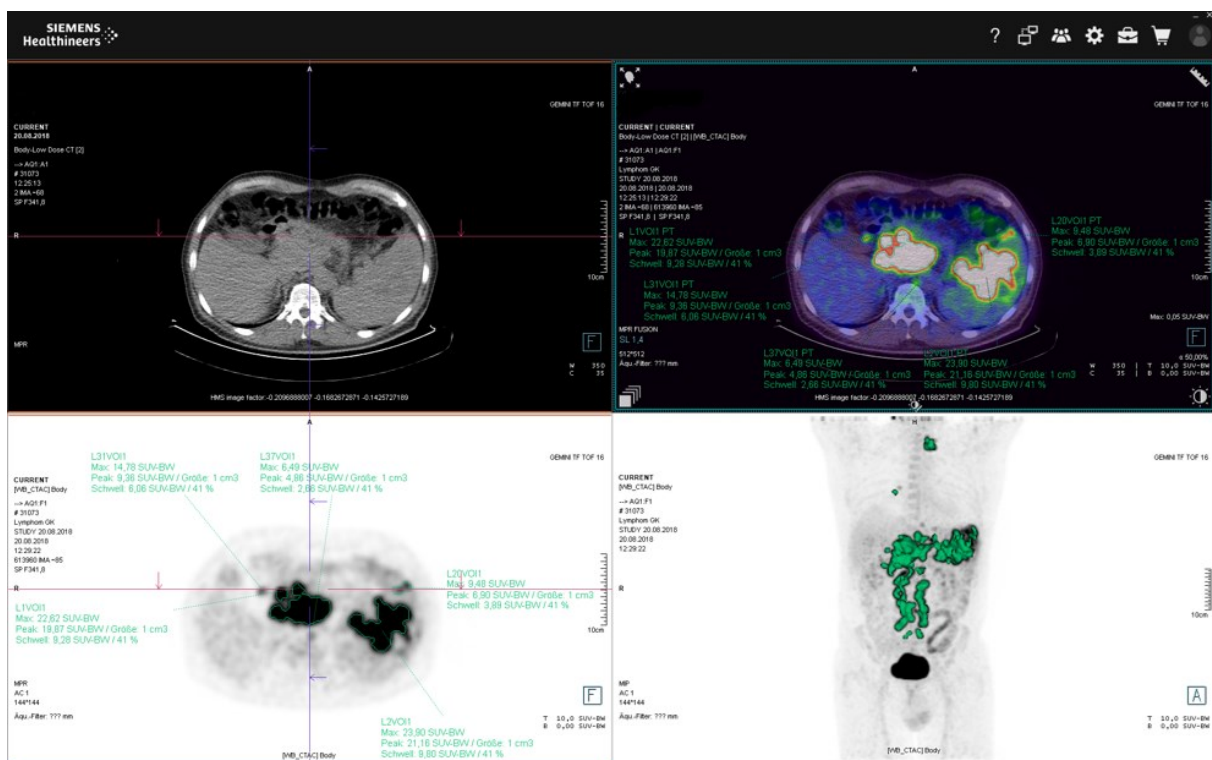


Abbildung 2 - Typische Darstellung der PET/CT bei der Auswertung, SUV-Skala: 0-10. links oben: low-dose CT, rechts oben: fusioniertes PET/CT-Bild, li.unten: PET, rechts unten.: Maximum-intensity-projection (MIP). Die in der MIP in Grün hinterlegten Volumina stellen das nach dem PERCIST-Schwellwert evaluierte metabolische Tumervolumen (MTV) dar.

Die Software syngo.via erlaubt das automatisch-generierte Setzen der Referenzregionen der Leber und der Aorta in „MM Onkologie“. Hierzu besteht die Möglichkeit mit Hilfe „multifokale Segmentierung“ (syngo.via-Version V06.04.0000.0050) bzw. „Läsions-Scout“ (syngo.via-Version V08.01.0000.0000) die jeweiligen Referenzregionen automatisch zu generieren. Die Volumina der Referenzregionen (VOIs) sind dabei konstant. Diese Option funktionierte im Rahmen der PET/CT-Auswertungen größtenteils problemlos, in Einzelfällen

jedoch erkannte der Algorithmus die korrekte anatomische Stelle nicht und die Position der Referenzregion musste manuell korrigiert bzw. gänzlich manuell gesetzt werden. Mit der „Läsions-Scout“-Funktion konnte im Zuge der Auswertungen der PET-Bilddatensätze eine gesteuerte Vorauswahl gemäß der PERCIST-Schwelle getroffen werden. Dieser beachtete in seiner Vorauswahl sowohl den von der Leberreferenzregion abhängigen PERCIST-Schwellwert als auch das Mindestvolumen von ≥ 1 mL und konnte bis zu 50 eigenständige Läsionen erkennen. Eine Überschreitung der 50 Läsionen war mit dem „Läsions-Scout“ nicht möglich, wobei dies mit einer passenden Informationsbox des Programms begleitet wurde. Die Auswahl des „Läsions-Scouts“-Algorithmus wurde im Anschluss vom Untersucher evaluiert, korrekt tumorös-verdächtige Läsionen übernommen, fehlende hinzugefügt und fehlerhafte bearbeitet bzw. entfernt. Die Konturierung erfolgte immer mit der relativen Konturierungsmethode 41% des SUV_{max} der ausgewählten Läsion (Abb. 3).

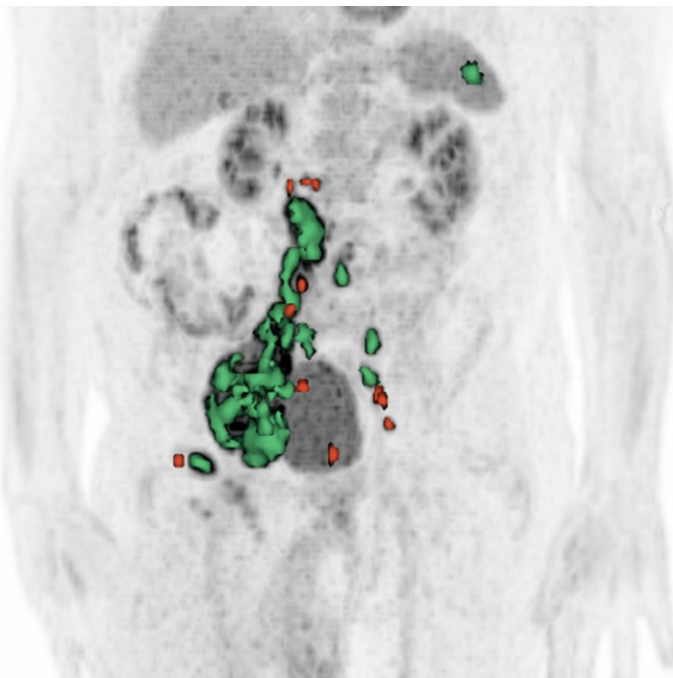


Abbildung 3 - MIP in Ventralansicht, SUV-Skala: 0-10. Exemplarisch sind in grün alle dem PERCIST-Schwellwert und Mindestvolumen gerechten Läsionen dargestellt, während in rot tumorverdächtige Läsionen hinterlegt sind, welche nicht dem Mindestvolumen/dem PERCIST-Schwellwert gerecht werden. Nur die in grün hinterlegten Läsionen wurden in die Berechnung MTV miteinbezogen.

Um eine adäquate Stichprobengröße zu gewährleisten, wurden im nachfolgenden Ergebnisteil im Rahmen des PERCIST-Schwellwerts und SUV_{mean} -Vergleichs der verschiedenen PET-Standorte und PET-Scannermodelle nur jeweils Standorte bzw. PET-Scannermodelle eingeschlossen, welche mehr als 20 verschiedenen Fällen beinhalteten. Der Export der PET-Scannerdaten aus syngo.via ergab bei einem der in der Analyse hinzugefügten Modelle eine unwesentlich voneinander abweichende Bezeichnung. Durch eine fast vollständige

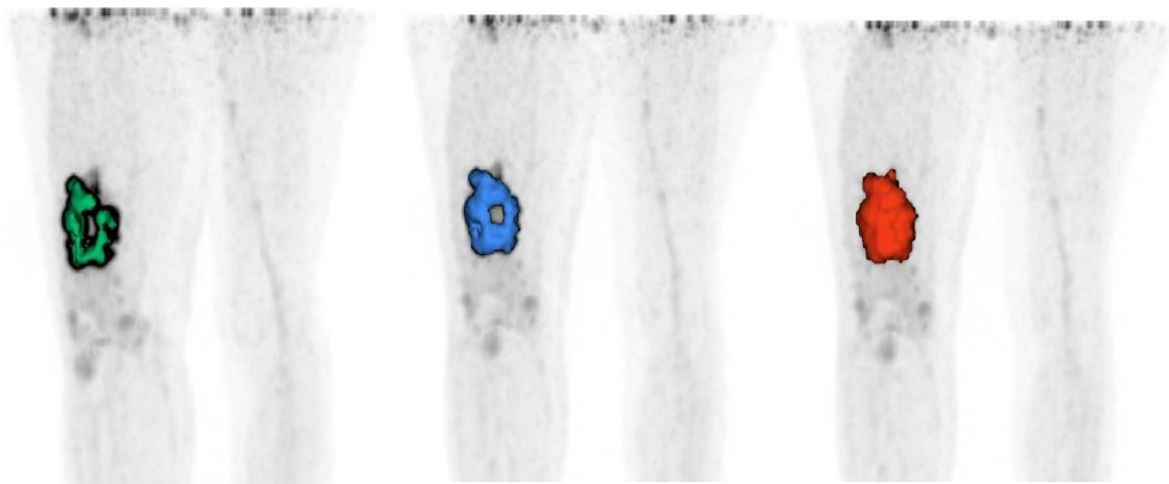
Übereinstimmung der ausgegebenen Bezeichnung konnten diese für die Analyse jedoch problemlos zusammengefügt werden. Als exemplarisches Beispiel dient hier der Fall der beiden Bezeichnungen „Biograph 40“ und „Biograph 40_mCT“ für das verwendete PET-Scannermodell, welche in den cmv-Dateien ausgegeben wurden. Diese wurden in der Analyse als „Biograph 40“ zusammengefügt.

Eine visuelle Beurteilung der bestimmten MTV ist ebenfalls im Auswertungsmechanismus integriert und beinhaltet einen semiquantitativen 5-stufigen Score. Zwecks des Versuches einer Vereinheitlichung dieses von Natur aus subjektiven Scores wurde vom Autor selbst folgende Richtlinie definiert (Tab. 6).

(-2)	MTV wirkt um mehr als 50% unterschätzt
(-1)	MTV wirkt um 11-50% unterschätzt
(0)	MTV $\pm$ 0-10% im Sollwert
(1)	MTV wirkt um 11-50% überschätzt
(2)	MTV wirkt um mehr als 50% überschätzt

Tabelle 6 - Prinzip des visuellen, semiquantitativen Scorings

Die hier gewählte Konturierungsmethode SUV_{max} (41%) ist eine von vielen Möglichkeiten der Konturierung von Lymphomherden. In der nachfolgenden Abbildung (Abb. 4) wird der Unterschied in den kalkulierten MTV exemplarisch dargestellt.



*Abbildung 4 - MIP einer FDG-PET der Beine, SUV-Skala: 1-10. Hier erfolgte exemplarisch die Konturierung einer Läsion mit einem hohen lokalen SUV_{max} , dieser lag in jedem dieser drei Aufnahmen bei 42,1.
Links: kalkuliertes MTV in grün bei SUV_{max} (41%)-Methodik: 33,9mL
Mitte: kalkuliertes MTV in blau bei SUV_{max} (20%)-Methodik: 87,2mL
Rechts: kalkuliertes MTV in rot bei $SUV_{\geq 4}$ -Methodik: 151,9mL*

Das Setzen der anatomischen Referenzregion erfolgte standardmäßig mit der Option „multifokale Segmentierung“ bzw. „Läsions-Scout“ in syngo.via, bei Leberinfiltrationen in die Referenzregion wurde jedoch eine Korrektur durch den Untersucher erforderlich. Teilweise war eine Korrektur durch das konstante Volumen von 14,1 mL jedoch nicht gänzlich möglich, so wie in folgender Abbildung dargestellt. (Abb. 5).

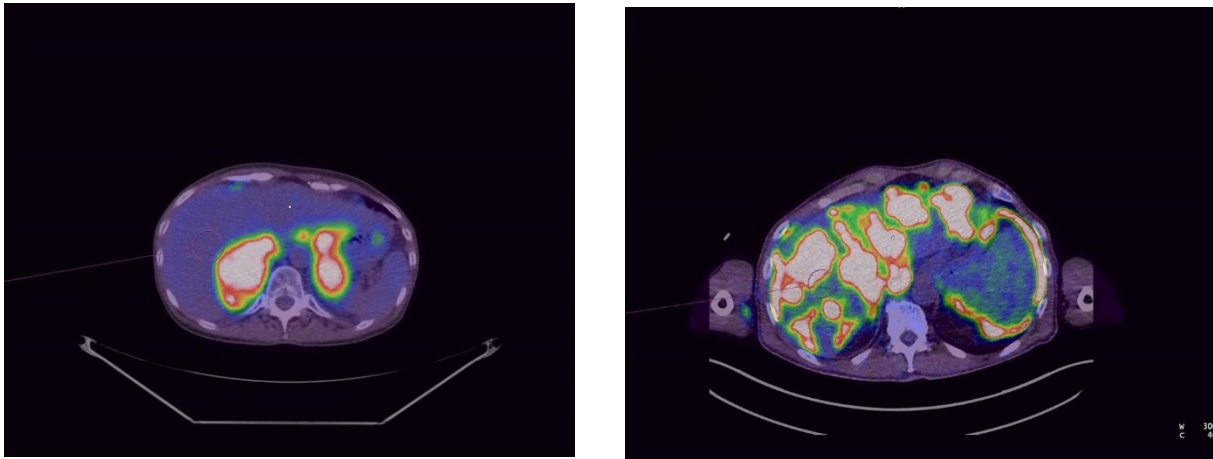


Abbildung 5 - Vergleich verschiedener "spill-over"-Situationen in der fusionierten FDG-PET/CT, SUV-Skala: 0-10 links: leicht korrigierbarer spill-over in die Leberreferenzregion. rechts: schwer korrigierbarer spill-over durch massive Infiltration der Leber. In Magenta ist jeweils die Position der Leberreferenzregion nach dem automatischen Setzen durch „syngo-via“ dargestellt.

3.6 Methodik der Interobserveranalysen

Zusätzlich dazu wurden im Rahmen dieser Dissertation zwei Interobserver-Analysen zur durchgeführt.

3.6.1 Methodik der ersten Interobserveranalyse mit dem Programm syngo.via

In der ersten Interobserver-Analyse wurden mittels einer repräsentativen Stichprobe von 60 der insgesamt 661 Fälle ausgewählt, welche von einem weiteren Observer mit exakt denselben PERCIST-Schwellwerten und Programm ausgewertet wurden. Das Selektionsverfahren mit dem Mindestwert des Volumens der Läsion von 1 mL und der Erfüllung des PERCIST-Schwellwerts war bei beiden Observer ident, eine Konturierung mittels 41% des SUV_{max} der Läsion wurde ebenfalls bei beiden Observer ident durchgeführt.

3.6.2 Methodik der zweiten Interobserver- und Interprogrammanalyse mit „syngo-via“ und „ROVER“

Eine zweite Interobserveranalyse wurde mit zur Verfügung gestellten Daten durchgeführt, welche mit dem Programm „ROVER“ (Version: 3.0.58h) des Unternehmens „advanced biochemical compounds GmbH“ ausgewertet wurden. Die Auswertung im Programm

„ROVER“ erfolgte auf Basis der PERCIST-Kriterien, hier wurde ebenfalls der SUV zur Quantifizierung genutzt.

Die Bestimmung der Leberreferenzregion in „ROVER“ erfolgte mittels einer Kugel mit einem Durchmesser von 3 mL und einem Volumen von 14 mL und wurde standardmäßig in den rechten Leberlappen gesetzt. Der SUV_{peak} errechnet sich aus dem mittleren SUL in einem ein Milliliter messenden Volumen um das „Voxel“ (Datenpunkt in einem dreidimensionalen Gitter) mit der höchsten FDG-Aufnahme der Läsion.

Die Konturierung der Volumengrenzen einer Lymphomläsion wurde mit Hilfe eines iterativen Hintergrund-korrigierten Algorithmus der „ROVER“-Software durchgeführt, welcher in der folgenden Formel dargestellt wird.

$$S = T * (C_{ref} - Bg) + Bg$$

„S“ steht hierbei für den errechneten Schwellwert, „ C_{ref} “ das SUV-Maximum der Läsion, „T“ für die relative Schwelle mit Hintergrundkorrektur und „Bg“ für den SUV des Hintergrundes.

Folgendes STARD-Diagramm liefert einen Überblick, welche Werte miteinander verglichen wurden (Abb. 6).

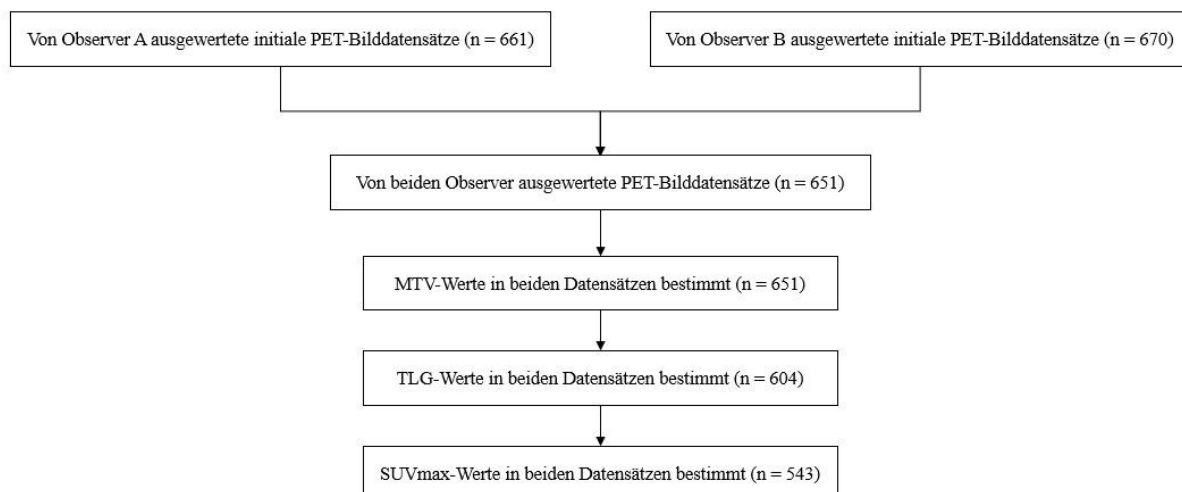


Abbildung 6 – STARD-Diagramm der Datengrundlage der zweiten Interobserveranalyse

3.7 Statistische Analyse

Die statistische Auswertung der zur Verfügung gestellten bzw. der in dieser Arbeit zusätzlich gewonnenen Daten erfolgte in Kombination mit den Programmen „Excel“ von *Microsoft* und „SPSS Statistics 26“ von *IBM*. Die verschiedenen Diagramme und Graphen dieser Dissertation wurden durch Import der zugrundeliegenden „Excel“-Tabellen mit „Python 3.11“ und den

zugehörnden Bibliotheken „pandas“, „matplotlib“ und „plotly“ erstellt. Insgesamt wurden in der statistischen Analyse die adäquaten Diagrammtypen verwendet: Histogramme, QQ-Diagramme, logarithmierte QQ-Diagramme Streudiagramme, Streudiagrammmatrizen, Regressionsanalysen und Bland-Altman-Plots.

Mithilfe der Histogramme und QQ-Diagramme bzw. logarithmierten QQ-Diagramme sollte die visuelle Analyse einer möglichen Normalverteilung bzw. logarithmischen Normalverteilung erfolgen. Die eingesetzten Streudiagramme stützten die visuelle Aufbereitung von Korrelationsanalysen, wie sie beispielsweise in Kap. 4.4 verwendet werden. Die Regressionsanalyse wurde zur Darstellung der extrahierten Messwerte für MTV und TLG in der Interobserveranalyse in syngo.via verwendet, während die Bland-Altman-Plots zur Analyse der Intersoftware- und Interobserveranalyse mit ROVER dienten. Der Bland-Altman-Plot vergleicht zwei Messmethoden, indem er die Differenzen der Einzelmesswerte gegen deren Durchschnitt darstellt. Dabei zeigt der Mittelwert der Differenzen den systematischen Bias, während die Grenzen der Übereinstimmung (Arithmetisches Mittel $\pm 1,96 \cdot SD$) den Bereich angeben, in dem 95 % der Differenzen liegen. So lassen sich systematische Abweichungen und zufällige Messfehler anschaulich identifizieren und quantifizieren.

Die hier verwendeten Angaben zur Korrelation beziehen sich auf die Pearson-Korrelationen (r). Mit diesem Koeffizienten wird der lineare Zusammenhang zwischen zwei Messvariablen quantifiziert. Der Wert von r liegt zwischen -1 (perfekte negative lineare Beziehung) und 1 (perfekte positive lineare Beziehung), wobei 0 auf keinen linearen Zusammenhang hinweist. Zusätzlich wurde noch der

Der Kolmogorov-Smirnov-Test (KS-Test) fand in dieser Arbeit Anwendung, um zu überprüfen, ob die Verteilungen der untersuchten Variablen von einer Normalverteilung abweichen.

Zur Überprüfung signifikanter Unterschiede zwischen den untersuchten PET-Standorten und PET-Scannern wurde eine einfaktorielle Varianzanalyse („ANOVA“) durchgeführt. Dieses Verfahren erlaubt es, Mittelwertunterschiede zwischen mehr als zwei unabhängigen Gruppen zu analysieren und festzustellen, ob eine systematische Varianz zwischen den Gruppen besteht, die über zufällige Schwankungen hinausgeht. Zur Bestimmung der Effektstärke wurde η^2 berechnet, welches das Verhältnis der erklärten Varianz zur Gesamtvarianz angibt und somit die Größe des Effekts beschreibt.

4. Ergebnisse

591 der 661 initialen FDG-PET/CTs (89,4%) zeigten gemäß der Auswertungskriterien und der Einschätzung des Untersuchers ein MTV > 0 mL in der PET, was bedeutet, dass zumindest eine tumorsuspekte FDG-anreichernde Läsion ≥ 1 mL, welche PERCIST-Kriterien konform über dem Schwellwert lag, im Bilddatensatz ermittelt wurde. Im Gegenzug dazu zeigten 70 von 661 Fälle (10,6 %) nach den vordefinierten Kriterien ein MTV von 0 mL.

4.1 Konturierungsqualität im visuellen Scoring

In 29 (4,4%) der Fälle entstand der visuelle Eindruck, dass das MTV um mehr als 50 % unterschätzt wurde, 15 dieser 29 Bilddatensätze wurden mit einem MTV von 0 mL quantifiziert. Diese Häufung lässt sich durch den Umstand erklären, dass eine Läsion nur bei einer konturierten Größe von ≥ 1 mL und Erfüllen des PERCIST-Schwellwertes miteinzukalkulieren war. Bei 56 (8,4%) der Fälle entstand der visuelle Eindruck, dass das MTV um 11-50 % unterschätzt wurde, in 557 (84,3%) der Fälle stimmte der visuelle Eindruck mit dem ermittelten MTV nach Erachten des Untersuchers überein.

Eine Überschätzung des MTV trat deutlich seltener auf als eine Unterschätzung. In nur 16 (2,4%) der Fälle entstand der visuelle Eindruck, dass das MTV um 11-50% überschätzt wurde, während in nur in 3 (0,4%) der Fälle der Eindruck entstand, dass es um mehr als 50% überschätzt wurde (Tab. 7).

Visuelle Bewertung	MTV = 0 mL	MTV > 0 mL	Gesamt
(-2)	15	14	29
(-1)	0	56	56
(0)	55	502	557
(1)	0	16	16
(2)	0	3	3
Anzahl	70	591	661

Tabelle 7 - Visuelle Bewertung nach MTV-Einteilung

4.2 Ermittelte Ursachen der Unterschätzung des Tumervolumens mit scheinbarem MTV von 0 mL

Auf die Subgruppe der Fälle, welche eine visuelle Bewertung von „(-2)“ und, laut den Kriterien des in dieser Arbeit angewandten Selektionsverfahrens, ein MTV = 0 mL aufweisen, kann ein besonderes Augenmerk gelegt werden. Eine Analyse der zugrundeliegenden Daten zeigt, dass bei 11/15 der PET-Bilddaten eine VOI < 1 cm³ der Grund war dieses Areal nicht in die Endauswertung einzuschließen, während bei 4/15 Fälle der Grund das Nichterreichen des PERCIST-Schwellwerts der VOIs war. In nur drei Fällen beträgt die Anzahl der nicht in die Auswertung aufgenommenen VOIs mehr als 1 (Maximum: 3), während das durchschnittliche MTV dieser 15 Bilddatensätzen unter Missachtung der PERCIST- und Volumensselektionskriterien bei 1,66 mL liegen würde.

4.3 SUV_{mean} der Leberreferenzregion und die PERCIST-Schwellwerte in der Gesamtanalyse

Die Auswertung der SUV_{mean} der Leberreferenzregionen, welche mit der Standardabweichung die Grundlage des PERCIST-Schwellwertes bilden, brachte folgendes Ergebnis (n = 651): Das arithmetische Mittel befand sich hier bei einem SUV_{mean} von 2,34, der Median bei 2,33 und die Standardabweichung bei 0,53. Das Minimum der SUV_{mean} in der Leberreferenzregion betrug 0,94, das Maximum 4,41 (Tab. 8). Die visuelle Analyse des Histogramms und das QQ-Diagramm bestätigen den Eindruck einer Normalverteilung (Abb. 7-8), der KS-Test ebenfalls (p = 0,91)

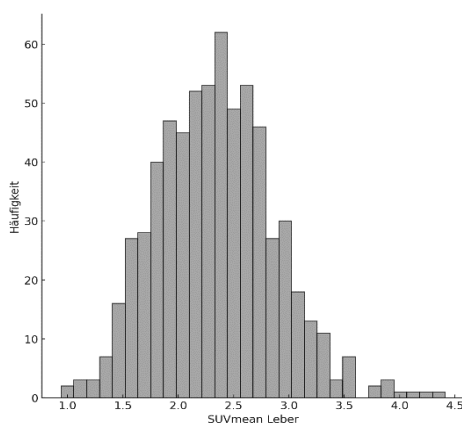


Abbildung 7 - Histogramm der SUV_{mean}-Werte

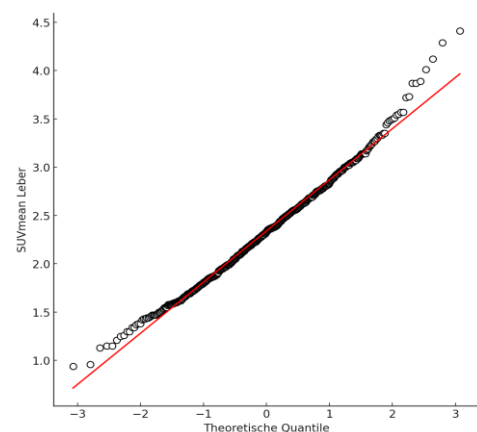


Abbildung 8 - QQ-Diagramm der SUV_{mean}-Werte

Die statistischen Auswertungen zum PERCIST-Schwellwert zeigten ein ähnliches Bild (n=651): Das arithmetische Mittel befand sich bei einem SUV von 4,23, 56,8% der Fälle lagen über der SUV_{≥4}-Schwelle. Der Median lag bei 4,18 und die Standardabweichung bei 1,00. Das

Minimum der PERCIST-Schwellwerte betrug 1,56, während das Maximum bei einem SUV von 9,34 lag (Abb. 9-10) (Tab. 8).

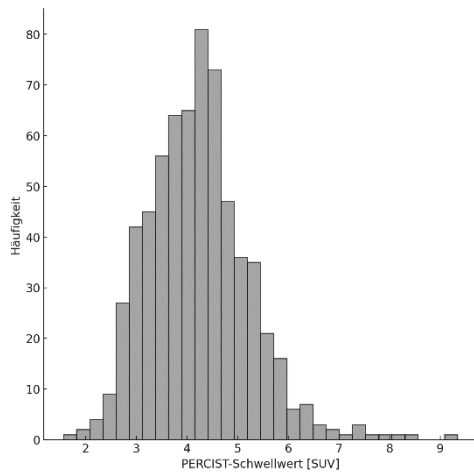


Abbildung 9 - Histogramm der PERCIST-Schwellwerte

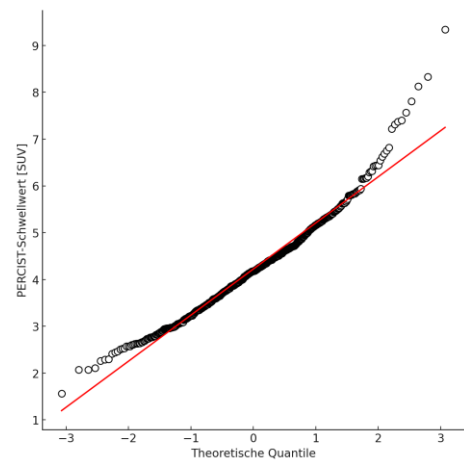


Abbildung 10 - QQ-Diagramm der PERCIST-Schwellwerte

Auch hier bestätigen sowohl das Histogramm als auch das QQ-Diagramm den visuellen Eindruck einer Normalverteilung, der KS-Test lehnt diese aber Trotz Tendenz ab ($p=0,04$).

	SUV _{mean}	PERCIST
Mittelwert	2,34	4,23
Median	2,33	4,18
Standardabweichung	0,53	1,00
Minimum	0,94	1,56
Maximum	4,41	9,34
25. Perzentile	1,96	3,56
50. Perzentile	2,33	4,19
75. Perzentile	2,68	4,76
Schiefe	0,37	0,76
Kurtosis	0,44	1,80

Tabelle 8 - Deskriptive Analyse der SUV_{mean} der Leberreferenzregion und der PERCIST-Schwellwerte

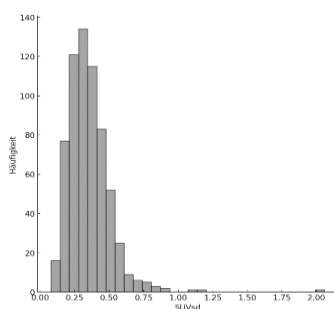


Abbildung 11 - Histogramm der SUV_{SD}-Werte

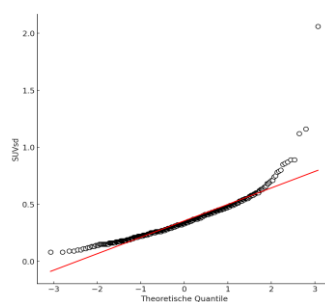


Abbildung 12 - QQ-Diagramm der SUV_{SD}-Werte

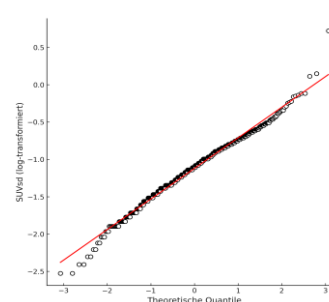


Abbildung 13 - QQ-Diagramm der logarithmierten SUV_{SD}-Werte

Insgesamt ist das Normalverteilungsmuster der SUV_{mean} etwas prägnanter als die der errechneten PERCIST-Schwellwerte. Da die errechneten PERCIST-Schwellwerte nicht nur vom SUV_{mean} der Leberreferenzregion, sondern auch von der SUV_{SD} , also der Standardabweichung der gemessenen SUV der Leberreferenzregion, abhängig sind, konnte hier sinngemäß eine Korrelation nachgewiesen werden. Stark erhöhte SUV_{SD} -Werte der Leber waren insbesondere bei nicht lagekorrigierbaren Tumordinfiltrationen in der Leberreferenzregion zu finden, wie auch das Histogramm ($n = 651$) (Abb. 11) und QQ-Diagramm (Abb. 12) zeigt. Die logarithmierten Werte der SUV_{SD} der Leber imponieren grafisch eher als normalverteilt, als die SUV_{SD} -Rohwerte (Abb. 13), der KS-Test bestätigt diese ($p=0,13$). Der erhöhte SUV_{mean} , gemeinsam mit einer erhöhten Standardabweichung des gemessenen SUV der Leberreferenzregion, führte durch ihre direkte Korrelation bei der Berechnung des PERCIST-Wertes zu höheren Ausreißern. Dies geschah sporadisch bei Tumordinfiltration der Leberreferenzregion. Ein quantitativ ähnlich häufig auftretendes Phänomen konnte bei niedrigen SUV_{mean} -Werten nicht beobachtet werden.

4.4 Einfluss von Körpergewicht, Radioaktivitätsmenge und Distributionszeit

Die Abhängigkeit der SUV_{SD} zu weiteren dokumentierten Messwerten wird in einer Streudiagrammmatrix dargestellt (Abb. 14). Wie in der Streudiagrammmatrix erkennbar, zeigt das Körpergewicht eine leichte bis moderate positive, signifikante ($p<0,001$) Korrelation mit der SUV_{SD} der Leberreferenzregion, der Pearson Korrelationskoeffizient r betrug 0,44. Die Aktivität an F-18-Fluorodeoxyglukose, welche dem Patienten verabreicht wurde, die Radionuklid-Gesamtdosis in MBq, wird vor der Gabe an das Körpergewicht angepasst und zeigt deshalb eine sinngemäß schwache, signifikante ($p=0,002$) Korrelation zum Körpergewicht von $r=0,12$. Die schwache, inverse Korrelation von $r=-0,07$ der applizierten Aktivität mit dem SUV_{SD} -Wert wird hingegen als nicht signifikant ($p=0,09$) eingestuft (Abb. 14).

Ähnliches kann beim Vergleich der Werte für SUV_{mean} und der PERCIST-Schwellwerte mit dem Körpergewicht der Patienten festgestellt werden. Die Korrelation von SUV_{mean} mit dem Körpergewicht betrug $r=0,30$ ($p<0,001$) (Abb. 15), zudem betrug die Korrelation des PERCIST-Schwellwert mit dem Körpergewicht $r=0,37$ ($p<0,001$). Eine Korrelation zwischen SUV_{mean} und applizierte Radionukliddosis konnte jedoch nicht beobachtet werden.

Die Bildgebungsverzögerung, welche in dieser Analyse dokumentierte Werte von 29 min bis 193 min annahm, zeigte keine signifikante Korrelation mit dem SUV_{mean} ($p=0,64$) und dem PERCIST-Schwellwerten ($p=0,31$).

Nach Ausschluss offensichtlich fehlerhafter bzw. fehlender Angaben zur Bildgebungsverzögerung am Datenblatt betrug bei $n=617$ ausgewerteten Bilddaten das arithmetische Mittel der Bildgebungsverzögerung 75,3 min. Ein Ausschluss einer vermuteten fehlerhaften Angabe der übermittelten Daten erfolgte bei einer dokumentierten Bildgebungsverzögerung von unter 15 Minuten.

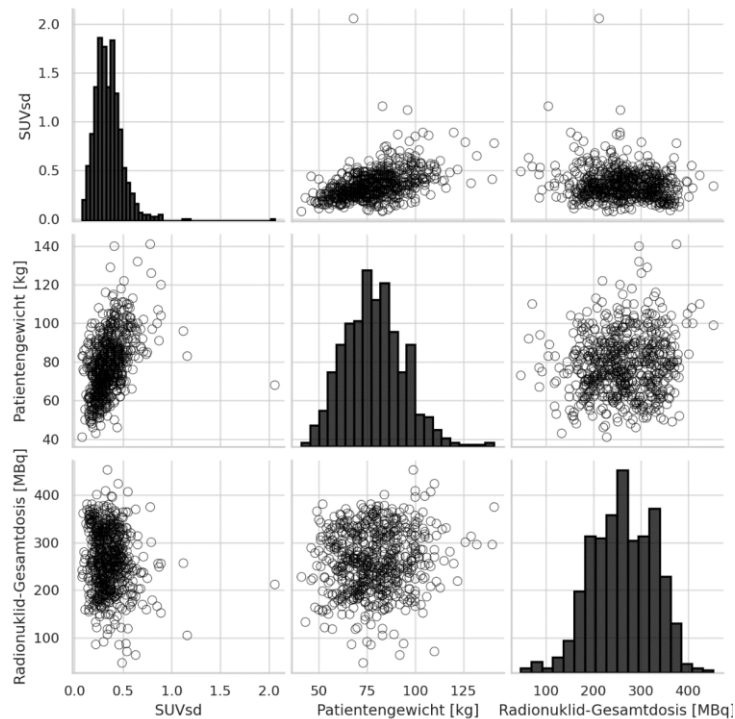


Abbildung 14 - Streudiagramm von SUV_{SD} , Patientengewicht [kg] und Radionuklid-Gesamtdosis [MBq]

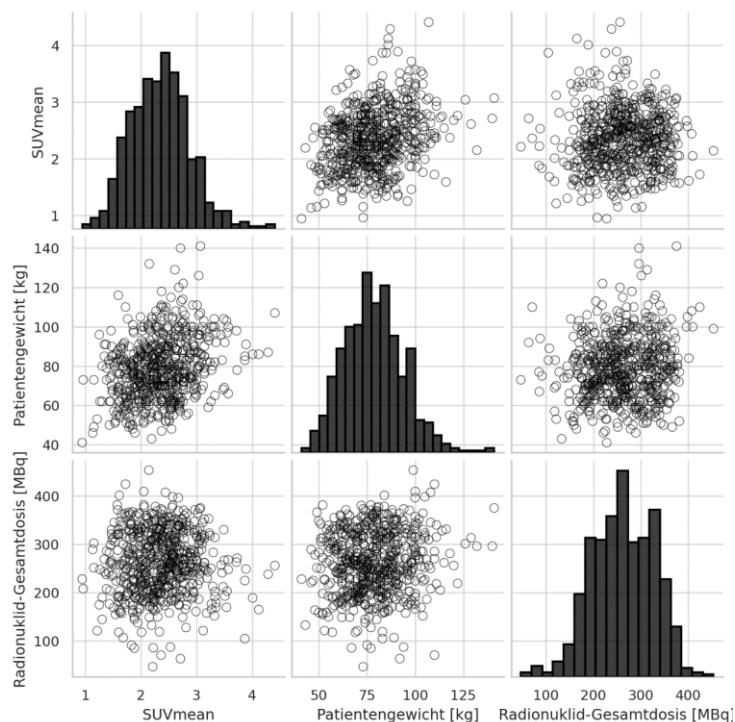


Abbildung 15 - Streudiagramm von SUV_{mean} , Patientengewicht [kg] und Radionuklid-Gesamtdosis [MBq]

4.5 Standortvergleich der SUV_{mean} und PERCIST-Schwellwerte

Die eingeschlossenen PET-Standorte wurden im Rahmen des Datenschutzes pseudonymisiert. Die statistische Analyse des Standortvergleich der SUV_{mean} und PERCIST-Werte zeigten für die beiden Werte unterschiedliche Ergebnisse.

Eingeschlossen in diese Analyse wurden jene Standorte, welche mehr als 20 auswertbare PET-Bilddatensätze der OPTIMAL>60-Studie zur Verfügung stellten, um eine adäquate Stichprobengröße zu gewährleisten. In den nachfolgenden Tabellen finden sich die Ergebnisse der statistischen Analyse, die Fallzahl wird in den eckigen Klammern angegeben (Tab. 9) (Tab. 10). Die meisten Standorte befanden sich mit ihrer Anzahl an Fällen nur knapp über dem Mindestwert von 21 Fällen.

Zentrums-ID	A	B	C	D	E	F	G	H	I
[Anzahl]	[54]	[39]	[24]	[45]	[26]	[21]	[22]	[33]	[22]
Mittelwert	2,51	2,46	2,19	2,43	2,16	2,27	2,40	2,42	2,20
Median	2,56	2,54	2,05	2,42	2,05	2,19	2,40	2,49	2,20
Standardabweichung	0,66	0,38	0,60	0,46	0,59	0,53	0,44	0,57	0,49
Standardfehler	0,09	0,06	0,12	0,07	0,12	0,12	0,09	0,10	0,10
Varianz	0,44	0,15	0,36	0,21	0,34	0,28	0,20	0,32	0,23
Minimum	0,96	1,78	1,55	1,38	1,25	1,38	1,26	1,45	1,44
Maximum	3,87	3,22	4,41	3,14	3,18	3,48	3,14	3,49	3,14

Tabelle 9 - Deskriptive Analyse der SUV_{mean} der Leberreferenzregion nach Standort

Zentrums-ID	A	B	C	D	E	F	G	H	I
[Anzahl]	[54]	[39]	[24]	[45]	[26]	[21]	[22]	[33]	[22]
Mittelwert	4,69	4,45	4,04	4,45	3,93	4,35	4,23	4,39	3,94
Median	4,70	4,46	3,74	4,47	3,92	4,14	4,17	4,45	3,94
Standardabweichung	1,14	0,71	1,17	0,84	1,09	1,42	0,77	1,02	0,86
Standardfehler	0,16	0,11	0,24	0,13	0,22	0,31	0,16	0,18	0,18
Varianz	1,31	0,50	1,36	0,71	1,19	2,04	0,59	1,04	0,74
Minimum	2,07	3,23	2,73	2,69	2,26	2,57	2,30	2,63	2,51
Maximum	7,31	6,17	8,33	6,15	5,83	9,34	5,67	6,31	5,44

Tabelle 10 - Deskriptive Analyse der PERCIST-Werte nach Standort

Mithilfe der einfaktoriellen Varianzanalyse (ANOVA) kam man zu dem Ergebnis, dass beim Standortvergleich der SUV_{mean} zwar ein Trend, aber formal kein signifikanter Unterschied (α -

Niveau = 0,05) der Mittelwerte feststellbar war ($p=0,055$), während die Mittelwerte der PERCIST-Schwellwerte je nach Standort sehr wohl statistisch signifikant unterschiedlich waren ($p=0,023$). Die Effektstärken (η^2) waren jeweils tendenziell niedrig (Tab. 11).

		Quadratsumme	Mittel d. Quadrate	Signifikanz (p)	Effektstärke (η^2)
SUV_{mean} Leber	Zwischen den Gruppen	4,47	0,56	0,055	0,053
	Innerhalb der Gruppen	79,98	0,29		
	Gesamt	84,45			
PERCIST	Zwischen den Gruppen	18,41	2,30	0,023	0,061
	Innerhalb der Gruppen	281,43	1,02		
	Gesamt	299,84			

Tabelle 11 - Ergebnisse der einfaktoriellen Varianzanalyse im Standortvergleich

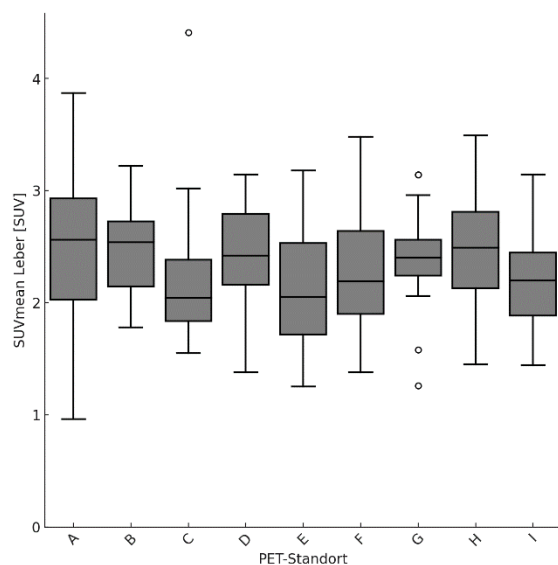


Abbildung 16 - Boxplots des SUVmean der Leber nach PET-Standort

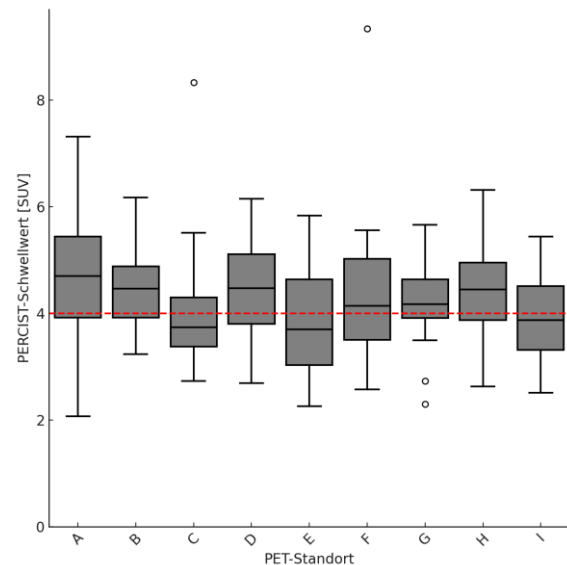


Abbildung 17 - Boxplots der PERCIST-Schwellwerte nach PET-Standort

Die angefertigten Boxplots des Standortvergleichs (Abb. 16-17) zeigen des Weiteren, dass nur in Ausnahmefällen Werte zu finden waren, welche unterhalb bzw. oberhalb der beiden „Whisker“ zu finden waren, sodass die Anzahl an statistischen Ausreißern gering war. Die rote Schraffierung im Boxplot der PERCIST-Schwellwerte (Abb. 17), soll grafisch Aufschluss darüber geben, wo der in Studien häufig gewählte absolute SUV=4-Grenzwert läge.

4.6 PET-Scannermodellvergleich der SUV_{mean} und PERCIST-Schwellwerte

Zusätzlich zur Auswertung der verschiedenen PET-Standorte ist eine Analyse der gemessenen SUV in der Leberreferenzregion und der berechneten PERCIST-Schwellwerte nach verschiedenen PET-Scannermodellen möglich. Die hohe Anzahl an teilnehmenden PET-Standorten bei der OPTIMAL>60-Studie gewährleistet hierbei eine hohe Dichte an verschiedenen PET-Scannern und Herstellern. Insgesamt wurden 7 (1,1%) Bilddatensätze des Herstellers „CPS“, 58 (8,8%) des Herstellers „GE Medical Systems“, 207 (31,3%) des Herstellers „Philips Medical Systems“ und 389 (58,8%) des Herstellers „SIEMENS“ ausgewertet. Die PET-Scanneranalyse schließt Scannermodelle mit mehr als 20 ausgewerteten PET-Bildern ein (Tab. 12).

Hersteller	Scannermodell	Fallzahl
SIEMENS	1080 ¹	25
SIEMENS	1093 ²	57
SIEMENS	Biograph 128	69
SIEMENS	Biograph 40	114
SIEMENS	Biograph 6	29
SIEMENS	Biograph 64	34
GE Medical Systems	Discovery 600	37
Philips Medical Systems	GEMINI TF TOF 16	68
Philips Medical Systems	Guardian Body	43
Philips Medical Systems	TruFlight Select	30

Tabelle 12 - Tabellarische Auflistung der eingeschlossen PET-Scannermodelle mit der Anzahl an ausgewerteten Fällen

Die Ergebnisse der statistischen Analyse im PET-Scannermodellvergleich finden sich auf der nächsten Seite (Tab. 13-14).

¹ Auch bezeichnet als Biograph 16

² Auch bezeichnet als Biograph TruePoint

PET-Scannermodell [Anzahl]	1080 [25]	1093 [57]	Biograph 128 [69]	Biograph 40 [114]	Biograph 6 [29]	Biograph 64 [34]	Discovery 600 [37]	GEMINI TF TOF 16 [68]	Guardian Body (C) [43]	TruFlight Select [30]
Mittelwert	2,23	2,65	2,30	2,40	2,14	2,18	2,43	2,31	2,15	2,27
Median	2,25	2,64	2,28	2,46	2,17	2,09	2,38	2,30	2,19	2,23
Standardabweichung	0,39	0,55	0,50	0,44	0,52	0,54	0,66	0,50	0,43	0,68
Standardfehler	0,08	0,07	0,06	0,04	0,10	0,09	0,11	0,06	0,07	0,12
Varianz	0,15	0,31	0,25	0,20	0,27	0,29	0,43	0,25	0,18	0,46
Minimum	1,49	1,15	1,38	1,34	0,96	1,55	1,15	1,26	1,42	1,25
Maximum	3,14	3,87	3,48	3,44	2,93	4,41	3,89	4,01	3,23	4,29

Tabelle 13 - Deskriptive Analyse der SUV_{mean} nach PET-Scannermodell

PET-Scannermodell [Anzahl]	1080 [25]	1093 [57]	Biograph 128 [69]	Biograph 40 [114]	Biograph 6 [29]	Biograph 64 [34]	Discovery 600 [37]	GEMINI TF TOF 16 [68]	Guardian Body (C) [43]	TruFlight Select [30]
Mittelwert	3,95	4,72	4,33	4,36	4,02	3,98	4,54	4,12	3,60	4,09
Median	4,05	4,63	4,27	4,43	4,17	3,74	4,37	4,08	3,69	3,98
Standardabweichung	0,71	1,07	1,04	0,85	0,84	1,05	1,37	0,88	0,68	1,24
Standardfehler	0,15	0,14	0,13	0,08	0,16	0,18	0,23	0,11	0,10	0,23
Varianz	0,51	1,14	1,08	0,73	0,71	1,10	1,88	0,78	0,47	1,54
Minimum	2,62	2,11	2,57	2,52	2,07	2,73	2,07	2,30	2,29	2,26
Maximum	5,21	7,31	9,34	7,40	5,49	8,33	8,13	7,22	5,38	7,81

Tabelle 14 - Deskriptive Analyse der PERCIST-Werte nach PET-Scannermodell

Wie schon beim Standortvergleich zeigte die einfaktorielle Varianzanalyse (ANOVA) folgenden Ergebnisse: Beim PET-Scannermodellvergleich der SUV_{mean} konnte ein signifikanter Unterschied der Mittelwerte festgestellt werden ($p=1,48 \cdot 10^{-5}$). Die Mittelwerte der PERCIST-Schwellwerte waren je nach PET-Scannermodell ebenso statistisch signifikant unterschiedlich ($p=1,76 \cdot 10^{-6}$), die Effektstärken (η^2) sind mittelhoch. Während im Standortvergleich die einfaktorielle Varianzanalyse keinen signifikanten Unterschied der SUV_{mean} -Werte feststellen konnte, so war der Unterschied der Mittelwerte der SUV_{mean} und auch der PERCIST-Schwellwerte hier klar statistisch signifikant (Tab. 15) (Abb. 18-19).

		Quadratsumme	Mittel d. Quadrate	Signifikanz (p)	Effektstärke (η^2)
SUV_{mean} Leber	Zwischen den Gruppen	10,38	1,15	$1,48 \cdot 10^{-5}$	0,074
	Innerhalb der Gruppen	129,99	0,26		
	Gesamt	140,37			
PERCIST	Zwischen den Gruppen	42,85	4,76	$1,76 \cdot 10^{-6}$	0,083
	Innerhalb der Gruppen	471,13	0,95		
	Gesamt	513,98			

Tabelle 15 - Ergebnisse der einfaktoriellen Varianzanalyse im PET-Scannermodellvergleich

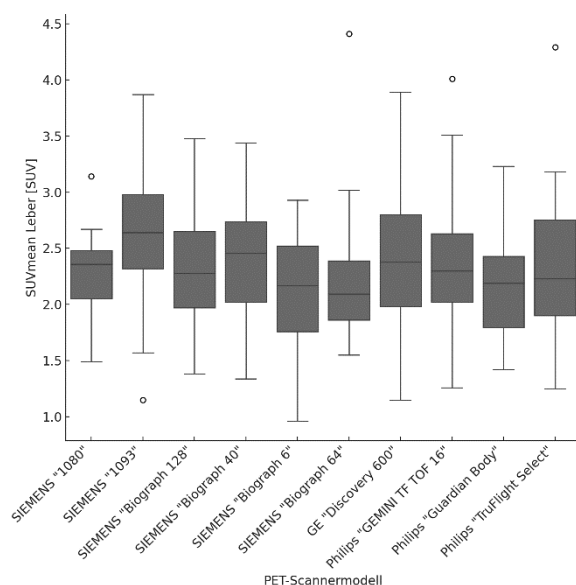


Abbildung 18 - Boxplots der SUV_{mean} -Werte nach PET-Scannermodell

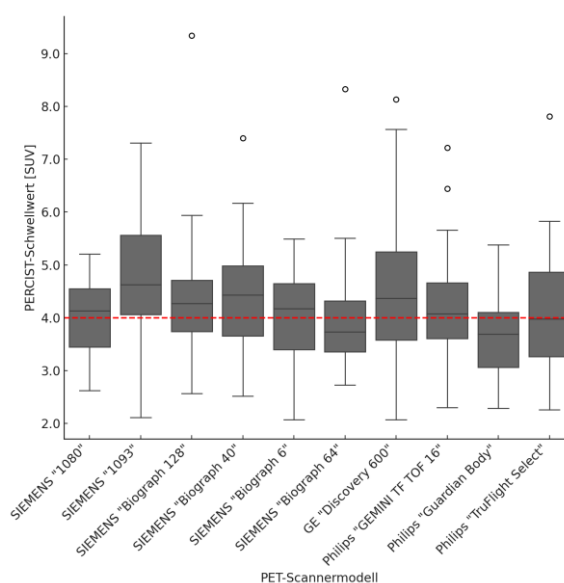


Abbildung 19 - Boxplots der PERCIST-Schwellwerte nach PET-Scannermodell

4.7 Vergleich von SUV_{mean} mediastinal und dem SUV_{mean} der Leber

Der Vergleich der SUV_{mean} -Werte des Mediastinums, welche dem SUV_{mean} der Aorta entsprechen, mit dem SUV_{mean} der Leber brachte folgendes Ergebnis, welches anhand eines Boxplots visualisiert wurde ($n=651$) (Abb. 20): Die Pearson-Korrelation zwischen den beiden SUV_{mean} -Werten beträgt 0,72, was für eine starke Korrelation zueinander spricht ($p<0,001$). Der Mittelwert der SUV_{mean} der Leber liegt mit 2,34 etwas höher als der der Aorta mit 1,72. Die Standardabweichung des SUV_{mean} der Aorta fiel mit 0,45 geringer aus als die der Leber mit 0,53. Des Weiteren nahm das Minimum des SUV_{mean} der Aorta mit 0,59 einen geringeren Wert an als der das Minimum des SUV_{mean} der Leber mit 0,94. Zudem lag auch das Maximum der Aorta mit 3,56 niedriger als das Maximum der Leber mit 4,41 (Tab. 16). Sowohl im Boxplot als auch tabellarisch ist ersichtlich, dass die Diskrepanz zwischen Median und arithmetischen Mittel beim SUV_{mean} der Aorta stärker ausgeprägt ist als beim SUV_{mean} der Leber.

	SUV_{mean} Aorta	SUV_{mean} Leber
Mittelwert	1,72	2,34
Median	1,70	2,34
Standardabweichung	0,45	0,53
Maximum	3,56	4,41
Minimum	0,59	0,96

Tabelle 16 - Deskriptive Analyse der beiden SUV_{mean} -Gruppen

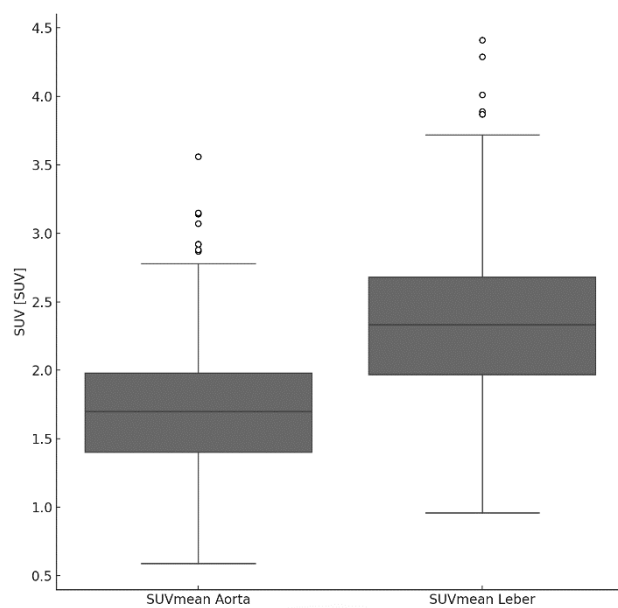


Abbildung 20 - Boxplots der SUV_{mean} von Aorta und Leber

4.8 Interobserver-Analyse der repräsentativen Stichprobe in syngo.via

Die Auswahl der repräsentativen Stichprobe dieser Interobserver-Analyse wurde zuvor bereits im Methodenteil beschrieben und zielte auf eine größtmögliche Abdeckung aller möglichen Tumervolumina in den PET-Bilddatensätzen ab. Die ausgewählte Stichprobe beinhaltet $n=60$ verschiedene FDG-PET/CTs, welche nach der initialen Auswertung durch „Observer 1“ dem „Observer 2“ zur Verfügung gestellt wurden. Die Analyse dieser beiden Auswertungen zielte darauf ab, zu überprüfen, wie reproduzierbar die extrahierten metabolischen Biomarker MTV, TLG und SUV_{max} bei der Auswertung verschiedener fachkundiger Personen sind. Gleichzeitig sollte untersucht werden, wie robust sich die Selektion mittels des PERCIST-Schwellwert darstellt. Im Zuge dieser Interobserver-Analyse konnten folgende Auswertungsergebnissen erzielt werden:

Die berechneten Mittelwerte für das metabolischen Tumervolumina (MTV) und der TLG lagen bei beiden Untersuchern dicht aneinander, genauso wie sich die Einzelergebnisse von MTV und TLG bei 28/60 Fällen (46,7%) exakt glichen und in 48/60 Fällen (80%) sich nur um höchstens 10% voneinander unterschieden. Gemeinsam mit dem Pearson Korrelationskoeffizient r , welcher für MTV $r=0,999$ ($p=2,6 \cdot 10^{-87}$) und für TLG $r=0,999$ ($p=5,2 \cdot 10^{-103}$) betrug und statistisch hochsignifikant ist sprechen diese Ergebnisse für eine außerordentlich hohe Reproduzierbarkeit dieses Auswertungsalgorithmus (Tab. 17). Die Einzelergebnisse wurden in Abb. 21 und 22 jeweils grafisch visualisiert. Während Abb. 21 die Regressionsgerade zwischen den Ergebnissen für MTV visualisiert, zeigt in Abb. 22 die Regressionsgerade der bestimmten TLG-Werte.

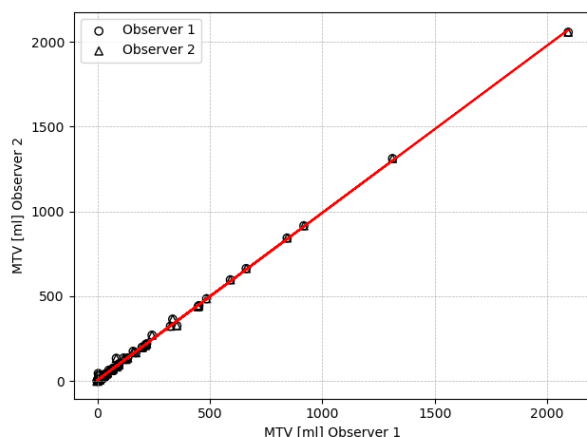


Abbildung 21 - Regressionsanalyse MTV im Interobserver-Vergleich

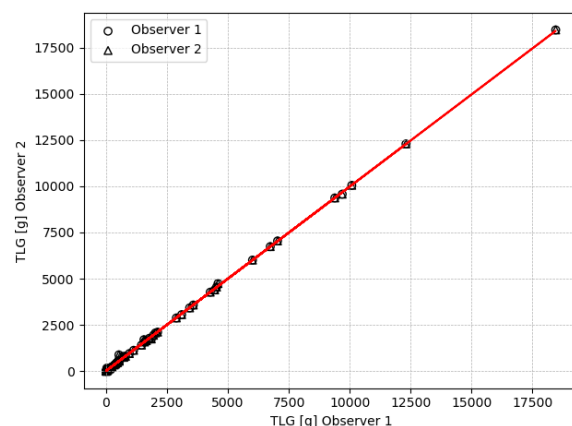


Abbildung 22 - Regressionsanalyse TLG im Interobserver-Vergleich

Der Mittelwertvergleich zwischen den ermittelten SUV_{max} in den zugrundeliegenden PET-Bilder zeichnete ein ähnliches Bild ($r=0,997$, $p=5,9*10^{-65}$). Der bestimmte PERCIST-Schwellwert ist in 45/60 Fällen ident und zeigte in 52/60 Fällen nur maximal einen Unterschied von 10% zueinander. Die Mittelwerte unterschieden sich mit 20,58 für Observer 1 und 20,62 für Observer 2 nur minimal voneinander, das Minimum und Maximum war ident.

	MTV Observer 1 [mL]	MTV Observer 2 [mL]	TLG Observer 1 [g]	TLG Observer 2 [g]	SUV_{max} Observer 1	SUV_{max} Observer 2
Anzahl	60	60	60	60	60	60
Mittelwert	199,9	202,8	2278,5	2292,8	20,6	20,6
SD	356,2	351,7	3571,4	3557,9	12,9	12,7
Minimum	0,0	0,0	0,0	0,0	0,0	0,0
Maximum	2092,8	2056,5	18457,5	18457,5	50,5	50,5

Tabelle 17 – Deskriptive Analyse von MTV, TLG und SUV_{max} in der ersten Interobserveranalyse (n=60)

Die SUV_{mean} , SUV_{SD} , und SUV_{max} -Werte der Leberreferenzregion zeigten bei beiden Untersuchern eine hohe Pearson-Korrelation zueinander (n=60), welche jeweils $r=0,95$ für den SUV_{mean} , $r=0,89$ für den SUV_{SD} und $r=0,89$ für SUV_{max} betrug. Der zugehörigen SUV der Aorta verhielten sich ähnlich (n=60). Der SUV_{mean} der Aortenreferenzregion zeigte einen Korrelationskoeffizienten $r=0,87$, SUV_{SD} von $r=0,89$ und der SUV_{max} von $r=0,90$ (alle $p<0,001$).

Die ermittelten PERCIST-Schwellwerte zeigen zudem ebenfalls eine statistisch signifikante Pearson-Korrelation zueinander ($p=1,5*10^{-31}$), welche $r=0,96$ beträgt.

Die visuellen Einschätzungen, ob das metabolische Tumervolumen mit dem Auswertungsalgorithmus nun unter-, über- oder richtig geschätzt wurde divergieren hingegen stärker. Während auch hier eine statistisch signifikante Korrelation ($p=0,003$) feststellbar ist, fiel diese im Verhältnis zu den weiteren Ergebnissen in der Interobserveranalyse mit $r=0,37$ geringer aus. Im direkten Vergleich tendierte Observer 2 häufiger zu der Entscheidung einer visuellen Unterschätzung des metabolischen Tumervolumens als Observer 1. Die Einschätzung, dass das konturierte metabolische Tumervolumen visuell überschätzt erscheint wurde jedoch von beiden Untersuchern kaum bzw. gar nicht getroffen (Tab. 18).

	Observer 1	Observer 2
(-2)	3	12
(-1)	1	14
(0)	56	32
(1)	0	1
(2)	0	1

Tabelle 18 - Interobserververgleich visuelle Einschätzung der MTVs

Aufgrund des standardisierten Auswertungsalgorithmus glich die Anzahl an identifizierten VOIs („volumes of interest“), also die Anzahl der FDG-Anreicherungen welche tumorsuspekt erschienen, in großen Teilen der Stichprobe einander bzw. war im Vergleich nur marginal voneinander entfernt. So war die Anzahl der VOIs von „Observer 1“ und „Observer 2“ in 31/60 (51,6%) Fällen vollkommen ident und in 54/60 (90,0%) Fällen um maximal 2 VOIs voneinander entfernt.

Es zeigten sich in 13/60 Messungen des PERCIST-Schwellwerts Differenzen, wie die nachfolgende Abbildung darstellt (Abb. 23). Diese traten mit 4 von 5 Fällen (80%) vermehrt beim manuellen Setzen der Leberreferenzregion auf, beim Setzen der Referenzregion durch das Programm syngo.via in 9 von 55 Fällen (16,4%) war diese Zahl geringer.

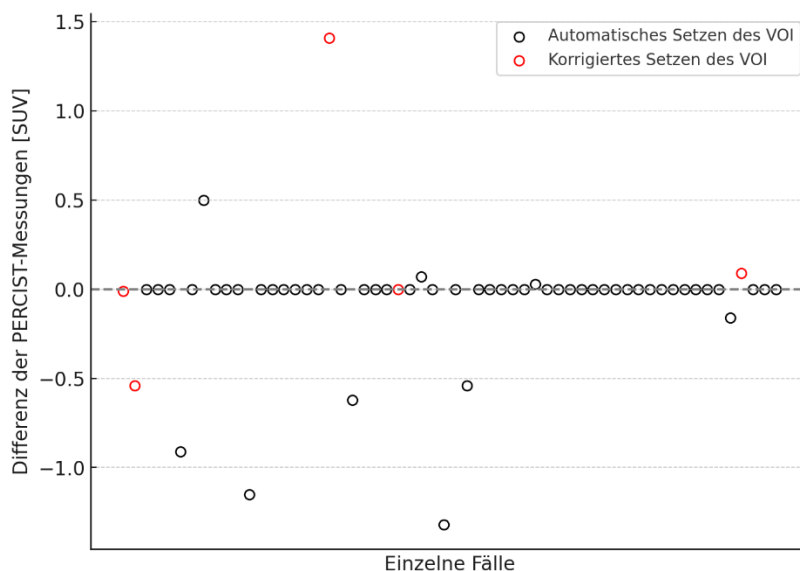


Abbildung 23 - Grafische Darstellung der Differenzen in den Messungen des PERCIST-Schwellwerts in der Interobserveranalyse. Die in rot hinterlegten Datenpunkte stellen die VOIs dar, welche von den Observern manuell korrigiert werden mussten.

4.9 Interobserver-Analyse des MTV und der TLG bei Nutzung verschiedener Softwares und Konturierungsmethoden

Folgender Abschnitt widmet sich nun der zweiten Interobserveranalyse. Ziel dieser war zu überprüfen, ob das MTV und die TLG bei Nutzung verschiedener Konturierungsmethoden und verschiedener Software auch reproduzierbar sind. Bei beiden Analysen erfolgte eine Selektion potenzieller Herde mit dem PERCIST-Schwellwert. Zur Vermeidung von Missverständnissen wird hier statt der Bezeichnungen „Observer 1 und 2“ wie in der ersten Interobserveranalyse die Bezeichnungen „Observer A und B“ gewählt. Zur grafischen Darstellung wurden hier zwei Bland-Altman-Plots, einmal für das MTV und einmal für die TLG, gewählt (Abb. 24-25). Die Ergebnisse der statistischen Analyse finden sich zudem in den nachfolgenden Tabellen (Tab. 19).

Die analysierte Fallzahl betrug bei der Analyse des MTV 651 PET-Bilder, und bei der Analyse der TLG 604 PET-Bilder. Sowohl das MTV ($r=0,95$), als auch die TLG ($r=0,96$) zeigten eine hochsignifikante Pearson-Korrelation zueinander (beide $p<0,001$) (Tab. 19).

	MTV Observer A [mL]	MTV Observer B [mL]	TLG Observer A [g]	TLG Observer B [g]	SUV _{max} Observer A	SUV _{max} Observer B
Anzahl	651	651	604	604	543	543
Mittelwert	179,3	192,6	2113,7	2205,3	24,5	23,6
SD	304,8	320,1	3326,7	3546,4	12,9	12,1
Minimum	0,0	0,0	0,0	0,0	0,0	0,0
Maximum	2503,2	2348,1	23582,9	23386,0	87,6	80,2

Tabelle 19 – Deskriptive Analyse für MTV, TLG und SUV_{max} in der zweiten Interobserveranalyse

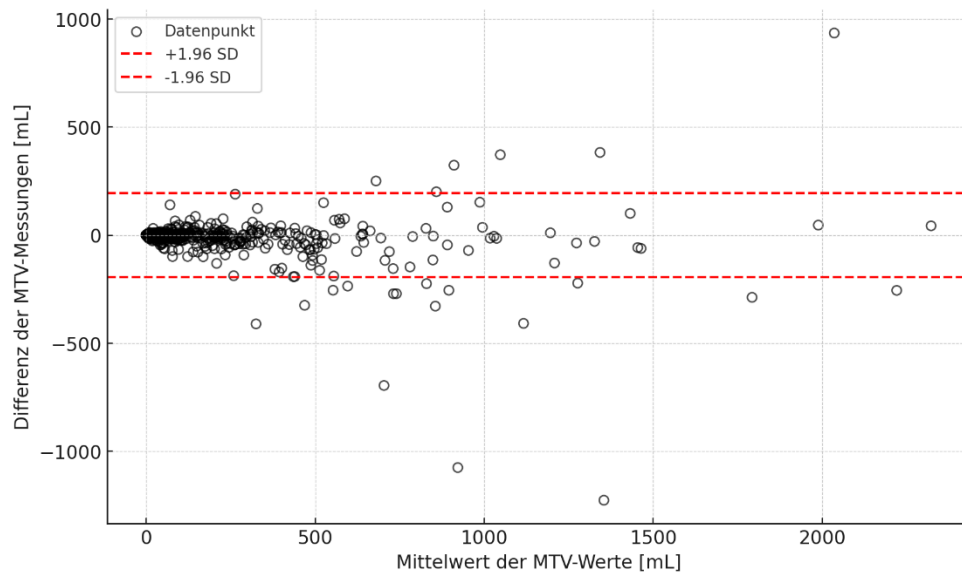


Abbildung 24 - Bland-Altman-Diagramm der MTV in der zweiten Interobserveranalyse (n=651)

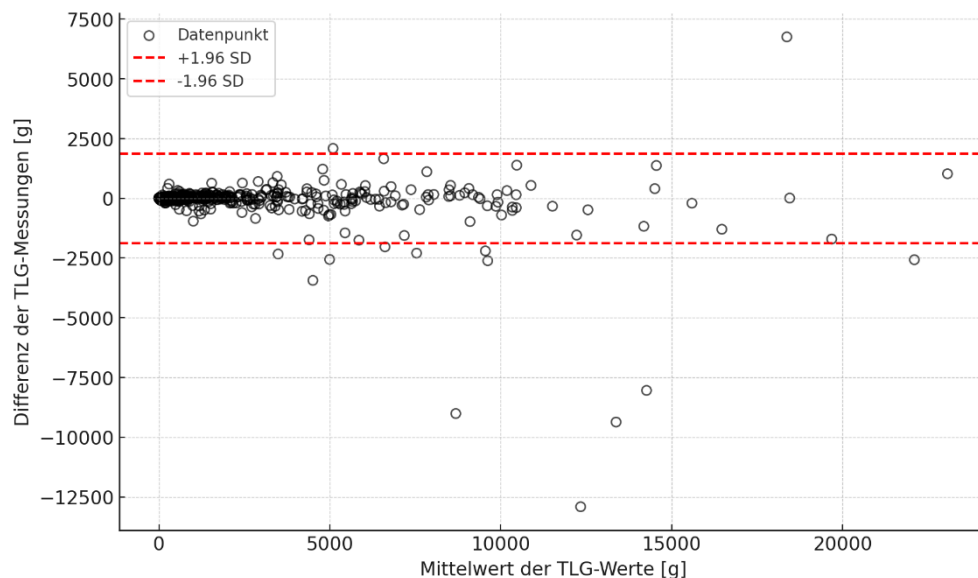


Abbildung 25 - Bland-Altman-Diagramm der TLG in der zweiten Interobserveranalyse (n=604)

Insgesamt 96,4% (628 von 651) der Datenpunkte des MTV und 97,7% (590 von 604) der Datenpunkte befanden sich in den Grenzen des Bland-Altman-Plots, was für eine hohe methodische Konstanz spricht, und dafür, dass die kalkulierten Tumervolumina auch bei den beiden unterschiedlichen Konturierungsmethoden vergleichbare Ergebnisse lieferten. Auch die Ergebnisse für den SUV_{max} zeigten bei (n=543) mit einer Pearson-Korrelation $r=0,97$ eine gute Reproduzierbarkeit.

4.10 MTV, TLG und SUV_{max} in den ausgewerteten Fällen

Das totale MTV wurde routinemäßig bei allen auswertbaren PET-Bilddatensätzen bestimmt (Abb. 26). Die Ergebnisse der bisherigen Analyse (n=661): Das arithmetische Mittel betrug 178,3 mL, der Median lag bei 57,3 mL, die Standardabweichung lag bei 303,7 mL. Das Minimum lag durch die PET ohne Tumornachweis in der Bildgebung bei 0 mL, während das Maximum der Analyse bei 2503,2 mL lag. 70 von 661 Fällen zeigten ein MTV von 0 mL.

Die gesonderte Betrachtung der PET-positiven Bilddatensätze (n=591) geht einher mit der Veränderung der statistischen Werte. Das arithmetische Mittel betrug 199,4 mL, der Median lag bei 72,5 mL, die Standardabweichung bei 314,6 mL. Das Minimum lag bei 1 mL, während das Maximum weiterhin bei 2503,2 mL lag (Tab. 20).

TLG wird als weiterer potenziell prognostischer Parameter für PFS und OS in der Forschung gehandelt [64] und wurde daher bei der Auswertung der Bilddatensätze mitberücksichtigt (Abb. 27). Die Ergebnisse stellen sich bei n=651 folgendermaßen dar:

Das arithmetische Mittel der TLG-Last bei Miteinbeziehung der aller Datensätze betrug 2104,9 g, der Median beträgt 667,8 g, die Standardabweichung lag bei 3302,1 g. Das Minimum lag bei 0g, das Maximum bei 23582,9 g (Abb. 27).

Eine gesonderte Betrachtung ausschließlich PET-positiver Bilddatensätze (n=591) geht mit der Veränderung statistischen Werten einher. Das arithmetische Mittel lag bei 2354,1 g, der Median bei 833,2 g, die Standardabweichung bei 2407,2 g. Das Minimum lag bei 4,7 g, das Maximum weiterhin bei 23582,9 g (Tab. 20).

	MTV (alle Fälle)	MTV (nur MTV>0mL)	TLG (alle Fälle)	TLG (nur TLG>0g)
Anzahl	661	591	651	591
Mittelwert	178,3	199,4	2104,9	2354,1
Median	57,3	72,5	667,8	833,2
SD	303,7	314,6	3302,1	3407,2
Minimum	0,0	1,0	0,0	4,7
Maximum	2503,2	2503,2	23582,9	23582,9

Tabelle 20 - Tabellarische Auflistung statistischer Werte für MTV und TLG

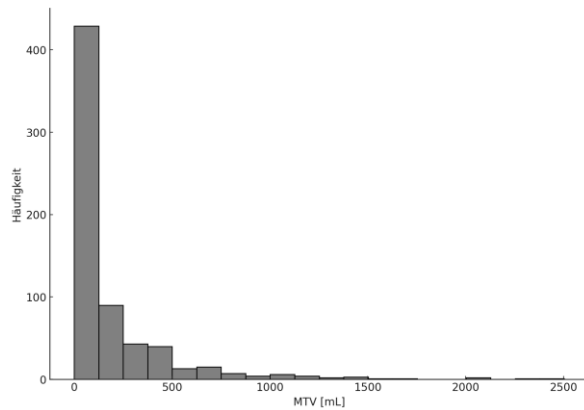


Abbildung 26 - Histogramm der MTV-Werte im gesamten Datensatz

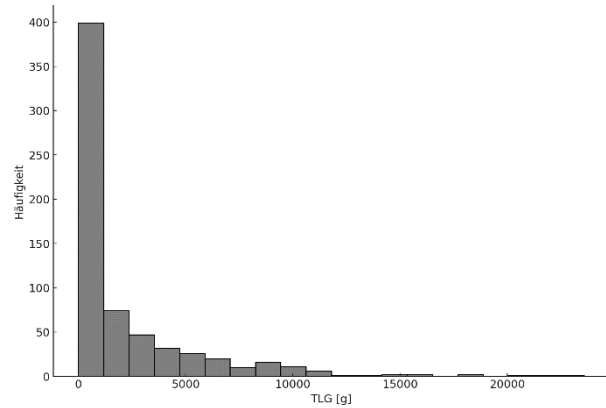


Abbildung 27 - Histogramm der TLG-Werte im gesamten Datensatz

Insgesamt zeigte sich sowohl bei den MTV als auch bei der TLG ein deutliches Verteilungsmuster zugunsten niedriger Werte. Sowohl das MTV als auch die TLG zeigten zudem eine relevante Erhöhung bei Nichteinbeziehung von Bilddatensätzen mit metabolischen Tumorumfängen von 0mL.

Der ermittelte SUV_{max} der Lymphomherde im Bildbefund nahm in 346/651 (53,1%) der Fälle Werte zwischen 10 und 30 an. Wenn man ausschließlich PET-positive Befunde in diesen Vergleich miteinbezieht, so kommt man auf das Ergebnis, dass der SUV_{max} in nur 130/651 (20,0%) der Fälle Werte kleiner als 10 annahm, und so eine verhältnismäßig niedriger SUV_{max} seltener vorkam als ein verhältnismäßig hoher SUV über 30. Der ermittelte Mittelwert der SUV_{max} in allen Fällen ($n=651$) betrug 22,3, der Median 21,5 und die Standardabweichung lag bei 14,1. Der niedrigste beobachtete Wert lag bei 0,0, der höchste bei 87,6.

4.11 MTV, TLG und SUV_{max} je nach Risikostratifizierung

Die Einteilung der Patienten gemäß ihrem Risiko, wie dies im Rahmen der OPTIMAL>60-Studie geschah, ermöglicht die Untersuchung der metabolischen Biomarker im Vergleich zueinander. Hier wurde die Einteilung der OPTIMAL>60-Studie übernommen und die Patienten in zwei Kollektive, dem „favourable“-Kollektiv (n=183; IPI=1) und „less favourable“-Kollektiv (n=472; IPI>1), zugeordnet und auf Unterschiede untersucht.

	MTV	MTV	TLG	TLG	SUV _{max}	SUV _{max}
IPI	1	>1	1	>1	1	>1
Anzahl	183	472	183	472	183	472
Mittelwert	17,2	242,1	222,7	2853,1	14,6	25,2
Median	7,6	116,5	54,2	1407,8	12,0	24,4
Minimum	0,0	0,0	0,0	0,0	0,0	0,0
Maximum	200,3	2503,2	2888,4	23582,9	72,8	87,6

Tabelle 21 - Deskriptive Analyse von MTV, TLG und SUV_{max} je nach vorangegangener Risikostratifizierung

Die Werte der MTV und TLG wiesen drastische Unterschiede auf (Abb. 28) (Tab. 21). Der Mittelwert der gemessenen MTV und TLG-Werte zeigt eine deutliche Diskrepanz, der Median ebenso ($p<0,001$). Der Unterschied der gemessenen SUV_{max} von tumorsuspekten Läsionen unterschied sich im Verhältnis zu den Unterschieden der MTV und TLG-Werte weniger deutlich (Tab. 21).

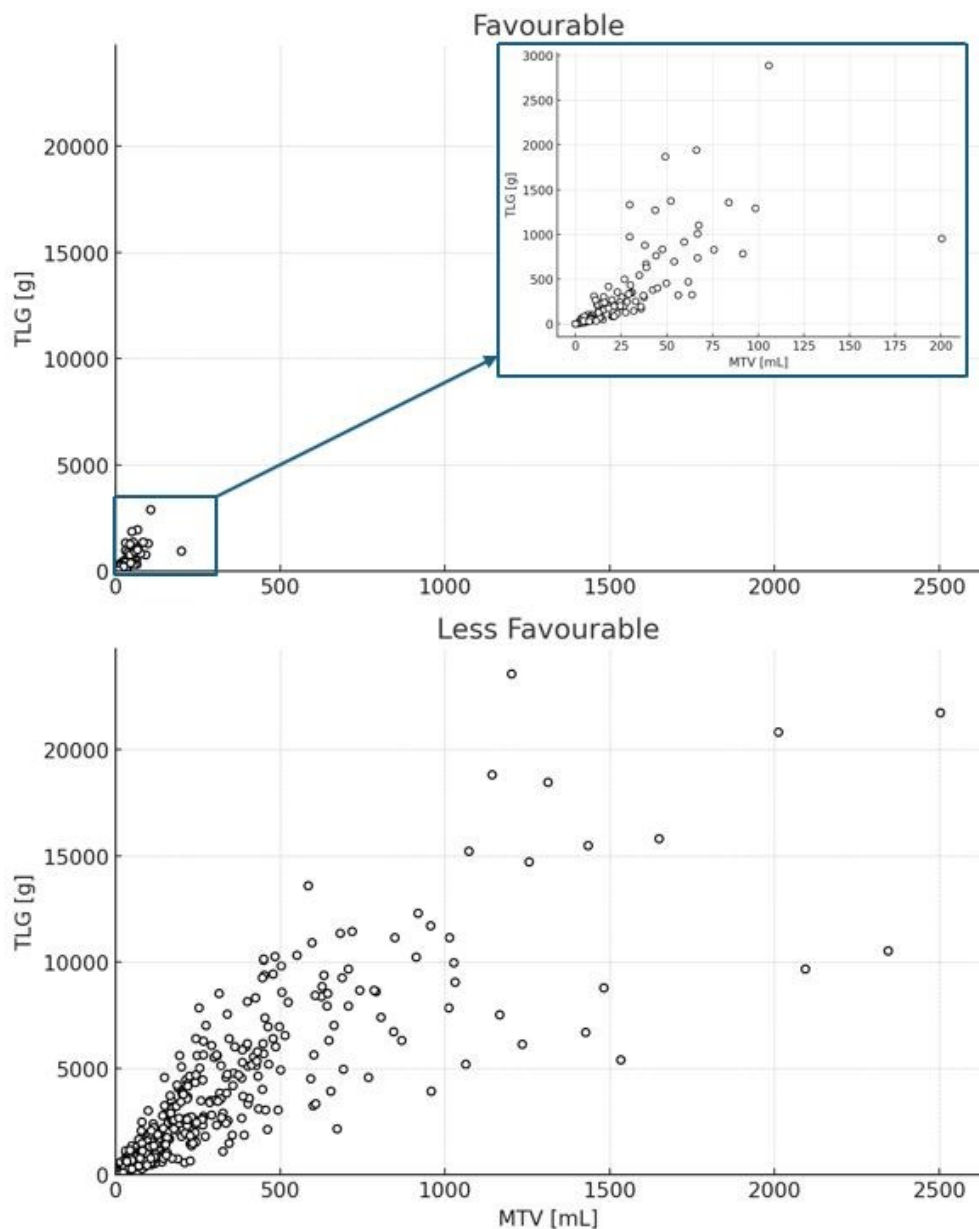


Abbildung 28 - Streudiagramm der MTV-Werte (x-Achse) und TLG-Werte (y-Achse) nach vorangegangener Risikostratifizierung in „favourable“ (n=183) und „less-favourable“ (n=472)

Während ein Befund von MTV = 0 mL im Kollektiv der „less-favourable“ (IPI>1) in nur 20/472 (4,2%) der Fälle vorlag, war bei 48/183 (26,2%) im „favourable“-Kollektiv (IPI=1) gemäß des Auswertungsalgorithmus kein metabolisches Tumervolumen kalkulierbar. In der Subgruppe der „favourables“ befanden sich alle MTV-Werte unter 201 mL, in der Subgruppe der „less-favourables“ 62,1%.

4.12 Statistische Ausreißer in der Bestimmung des MTV und der TLG

Sowohl das MTV als auch die TLG zeigten eine Verteilung zugunsten niedriger Messwerte, sowohl bei Betrachtung des Gesamtkollektives als auch nach Ausschluss aller Fälle mit MTV=0mL. Ausreißer nach oben kamen bei beiden Analysen vor, während das metabolische Tumervolumen bei 17/661 (2,3%) Patienten mehr als drei Standardabweichungen über dem arithmetischen Mittel (1058,1 mL) liegt. Die Zahl sehr hoher TLG-Werte lag mit 11/651 (1,5%) der Fälle etwas unter der Zahl der MTV-Ausreißer, ein Wert von TLG =11.846,3 g entsprach einem Wert von 3 Standardabweichungen über dem Mittelwert. Unter den 17 Fällen mit einem MTV von über 1058,1mL fanden sich ebenfalls 9 von den 11 Fällen wieder, welche ein TLG von über 11 846,3 g aufwiesen. Eine absolute Deckungsgleichheit dieser statistischen Ausreißer kann daher nicht hergestellt werden. Die Anzahl der VOIs schwankte jedoch bei den statistischen Ausreißern der MTV und TLG-Werte stark [7;107].

4.13 Einfluss der MTV-Last auf den SUV_{mean} der Leber und PERCIST-Schwellwert

Ein weiterer Aspekt, welcher analysiert werden kann, ist der Einfluss des metabolischen Tumervolumens und der TLG auf den FDG-Uptake in der Leber. Der Zusammenhang wurde in Abbildung 29 und 30 mithilfe eines Streudiagramms visualisiert, die Linien in lila visualisieren die Regressionsgerade des SUV_{mean} (Abb. 29) und des PERCIST-Schwellwerts (Abb. 30) inklusive deren Vertrauensbereich.

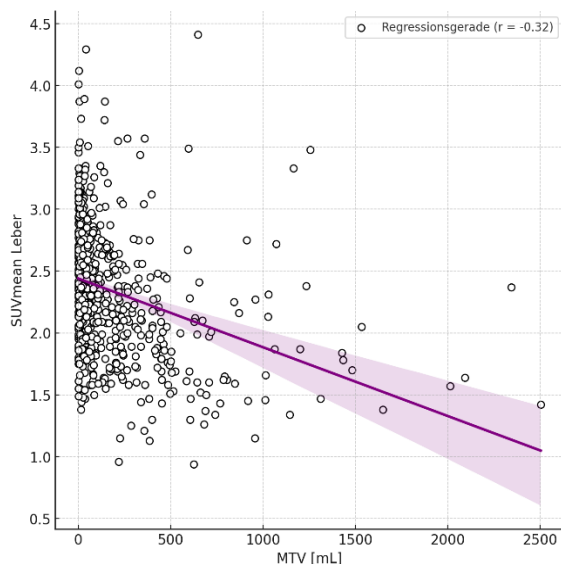


Abbildung 29 - Streudiagramm SUV_{mean} der Leber und dem MTV mit Regressionsgerade (lila) und Vertrauensbereich

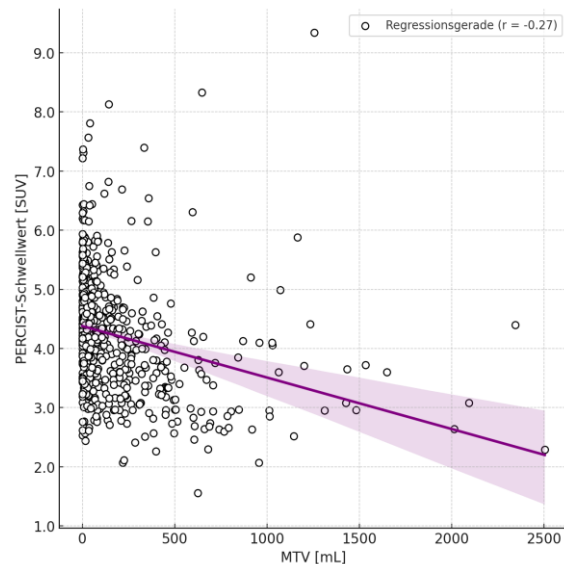


Abbildung 30 - Streudiagramm des PERCIST-Schwellwerts und dem MTV mit Regressionsgerade (lila) und Vertrauensbereich

Das metabolische Tumervolumen zeigt eine signifikante negative Pearson-Korrelation zum SUVmean ($r=-0,32$, $p<0,001$) und zum PERCIST-Schwellwert ($r=-0,27$, $p<0,001$). Auffallend an dieser Darstellung ist zudem, dass besonders hohe metabolische Tumervolumina ($MTV \geq 1000$ mL) zum überwiegenden Anteil unter dem arithmetischen Mittel der jeweiligen Werte liegen (16/21, 76,1 %). Ein Erklärungsansatz hierfür kann der „Tumor-sink“-Effekt sein, auf welchen in Kap. 5.4 im Diskussionsteil näher eingegangen wird (Abb. 31).

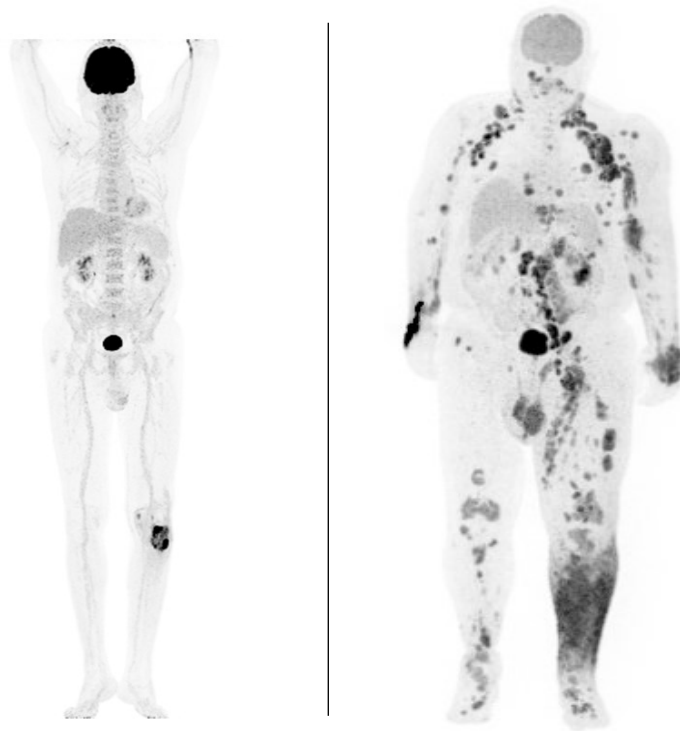


Abbildung 31 - MIP. Anhand dieser beiden MIPs kann der "tumor-sink"-Effekt dargestellt werden. Im Bild links kann man eine verhältnismäßig stark FDG-anreichernde Leber erkennen (SUV_{mean} : 2,59; 71. Perzentile) bei moderaten MTV (23,0 mL; 34. Perzentile). Im rechten Bild sieht man einen Patienten mit einem ausgesprochen hohen MTV (2092,3 mL; 99. Perzentile) und einer geringen Anreicherung von FDG im Lebergewebe (SUV_{mean} : 1,64; 9. Perzentile)

5. Diskussion

5.1 Evaluation der Ergebnisse der Interobserver-Analysen

Die hier eingesetzte kombinierte Auswertungsmethode mit Selektion über den PERCIST-Schwellwert und Konturierung über 41% des tumoralen SUV_{max} mit manueller Selektion der durch das Programm markierten VOIs unterliegt, wie andere Verfahren zur Bestimmung des metabolischen Tumorumfanges, gewissen Limitationen. Jedoch erweist sie sich auf Basis einer hohen Übereinstimmung in der hier gezeigten ersten Interobserver-Analyse (Kap. 4.5) als höchst zuverlässig in der Bestimmung von MTV, TLG (beide $r=0,999$) sowie SUV_{max} ($r=0,997$). Die zur Bestimmung des PERCIST-Schwellwerts notwendigen SUV_{mean} und SUV_{SD} in der Leber zeigten bei dieser ersten Interobserver-Analyse ebenfalls eine hohe Korrelation zueinander ($r_{SUV_{mean}}=0,95$ bzw. $r_{SUV_{SD}}=0,89$). Begünstigend wirkte hier die Funktion „multifokale Segmentierung“ des Programms syngo.via, das die Leberreferenzregion automatisiert und standardisiert in das anatomisch korrekte Areal setzte. In 5/60 Fällen war jedoch eine Korrektur der Leberreferenzregion notwendig, da das System diese nicht in die anatomisch korrekte Region setzen konnte, was die standardisierte Bestimmung des PERCIST-Schwellwerts tendenziell negativ beeinflusst (Kap. 4.8).

Bei der visuellen Bewertung der Konturierung des MTV stuften beide Observer bei einem relevanten Anteil an Fällen das Tumorumfang als leicht (10-50%) oder stark (über 50%) unterschätzt ein. Eine wesentliche Ursache für die Unterschätzung kann sein, dass aufgrund der geforderten Mindestgröße für VOIs einige Tumorherde nicht zum MTV gezählt wurden. Die Anzahl der als unterschätzt eingestuften metabolischen Tumorumfänge variierte zwar zwischen Observer 1 (4/60) und Observer 2 (26/60), wobei sich allerdings die Anzahl der zur Berechnung des MTV verwendeten VOIs in über 90 % der Fälle höchstens um 2 unterschied. Eine mögliche Ursache für diese Divergenz der subjektiven Einschätzungen könnte der Erfahrungsunterschied in der Bewertung von PET/CTs der beiden Observer sein, sowie das generell hohe Maß an Subjektivität, welche diese Scoring-Methode mit sich bringt. Eine Paralleldokumentation der Volumina verdächtiger VOIs, welche nicht in die Auswertung mitfließen, könnte zu einer Verbesserung der Übereinstimmungswerte beitragen.

In einem Versuch, die Interobserver-Reliabilität von verschiedenen Methoden und Programmen (PETTRA 2.5, HERMES 2.5) hinsichtlich des bestimmten MTV bei DLBCL-Patienten zu messen, beobachteten *Ilyas et. al* (2018) sowohl für SUV_{max} (41%) als auch den PERCIST-Schwellwert ein hohes Maß an Korrelation, jedoch wurden diese nur einzeln und nicht kombiniert überprüft [72]. Die hier gezeigten Ergebnisse der Interobserver- und

Interobserveranalyse im Vergleich zur Software „ROVER“ konnten ebenfalls ein hohes Maß an Korrelation der kalkulierten bildbasierten Biomarker bestätigen (MTV: $r=0,95$; TLG: $r=0,96$). Es zeigte sich jedoch ein erwartbar höherer Unterschied als in der ersten Interobserveranalyse, welche das gleiche Auswerteprogramm nutzte. Da mit 96,4% der MTV-Messergebnisse und 97,7% der TLG-Messergebnisse ein sehr hoher Anteil innerhalb der Grenzen der Standardabweichungen im Bland-Altman-Plot lagen, kann eine systematische proportionale Abweichung einer der beiden Observer und/oder Programme ausgeschlossen werden.

5.2 Evaluation der Ergebnisse der PERCIST-Schwellwerte nach PET-Scannermodell, Studienzentrum und Patientengewicht

Die Ergebnisse für SUV_{mean} und den PERCIST-Schwellwert zeigen in der Korrelationsanalyse (Kap. 4.4) eine Abhängigkeit vom PET-Scannermodell ($p=1,48 \cdot 10^{-5}$ bzw. $p=1,76 \cdot 10^{-6}$) mit relevanten Effektstärken η^2 von über 6%. Für den Einfluss des Standortes kann kein so ausgeprägter Effekt beobachtet werden (Effektstärke η^2 unter 6%), auch wenn grundsätzlich Signifikanz gegeben war für den Einfluss auf die PERCIST-Schwellwerte ($p=0,023$). Für SUV_{mean} war lediglich eine Tendenz zu beobachten ($p=0,055$).

Durch die lange Laufzeit der OPTIMAL>60-Studie von 2011 bis 2020 und den damit verbundenen prolongierten Zeitraum für das Anfertigen von FDG-PET/CT-Aufnahmen zeigte sich eine erhöhte Varietät in verwendeten PET-Scannern. Moderne PET/CT-Scanner stellen bei kleineren Pixelgrößen höhere SUV-Maxima dar als vor allem ältere PET/CT-Scanner mit größeren Pixelgrößen. Die Verwendung von SUV-Peak-Werten (Mittelwert in einer Kugel mit 1 mL Volumen um die stärkste Anreicherung) könnte in diesem Aspekt zu einem robusteren Ergebnis bei Messungen von tumoralen FDG-Anreicherungen kommen [3].

Im Vergleich zwischen den Standorten könnte die Ablehnung eines signifikanten Einflusses auf SUV_{mean} daher rühren, dass zwei der untersuchten Standorte („A“ und „I“) während der Studienlaufzeit den PET-Scanner erneuert haben. In weiterer Folge fällt auch die Diversität der verwendeten PET-Scannern untereinander etwas geringer aus als im PET-Scannervergleich. Die Modelle „Biograph 128“, „Biograph 40“ und „1093“ des Herstellers SIEMENS Healthineers wurden jeweils zweimal von verschiedenen PET-Standorten verwendet.

Trotz dieser Limitationen zeigen diese Ergebnisse, dass der mittlere FDG-Uptake im rechten Leberlappen statistisch signifikant schwankt. In weiterer Folge kann auf Basis dieser Information angenommen werden, dass die Evaluation des metabolischen Tumorzusatzes

mittels absoluter SUV-Grenzen wie z.B. der $SUV \geq 2,5$ -Methode oder $SUV \geq 4,0$ -Methode weniger stark auf die individuelle FDG-Stoffwechselaktivität des Patienten eingehen. Hierbei könnte das ermittelte arithmetische Mittel der SUV_{mean} -Werte des PET-Scannermodells „1093“ von SIEMENS angeführt werden, welches mit einem Wert von 2,65 bereits durchschnittlich über dem zur Konturierung kritisch betrachteten, aber durchaus üblichen Schwellwert von $SUV \geq 2,5$ läge [73] und daher nicht-tumoröses Lebergewebe in das metabolische Tumervolumen einfließen würde. Um den Einfluss statistischer Ausreißer auf diese Subanalyse bezüglich der Standorteffekte bzw. Scannereffekte zu minimieren, wurde eine Mindestzahl von 21 Fällen für jeden untersuchten Standort und jedes untersuchte PET-Scannermodell festgelegt. Durch diese quantitative Hürde soll über eine ausreichende Größe der Stichproben aus der Grundgesamtheit aller Patienten mit neu diagnostiziertem aggressivem Lymphom der Einfluss von patientenbezogenen Schwankungen minimiert werden.

Ein weiterer Einflussfaktor auf die gemessenen SUV_{mean} des rechten Leberlappens und in weiterer Folge auf den PERCIST-Schwellwert ist die Streuung der Vernichtungsstrahlung aus den Positronen gerade bei adipösen Patienten. Deswegen kann eine hohe Körpermasse einen Einfluss auf die Messunsicherheit des PET-Signals haben. Aus den zur Verfügung gestellten Daten eignete sich das Körpergewicht der untersuchten Patienten am besten als Surrogatmarker, um diesen Einflussfaktor zu evaluieren. Eine BMI-Berechnung mittels der Körpergröße würde im Vergleich zum Körpergewicht präferabel sein, jedoch wurden nur sporadisch vollständige Daten zur Körpergröße der Patienten übermittelt und dokumentiert. Gemäß der hier gezeigten Auswertung zeigt das Körpergewicht sowohl für den SUV_{mean} der Leber ($r=0,30$; $p<0,001$) als auch für den PERCIST-Schwellwert ($r=0,37$; $p<0,001$) eine moderat positive Korrelation. Aufgrund der hohen untersuchten Stichprobengröße kann dieser Korrelation hohen Wert zugeschrieben werden. Ein relevanter Einfluss der applizierten Radionuklidosis [MBq] auf die gemessene Standardabweichung der SUV-Messung im rechten Leberlappen konnte mit einer nicht signifikanten Pearson-Korrelation ausgeschlossen werden.

5.3 Evaluation der Ergebnisse zur Normalverteilung von SUV_{mean} , SUV_{SD} und dem PERCIST-Schwellwert

In dieser Arbeit wurde die Frage, ob SUV_{mean} sowie SUV_{SD} der Leber und/oder der PERCIST-Schwellwert normalverteilt sind, mit einem statistischen Test und einer visuellen Methode analysiert. Die zugehörigen Histogramme und QQ-Diagramme für den SUV_{mean} und PERCIST lassen den Eindruck einer Gaußschen Normalverteilung zu, während für den SUV_{SD} eine logarithmische Normalverteilung wahrscheinlicher erscheint. Der KS-Test bestätige bei

SUV_{mean} eine Gaußsche Normalverteilung und bei SUV_{SD} eine logarithmische Normalverteilung, beim PERCIST-Schwellwert wird ein Gaußsches Normalverteilungsmuster trotz des visuellen Eindrucks knapp abgelehnt ($p=0,04$). Da der PERCIST-Schwellwert sich sowohl aus SUV_{mean} als auch aus dem SUV_{SD} zusammensetzt, welche beide unterschiedlichen Normalverteilungsmustern folgen, ist die knappe Ablehnung durch den KS-Test wenig verwunderlich. Die visuelle Beurteilung einer Gaußschen Normalverteilung ist beim PERCIST dem KS-Test vorzuziehen. Fachliteratur, wie die Arbeiten von *Boellaard et. al* (2004) gehen von einer logarithmischen Normalverteilung der SUV-Werte aus, welche zumindest für den SUV_{mean} der Leber in dieser Arbeit nicht bestätigt werden konnte, für den SUV_{SD} der Leberreferenzregion jedoch schon [74].

5.4 Evaluation des visuellen Scorings und der ermittelten Werte der Biomarker

Der hier implementierte Bewertungsalgorithmus konnte in 97,7 % der Fälle das Vorliegen bzw. das Nichtvorliegen eines Tumors im Bilddatensatz dem visuellen Eindruck des Untersuchers entsprechen. In nur 2,3 % der Fälle wurde durch den Bewertungsalgorithmus ein MTV von 0 mL ermittelt, während die visuelle Evaluation zu einem $MTV > 0$ mL tendiert hätte. In Anbetracht dieser Zahlen zeigt die hier gewählte Auswertemethode eine hohe Übereinstimmung mit dem visuellen Eindruck des Untersuchers. Insgesamt dominierte in der Gesamtauswertung Eindruck, dass das metabolische Tumolvolumen häufiger unterschätzt als überschätzt wurde, insgesamt wurde der Eindruck einer Unterschätzung 4,7x häufiger getroffen als der Eindruck einer Überschätzung.

Insgesamt zeigte sich sowohl bei dem MTV als auch bei der TLG ein deutliches Verteilungsmuster zugunsten niedriger Werte. Das MTV zeigte mit der in dieser Arbeit eingesetzten Auswertemethode arithmetische Mittel und Mediane, die mit denen weiterer Studien vergleichbar sind, welche eine Auswertung solitär mit dem Kriterium SUV_{max} (41%) vollzogen hatten. Jedoch lagen die Mittelwerte etwas unter diesen anderer Arbeiten [67, 72, 75]. Zudem ist besonders in den Arbeiten von *Ilyas et. al* (2018) ersichtlich, dass die Konturierung mittels SUV_{max} (41%) deutlich niedrigere metabolischen Tumorumfänge erzeugt, als eine Selektion mittels absoluter Schwellwerte und dem PERCIST-Schwellwert allein [72]. Dies lässt vermuten, dass unter den eingesetzten Auswertungskriterien dieser Arbeit die relative Konturierungsmethode mittels SUV_{max} (41%) die größte Reduktion des ermittelten metabolischen Tumorumfanges verursacht. Eine Veranschaulichung verschiedener Konturierungsmethoden und deren Einfluss auf die metabolischen Tumorumfänge bietet die Abbildung 4 in Kapitel 3.5 im Methodenteil.

Bei multiplen kleinen Lymphomherden konnte das Mindestvolumen ($\geq 1\text{cm}^3$) einen starken Einfluss auf das ermittelte niedrige MTV haben. Letzteres wird durch die Analyse der reglementierenden Faktoren in der Subgruppe der Bilddatensätze MTV=0mL und einer Unterschätzung von „-2“ im visuellen Scoring verdeutlicht (Kap. 4.1.2). Das Nichterreichen des Mindestvolumens war für den Großteil der nicht in die Endauswertung aufgenommenen Läsionen verantwortlich.

Während Autoren wie *Martín-Saladich et. al (2020)* einen Vorteil in der Selektion mittels absoluten SUV-Schwellwerten durch ihre hohe Reproduzierbarkeit sehen [67], zeigen die Ergebnisse dieser Arbeit, dass die hier gewählte Selektionsmethode sich in der Interobserver-Analyse ebenfalls als sehr reproduzierbar erwiesen hat. Zusätzlich dazu zeigte sich der Vorteil, dass eine Fehlkonturierung von im PET-Bild unauffälligen Geweben besser unterbunden werden konnte, da nur 18/661 metabolische Tumorumfänge der Fälle einen für den Untersucher überschätzten Eindruck machten.

Besonders eindrücklich präsentiert sich der Unterschied der MTV und TLG-Werte, welche zuvor gemäß der Items des IPI (Kap. 2.2.5.1) in eine „favourable“ (bei IPI=1) und eine „non-favourable“-Gruppe (bei IPI von 2-5) in der OPTIMAL>60-Studie risikostratifiziert wurden (Kap. 4.10). Da der IPI, bzw. der R-IPI („revised International Prognostic Index“) durch mehrere Studien mit hohem Evidenzgrad als prognostischer Faktor bei Patienten mit DLBCL [64, 76] und follikulären Lymphomen („FLIPI“) [77] bestätigt wurde, ist die starke Korrelation dieser Risikostratifizierung mit den hier ermittelten Biomarkern MTV und TLG ein weiteres Indiz für deren Wertigkeit als prognostische Faktoren. Das kann allerdings erst mit Vorliegen von Überlebensdaten weiter analysiert werden.

Die extrahierten Werte für den SUV_{max} im PET-Bilddatensatz zeigten ebenfalls Unterschiede im arithmetischen Mittel und Median zwischen den Risikogruppen, jedoch kann diese im Zuge dieser Arbeit als weniger prägnant bewertet werden als für MTV und TLG. Im systematischen Review von *Frood et. al (2021)* bezüglich der prognostischen Wertigkeit von MTV, TLG und SUV_{max} erwies sich SUV_{max} in nur einem Viertel (5/20) der Studien als prognostischer Faktor, eine deutlich geringere Quote als für MTV und TLG [64]. Das MTV unterschied sich im arithmetischen Mittel um mehr als 200 mL zwischen den Risikogruppen, während sich der Median ebenfalls um mehr als 100 mL unterschied. Ähnliche Diskrepanzen ergeben sich bei der TLG (Kap. 4.6). Besonders eindrücklich ist auch, dass das höchste MTV in der Gruppe mit günstigem Risiko bei 200,3 mL lag und damit sogar unterhalb des arithmetischen Mittels der Gruppe mit ungünstigem Risiko. Die metabolischen Tumorumfänge der günstigen

Risikogruppe befanden sich gänzlich unter beliebten Cut-offs zur Risikostratifizierung wie z.B. 300 mL [78].

Eine ebenfalls interessante Beobachtung stellt die Anzahl der Fälle mit einem kalkulierten MTV von 0 dar. Häufiger kommt ein MTV von 0 mL in der Gruppe vor, welche in der vorangegangenen Risikostratifizierung als günstig bewertet wurde (48/183, 25,6 %). Im Vergleich mit der Gruppe mit ungünstigem Risiko ergab die Auswertung der als günstig stratifizierten Gruppe 5x häufiger ein MTV von 0 mL (20/473, 5,0 %).

Die hier ermittelten Werte für MTV und TLG zeigen starke Unterschiede je nach Gruppierung gemäß den IPI-Kriterien. Dies wirft die Frage auf, ob die Einbeziehung eines dieser beiden Biomarker in einen neuen Prognoseindex eine Verbesserung der prognostischen Aussagekraft nach sich ziehen könnte. Die retrospektive Kohortenstudie der Arbeitsgruppe um *Mikhaeel et. al* (2022) implementierte ein sogenanntes „MTV age stage model“, ein Vorhersagemodell, welches die Verwendung von MTV und einzelnen IPI-Komponenten kombinierte und bessere Ergebnisse bezüglich der OS und PFS erzielte, als der IPI [5].

5.5 Limitationen und potenzielle Fehlerquellen

Bei dieser Dissertation handelt es sich um eine Auswertung von Bilddaten zur Gewinnung bildbasierter Biomarker zur weiteren Verwendung in der prospektiv geplanten Analyse der multizentrischen OPTIMAL>60-Studie. Gegenüber vielen anderen Studien, welche eine retrospektive Selektion von Fällen vorgenommen haben, ist die hier genutzte Datenbasis grundsätzlich als hochwertig anzusehen. Dennoch mag es eine Reihe von Limitationen geben.

Lediglich 661 der 1152 Patienten, welche an der OPTIMAL>60-Studie teilnahmen, erhielten eine initiale FDG-PET/CT-Untersuchung. Während seit vielen Jahren die Basisdiagnostik mittels PET/CT internationaler Standard bei der Versorgung aggressiver Lymphome ist, steht diese Untersuchung in Deutschland erst seit 2021 als reguläre Leistung der gesetzlichen Krankenversicherung zur Verfügung. Daher ist zu hoffen, dass zukünftig die in der S3-Leitlinie empfohlene Ausbreitungsdiagnostik mittels PET durchgängig eingesetzt wird.

Im Vergleich zu weiteren wissenschaftlichen Arbeiten, welche zu dem metabolischen Tumolvolumen bei Patienten mit NHL arbeiteten, ist die Datengrundlage dieser Dissertation in der Gruppe der retrospektiven Kohortenstudien bzw. der retrospektiven Datenanalysen mit n=661 quantitativ als hochwertig einzustufen.

Ebenfalls als qualitativ hochwertig im Sinne der Versorgungsrelevanz einzuschätzen ist, dass die multizentrische Natur dieser Studie eine hohe Varietät verschiedener PET-Scannermodelle aufweist. Auch die insgesamt hohe Anzahl an 83 verschiedenen Standorten, an welchen die PET-Bilddatensätze gewonnen wurde, ist ein Vorteil zu einer unizentrischen Studie vergleichbarer Größe. Eine mögliche Limitation stellt jedoch der Ausschluss von Patienten mit aggressiven NHL jünger als 60 Jahre dar, was durch die Konzeption der OPTIMAL>60-Studie bedingt ist. Aus epidemiologischer Sicht tritt diese Erkrankung jedoch häufiger bei Älteren auf (Kap. 2.2.2).

Die Informationen über verwendete PET-Scanner stammen aus den DICOM-Metadaten, zusätzlich erfolgte eine weitere Dokumentation in CRFs (siehe Kap. 11). Wie in Kap. 3.5 beschrieben, wurden im Zuge der Auswertung Namen von PET-Scannern, welche sich lediglich durch minimale Unterschiede wie einem Sonderzeichen bzw. einem Buchstaben (z.B. „Biograph 40“ und „Biograph 40_mCT“) unterscheiden in dieselbe Gruppe zugeteilt, da anzunehmen ist, dass es sich hierbei um dasselbe PET-Scannermodell handelt. Die Möglichkeit zur eigenständigen Überprüfung existierte jedoch nicht, da man hier auf die Richtigkeit der Angaben der beitragenden Studienzentren angewiesen ist. Dies stellte in einigen wenigen Fällen ein Problem dar, da in einigen PET-Bilddatensätzen in den DICOM-Feldern der Begriff „SOMATOM DEFINITION AS“ in dem Feld für das angewandte PET-Scannermodell auftaucht. Es ist jedoch davon auszugehen, dass mit „SOMATOM DEFINITION AS“ das Modell des zugehörigen Computertomographen der PET/CT in das Feld des PET-Scanners gelistet wurde. Die Angabe über das verwendete PET-Scannermodell fehlt daher in diesen Fällen. Da die Listung von „SOMATOM DEFINITION AS“ nur bei zuweisenden Studienzentren verwendet wurde, welche in weiteren korrekt eingetragenen DICOM-Feldern das PET-Scannermodell „Biograph 128“ von SIEMENS Healthineers angaben, liegt die Vermutung nahe, dass es sich um PET-Aufnahmen mittels dem „Biograph 128“-Modell handelt. Eine Aufnahme dieser Subgruppe in die Analyse der SUV_{mean} und PERCIST-Schwellwert nach PET-Scannermodell erfolgte jedoch nicht, in die Auswertung der Biomarker MTV, TLG und SUV_{max} und in den Standortvergleich des SUV_{mean} und der PERCIST-Schwellwerte flossen sie jedoch ein.

Eine Schwierigkeit bei der Auswertung mittels syngo.via stellten einzelne PET-Bilddatensätze dar, welche die PET-Komponente nicht SUV-skaliert darstellten, sondern nur mittels als „counts“ oder der Einheit „g/mL“ (n=10). Trotz dieser unterschiedlichen Einheiten wurde sowohl bei einer Einheitenangabe nach „counts“, als auch bei einer Angabe nach „g/mL“

mittels des Verhältnisses, welches analog zur Berechnung des PERCIST-Schwellwerts herangezogen wird, ein metabolisches Tumervolumen bestimmt. Als Vergleichsvolumen für dieses relative Auswertungsverfahren diene analog zum PERCIST-Schwellwert die Anzahl an „counts“ bzw. der Wert in g/mL im rechten Leberlappen.

Ebenfalls eine Sonderstellung in der Auswertung waren PET-Bilddatensätze mit separaten Akquisitionen verschiedener Körperareale, z.B. Torso und Beine. Dies stellte ein Problem für den PERCIST-Algorithmus dar, da in diesen Fällen oftmals eine zeitliche Differenz zwischen den Aufnahmen lag und die messbare FDG-Aktivität von der verstrichenen Zeit nach intravenöser Applikation abhängig ist. In weiterer Folge ist die Anwendung des PERCIST-Schwellwerts bei PET-Bilddatensätzen, welche mit Verzögerung aufgenommen wurden, einer größeren Unsicherheit in der Selektion von Läsionen unterworfen. Aus Mangel an Alternativen wurden bei diesen PET-Bildern die Biomarker trotzdem mittels des PERCIST-Schwellwerts bestimmt und flossen in die Auswertung mit ein.

Ein weiteres in dieser Arbeit beobachtetes Phänomen ist, dass bei massiver Tumordinfiltration in der Leber der gemessene SUV_{mean} der anatomischen Referenzregion durch einen „spill-over“ fälschlich in die Höhe getrieben wird. Eine Korrektur bei fälschlicher Miteinbeziehung von Tumorherden in die Leberreferenzregion wurde bei der Auswertung zwar standardmäßig vorgenommen, durch das konstante Volumen von 14,1 mL der kugelförmigen Referenzregion konnten aber in seltenen Einzelfällen mit diffusem Leberbefall die ubiquitären Tumordinfiltrationen stören, wie in Kap. 3.5 beschrieben (Abb. 5). Abhilfe könnte hierbei eine durch Tumordinfiltration begründete Ausnahme schaffen, das Volumen der kugelförmigen Referenzregion beliebig weit zu senken, bis das physiologische Lebergewebe ohne Fremdpathologien abgebildet werden kann. Die Größe des VOI wurde in den Arbeiten von *Boellaard et. al (2009)* ebenfalls als großer Einflussfaktor auf die Höhe des SUV eingestuft [79].

SUV_{mean} verschiedener Leberpathologien nicht-malignen Ursprungs wie z.B. die NAFLD („non-alcoholic fatty liver disease“) können ebenfalls den SUV_{mean} im rechten Leberlappen beeinflussen [80]. In einer retrospektiven Studie mit 238 Patienten verschiedener, meist tumorassoziierter Pathologien konnten *Keramida et. al (2019)* ebenfalls eine Korrelation der gemessenen SUV_{mean} mit dem Blutglukosewert und den pathologischen Effekten des Diabetes mellitus Typ 1 und 2 und der alkoholischen Fettleber nachgewiesen werden [81]. Hier können mangels begleitender Dokumentation keine genauen Aussagen zum etwaigen Einfluss von Leberpathologien auf den SUV_{mean} im rechten Leberlappen getroffen werden. Mit den

Ausschlusskriterien der OPTIMAL>60-Studie sind allerdings wesentliche hepatische Begleiterkrankungen in diesem Kollektiv nicht enthalten.

Die Zuordnung von Speicherungen zum MTV unterliegt Schwankungen. Die Differenzierung von physiologisch FDG-speicherndem, leicht entzündeten und potenziell tumorösen Gewebe im Waldeyer-Rachenring stellte beispielsweise eine Schwierigkeit dar. Hierbei ergeben sich besonders bei leicht vermehrtem FDG-Uptake in den palatinalen Tonsillen Entscheidungsschwierigkeiten, welche bei Fehlen einer Computertomographie zusätzlich verschärft werden [82].

Bei männlichen Patienten sorgten FDG-Anreicherungen in der Prostata für ein gewisses Maß an Unsicherheit. Da primäre Lymphome der Prostata mit 0,1 % aller lymphoproliferativen Krankheiten als auch maligne Metastasierungen in der Prostata selten sind [83], wurden diese Areale bei Nichteindeutigkeit verworfen und flossen nicht in die Kalkulation des MTV ein, auch wenn diese das Mindestvolumen und den PERCIST-Schwellwert erreichten. Prostatakarzinome sind typischerweise nicht FDG-avide, jedoch kann besonders bei Prostata-Cas mit einem Gleason-Score von über 7 Punkten ein hoher FDG-Uptake beobachtet werden [84].

Im Vergleich zu einer Manifestation in der Prostata stellen NHL im Hoden einen epidemiologisch bedeutsameren Anteil dar, welcher mit ca. 1-2 % aller NHL und 1-9 % aller Malignome des Hodens betitelt wird [85]. Da im Hoden ebenfalls ein physiologischer FDG-Uptake beobachtet werden kann, ist die Befundung hier ebenfalls mit einer gewissen Unsicherheit behaftet. Die Arbeiten von *Harper et. al* (2021) versuchten in einer retrospektiven Studie die testikuläre FDG-Anreicherung bei 350 Patienten, welche aufgrund verschiedener Malignitäten hospitalisiert wurden, zu quantifizieren und so einen besseren Überblick über als physiologisch zu wertende SUV_{mean} und SUV_{max} zu bekommen [86]. Eine Limitation dieses Reviews ist die monozentrische Natur und ihrer damit limitierten Varietät an PET-Scannern, da wie in Kap. 4.3 dieser Dissertation festgestellt wurde, sich die SUV_{mean} der Leber verschiedener untersuchter PET-Scannermodelle statistisch signifikant voneinander unterscheiden. Versuche, physiologische SUV-Werte für anatomische Areale des Körpers zu definieren, werden dadurch erheblich erschwert.

Im Allgemeinen zeigte sich während der Auswertungen, dass PET-Bilddatensätze mit einer korrelierenden Bildgebung wie einer low-dose-CT, Kontrastmittel-CT oder einer MRT sich mit größerer Sicherheit befunden ließen als PET-Bilddatensätze, bei dem diese fehlten und nur das

PET-Bild als Grundlage zur Befundung verwendet wurde. Dem Großteil der ausgewerteten Bilddatensätze war eine low-dose-CT beigelegt. Zum Start der Studie waren noch alleinige PETs zulässig, während der Studienlaufzeit wurde jedoch im Amendment nur noch die Bereitstellung von PET/CTs und PET/MRTs gestattet. Eine Kontrastmittel-CT, eine MRT oder ein PET-Bilddatensatz ohne beigelegte korrelierende Bildgebung stellten daher eine Ausnahme dar, wobei jedoch auf eine Quantifizierung dieser verzichtet wurde. In einer Vergleichsstudie von *Fueger et. al* (2009) wurde die Genauigkeit der Erkennung von Tumormanifestationen bei Patienten mit indolenten NHL untersucht und die Überlegenheit des PET/CT zu einer einzelnen PET-Aufnahme bestätigt [87], was den subjektiven Eindruck der schlechteren Befundbarkeit der solitären PET stützt.

Tumordinfiltrationen bzw. Tumorumwucherungen der harnableitenden Organe konnten ein Problem bei der Konturierung verdächtiger Läsionen darstellen, da die FDG renal eliminiert wird und dadurch hier eine physiologisch hohe FDG-Anreicherung herrscht. Im Bild entstand so bei Tumordinfiltration bzw. -umwucherung das Bild einer „Verschmelzung“ physiologischer und tumoröser Areale, welche gelegentlich nicht klar voneinander trennbar waren und in weiterer Folge eine Überschätzung des metabolischen Tumorumfanges begünstigte. In der Intersoftware- und Interobserveranalyse mit ROVER war dieses Phänomen durch die erschwerte Konturierung der Lymphknotenherde in syngo.via für 3 von 4 MTV-Differenzen von über 500 mL verantwortlich.

Des Weiteren stellt die physiologische FDG-Anreicherung der glatten Darmmuskulatur bei Kontraktion ein differenzialdiagnostisches Problem dar, so wie die FDG-Anreicherung im Darm bei mit Metformin behandelten Patienten [88, 89].

Septische Leberläsionen, welche im FDG-avid imponieren, können auch zur Verfälschung der gemessenen MTV führen [90]. Auch der „Tumor-sink-Effekt“, welcher durch hochvolumige metabolisch aktive Tumormanifestationen verursacht wird und in weiterer Folge Tumorgewebe mit niedrigerer metabolischer Aktivität maskieren kann (Kap. 4.13), stellt eine mögliche Schwierigkeit bei der Erfassung metabolischer Tumorumfänge dar, sowohl bei aggressiven NHL als auch bei Tumoren anderer FDG-avider Entitäten.

Da ein Datenembargo bezüglich der Überlebensdaten erst zum 18.01.2025 ausgelaufen ist, können hier keine Überlebensanalysen berichtet werden. Diese sind erforderlich, um die prognostische Aussagekraft der hier erhobenen Parameter weiter zu bewerten. Dieses muss weitergehenden Analysen vorbehalten bleiben.

6. Zusammenfassung

Im Rahmen der vorliegenden Arbeit sollte die Validität und Reproduzierbarkeit eines schwellwertbasierten Konturierungsverfahrens zur Ermittlung des metabolischen Tumervolumen bei NHL-Patienten in der initialen PET/CT geprüft werden. Die Qualität der Konturierung wurde visuell mit einem semiquantitativen 5-stufigen Score beurteilt. Zudem wurde der Zusammenhang der metabolischen Tumervolumina mit der vorangegangenen Risikostratifizierung der Patienten gemäß den IPI-Kriterien geprüft und der mögliche Einflussfaktor des PET-Standorts bzw. des PET-Scannermodells auf die Höhe der PERCIST-Schwellwerte evaluiert.

Bei der Auswertung aller in die Gesamtauswertung einfließender Fälle zeigen sowohl MTV ($n=661$) als auch TLG ($n=651$) eine Verteilung zugunsten niedriger Werte, statistische Ausreißer wurden nur in 2,3% der Fälle der MTV und 1,5% der Fälle der TLG dokumentiert. Die Einstufung der metabolischen Tumervolumina gemäß dem visuellen Scoring brachte in 84,3% (557/661) der Fälle eine starke Übereinstimmung mit dem kalkulierten MTV, in 12,8% (85/661) der Fälle den Eindruck einer Unterschätzung des MTV und in 2,8% (19/661) den Eindruck einer Überschätzung des MTV. Eine geringe Zahl von 15 der 661 (2,4%) bearbeiteten Fälle wurde vom Observer als PET-positiv bewertet, während das Selektionsverfahren ein MTV von 0 mL kalkulierte. Grund hierfür war in 11/15 Fällen die Nichterfüllung des Mindestvolumens von 1 mL und in 4/15 Fällen die Nichterfüllung des PERCIST-Schwellwerts der Läsionen. Das MTV und die TLG unterscheiden sich signifikant voneinander je vorangegangener Risikostratifizierung ($p<0,001$), ein durch das Selektionsverfahren festgestelltes metabolisches Tumervolumen von 0 mL weisen 26,2% (48/183) der „favourable“-Gruppe und 4,2% (20/472) der „less favourable“-Gruppe auf.

Der Interobserver-Vergleich einer repräsentativen Stichprobe der Gesamtauswertung ($n=60$) zeichnet sich durch ein sehr hohes Agreement zu den extrahierten Biomarkern MTV ($r=0,999$; $p<0,001$), TLG ($r=0,999$, $p<0,001$) und SUV_{max} ($r=0,997$; $p<0,001$) aus. Der PERCIST-Schwellwert ($r=0,96$; $p<0,001$) und der SUV_{mean} ($r=0,95$; $p<0,001$) des rechten Leberlappens zeigten ebenfalls eine hohe Korrelation. Im Rahmen des visuellen Scorings tendierten beide Observer eher zu einer Unterschätzung des MTV als zu einer Überschätzung, die Absolutwerte unterschieden sich jedoch deutlich (Observer 1: 4/60; Observer 2: 26/60). Um das visuelle Scoring zu optimieren, erscheint eine Paralleldokumentation der Volumina von nicht eingeschlossenen Lymphomherden als sinnvoll. Zudem konnte eine hohe Pearson-Korrelation der extrahierten Messwerte für MTV ($r=0,95$; $p<0,001$), TLG ($r=0,96$; $p<0,001$) und SUV_{max}

($r=0,97$; $p<0,001$) in einer zweiten Interobserveranalyse mit der Software „ROVER“ bei leicht unterschiedlicher Konturierungsmethodik aufgezeigt werden.

Die Höhe des SUV_{mean} des rechten Leberlappens sowie die Höhe des PERCIST-Schwellwerts unterlagen einem signifikanten Unterschied nach eingesetztem PET-Scannermodell ($p<0,001$), derselbe Effekt konnte im Standortvergleich jedoch nur bei den PERCIST-Schwellwerten dargestellt werden ($p=0,023$). Eine moderate Korrelation des Körpergewichts mit dem zur Berechnung des PERCIST-Schwellwerts herangezogenen SUV_{SD} konnte hergestellt werden ($r=0,44$; $p<0,001$), für den SUV_{mean} ($r=0,30$; $p<0,001$) und PERCIST ($r=0,37$; $p<0,001$) ebenso. Der Wert der Bildgebungsverzögerung zeigt keine signifikante Korrelation zur Höhe des SUV_{mean} ($p=0,64$) und dem PERCIST-Schwellwert ($p=0,31$). Die Höhe der MTV-Last zeigte eine schwache, negative Korrelation zu den kalkulierten SUV_{mean} -Werten der Leber ($r=-0,32$; $p<0,001$) und dem PERCIST-Schwellwert ($r=-0,27$; $p<0,001$). Aufgrund der Vielzahl an Einflussfaktoren auf die Höhe der gemessenen SUV im PET-Bild erscheint eine relative Selektionsmethode auf Basis des PERCIST-Schwellwerts sinnvoller als eine Kalkulation mittels eines absoluten SUV-Grenzwerts wie $SUV \geq 2,5$ oder $SUV \geq 4$.

Die in dieser Arbeit aufwändig erzeugten bildbasierten Biomarker eignen sich für weitergehende Analysen. Diese werden jedoch erst mit aktuellem Wegfall des Datenembargos der OPTIMAL>60-Studie möglich, da bislang keine Überlebensdaten zur Verfügung stehen, um die prognostische Aussagekraft der hier erhobenen Parameter weiter zu bewerten.

7. Literatur

- 1 *International Atomic Energy Agency (IAEA)*. IMAGINE - PET scanners (per 1 mil) (2021). Im Internet: <https://www.iaea.org/resources/hhc/nuclear-medicine/databases/imagenet/pet-scanners>; Stand: 15.02.25
- 2 *Gallach M, Mikhail Lette M, Abdel-Wahab M et al*. Addressing Global Inequities in Positron Emission Tomography-Computed Tomography (PET-CT) for Cancer Management: A Statistical Model to Guide Strategic Planning. *Medical science monitor : international medical journal of experimental and clinical research* 2020; 26: e926544
- 3 *Hellwig D, Dührsen U*. Non-Hodgkin-Lymphome – PET in Diagnostik und Therapiesteuerung. *Onkologie* 2019; 25: 867 – 879
- 4 *Moskowitz AJ, Schöder H, Gavane S et al*. Prognostic significance of baseline metabolic tumor volume in relapsed and refractory Hodgkin lymphoma. *Blood* 2017; 130: 2196 – 2203
- 5 *Mikhaeel NG, Heymans MW, Eertink JJ et al*. Proposed New Dynamic Prognostic Index for Diffuse Large B-Cell Lymphoma: International Metabolic Prognostic Index. *JCO* 2022; 40: 2352 – 2360
- 6 *Storck K, Brandstetter M, Keller U et al*. Clinical presentation and characteristics of lymphoma in the head and neck region. *Head & face medicine* 2019; 15: 1
- 7 *Swerdlow SH, Campo E, Pileri SA et al*. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* 2016; 127: 2375 – 2390
- 8 *Robert Koch Institut*. Krebs in Deutschland | 2019/2020 (07.02.2023). Im Internet: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/krebs_in_deutschland_node.html; Stand: 15.02.2024
- 9 *International Agency for Research on Cancer*. Non-Hodgkin lymphoma. Im Internet: <https://gco.iarc.fr/today/data/factsheets/cancers/34-Non-hodgkin-lymphoma-factsheet.pdf>; Stand: 24.11.2021
- 10 *Shankland KR, Armitage JO, Hancock BW*. Non-Hodgkin lymphoma. *The Lancet* 2012; 380: 848 – 857
- 11 *Thandra KC, Barsouk A, Saginala K et al*. Epidemiology of Non-Hodgkin's Lymphoma. *Medical sciences (Basel, Switzerland)* 2021; 9
- 12 *Leval L de, Jaffe ES*. Lymphoma Classification. *Cancer J* 2020; 26: 176 – 185
- 13 *Smedby KE, Hjalgrim H, Askling J et al*. Autoimmune and chronic inflammatory disorders and risk of non-Hodgkin lymphoma by subtype. *JNCI: Journal of the National Cancer Institute* 2006; 98: 51 – 60

- 14 *Cinoku II, Verrou K-M, Piperi E et al.* Interferon (IFN)-stimulated gene 15: A novel biomarker for lymphoma development in Sjögren's syndrome. *Journal of autoimmunity* 2021; 123: 102704
- 15 *Young LS, Murray PG.* Epstein-Barr virus and oncogenesis: from latent genes to tumours. *Oncogene* 2003; 22: 5108 – 5121
- 16 *Morgner A, Bayerdörffer E, Thiede C et al.* Bedeutung der Helicobacter-pylori-Infektion für die Pathogenese und Therapie von MALT-Lymphomen des Magens. *Visc Med* 2002; 18: 24 – 28
- 17 *Boerma E-J, Siebert R, Baudis M et al.* Translocations Involving 8q24 in Burkitt Lymphoma and Other Malignant Lymphomas: A Historical Review of Cytogenetics in the Light of Today's Knowledge. *Blood* 2008; 112: 2814
- 18 *Saleh K, Michot J-M, Camara-Clayette V et al.* Burkitt and Burkitt-Like Lymphomas: a Systematic Review. *Current oncology reports* 2020; 22: 33
- 19 *Burlet B, Ramla S, Fournier C et al.* Identification of novel, clonally stable, somatic mutations targeting transcription factors PAX5 and NKX2-3, the epigenetic regulator LRIF1, and BRAF in a case of atypical B-cell chronic lymphocytic leukemia harboring a t(14;18)(q32;q21). *Cold Spring Harbor molecular case studies* 2021; 7
- 20 *Bakhshi TJ, Georgel PT.* Genetic and epigenetic determinants of diffuse large B-cell lymphoma. *Blood cancer journal* 2020; 10: 123
- 21 *Lai R, Medeiros LJ.* Pathologic diagnosis of mantle cell lymphoma. *Clinical lymphoma* 2000; 1: 197-206; discussion 207-8
- 22 *Roos AJ de, Schinasi LH, Miligi L et al.* Occupational insecticide exposure and risk of non-Hodgkin lymphoma: A pooled case-control study from the InterLymph Consortium. *International journal of cancer* 2021; 149: 1768 – 1786
- 23 *Wahner-Roedler DL, Witzig TE, Loehrer LL et al.* Gamma-heavy chain disease: review of 23 cases. *Medicine* 2003; 82: 236 – 250
- 24 *Gleissner B, Küppers R, Siebert R et al.* Report of a workshop on malignant lymphoma: a review of molecular and clinical risk profiling. *British journal of haematology* 2008; 142: 166 – 178
- 25 *Cancer Research UK.* Low grade NHL (20.11.2020). Im Internet: <https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/low-grade>; Stand: 29.11.2021

- 26 *Cancer Research UK*. High grade NHL (24.11.2020). Im Internet:
<https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/high-grade>; Stand: 29.11.2021
- 27 *Dreyling M, Hrsg*. Maligne Lymphome. Empfehlungen zur Diagnostik, Therapie und Nachsorge. 11. Aufl. München: W.Zuckerschwerdt Verlag GmbH München; 2019
- 28 *Liu Y, Barta SK*. Diffuse large B-cell lymphoma: 2019 update on diagnosis, risk stratification, and treatment. *American journal of hematology* 2019; 94: 604 – 616
- 29 *Ziepert M, Hasenclever D, Kuhnt E et al*. Standard International prognostic index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. *JCO* 2010; 28: 2373 – 2380
- 30 *Sehn LH, Chhanabhai M, Fitzgerald C et al*. Revised International Prognostic Index (R-IPI) Is a Better Predictor of Outcome Than the Standard IPI for Patients with Diffuse Large B-Cell Lymphoma (DLBCL) Treated with Rituximab and CHOP (R-CHOP). *Blood* 2005; 106: 492
- 31 *Klapper W, Kreuz M, Kohler CW et al*. Patient age at diagnosis is associated with the molecular characteristics of diffuse large B-cell lymphoma. *Blood* 2012; 119: 1882 – 1887
- 32 *Chapuy B, Stewart C, Dunford AJ et al*. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nature medicine* 2018; 24: 679 – 690
- 33 *Schalhorn B*. Double-Hit-Lymphome: intensive Erstlinientherapie entscheidend. Im *Focus Onkologie* 2018; 21: 22
- 34 *Kühnl A, Cunningham D, Counsell N et al*. Outcome of elderly patients with diffuse large B-cell lymphoma treated with R-CHOP: results from the UK NCRI R-CHOP14v21 trial with combined analysis of molecular characteristics with the DSHNHL RICOVER-60 trial. *Annals of oncology : official journal of the European Society for Medical Oncology* 2017; 28: 1540 – 1546
- 35 *Pfreundschuh M, Trümper L, Österborg A et al*. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *The Lancet Oncology* 2006; 7: 379 – 391
- 36 *Held G, Murawski N, Ziepert M et al*. Role of radiotherapy to bulky disease in elderly patients with aggressive B-cell lymphoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2014; 32: 1112 – 1118

- 37 *Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF (Leitlinienprogramm Onkologie)*. Therapie und Nachsorge für erwachsene Patient*innen mit einem diffusen großzelligen B-Zell-Lymphom und verwandten Entitäten, Langversion 1.0, 2022; AWMF-Registernummer: 018/038OL (2022). Im Internet: <https://www.leitlinienprogramm-onkologie.de/leitlinien/dlbcl>; Stand: 15.02.25
- 38 *Poeschel V, Held G, Ziepert M et al.* Excellent Outcome of Young Patients (18-60 years) with Favourable-Prognosis Diffuse Large B-Cell Lymphoma (DLBCL) Treated with 4 Cycles CHOP Plus 6 Applications of Rituximab: Results of the 592 Patients of the Flyer Trial of the Dshnhl/GLA. *Blood* 2018; 132: 781
- 39 *Dührsen U, Müller S, Hertenstein B et al.* Positron Emission Tomography-Guided Therapy of Aggressive Non-Hodgkin Lymphomas (PETAL): A Multicenter, Randomized Phase III Trial. *JCO* 2018; 36: 2024 – 2034
- 40 *Sohani AR, Hasserjian RP.* Diagnosis of Burkitt Lymphoma and Related High-Grade B-Cell Neoplasms. *Surgical pathology clinics* 2010; 3: 1035 – 1059
- 41 *Perkins AS, Friedberg JW.* Burkitt lymphoma in adults. *Hematology* 2008: 341 – 348
- 42 *Ferry JA.* Burkitt's lymphoma: clinicopathologic features and differential diagnosis. *The oncologist* 2006; 11: 375 – 383
- 43 *Nie M, Wang Y, Bi X-W et al.* Effect of rituximab on adult Burkitt's lymphoma: a systematic review and meta-analysis. *Annals of hematology* 2016; 95: 19 – 26
- 44 *Thomas DA, Faderl S, O'Brien S et al.* Chemoimmunotherapy with hyper-CVAD plus rituximab for the treatment of adult Burkitt and Burkitt-type lymphoma or acute lymphoblastic leukemia. *Cancer* 2006; 106: 1569 – 1580
- 45 *Wang MY, Qu Y, Hu DR et al.* Methotrexate-loaded biodegradable polymeric micelles for lymphoma therapy. *International journal of pharmaceutics* 2019; 557: 74 – 85
- 46 *June CH, Sadelain M.* Chimeric Antigen Receptor Therapy. *The New England journal of medicine* 2018; 379: 64 – 73
- 47 *Raje N, Berdeja J, Lin Y et al.* Anti-BCMA CAR T-Cell Therapy bb2121 in Relapsed or Refractory Multiple Myeloma. *The New England journal of medicine* 2019; 380: 1726 – 1737
- 48 *Hallek M.* Chronic lymphocytic leukemia: 2020 update on diagnosis, risk stratification and treatment. *American journal of hematology* 2019; 94: 1266 – 1287
- 49 *Glass B, Hasenkamp J, Wulf G et al.* Rituximab after lymphoma-directed conditioning and allogeneic stem-cell transplantation for relapsed and refractory aggressive non-

- Hodgkin lymphoma (DSHNHL R3): an open-label, randomised, phase 2 trial. *The Lancet Oncology* 2014; 15: 757 – 766
- 50 *Zahid U, Akbar F, Amaraneni A et al.* A Review of Autologous Stem Cell Transplantation in Lymphoma. *Current hematologic malignancy reports* 2017; 12: 217 – 226
 - 51 *Schmitz N, Pfistner B, Sextro M et al.* Aggressive conventional chemotherapy compared with high-dose chemotherapy with autologous haemopoietic stem-cell transplantation for relapsed chemosensitive Hodgkin's disease: a randomised trial. *The Lancet* 2002; 359: 2065 – 2071
 - 52 *Cheson BD.* Staging and response assessment in lymphomas: the new Lugano classification. *Chinese clinical oncology* 2015; 4: 5
 - 53 *Cheson BD, Fisher RI, Barrington SF et al.* Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *JCO* 2014; 32: 3059 – 3068
 - 54 *Mohnike W, Hör G, Hertel A, Schelbert H.* PET/CT-Atlas. Berlin, Heidelberg: Springer Berlin Heidelberg; 2016
 - 55 *Thanarajasingam G, Bennani-Baiti N, Thompson CA.* PET-CT in Staging, Response Evaluation, and Surveillance of Lymphoma. *Curr. Treat. Options in Oncol.* 2016; 17: 24
 - 56 *Kaddu-Mulindwa D, Altmann B, Held G et al.* FDG PET/CT to detect bone marrow involvement in the initial staging of patients with aggressive non-Hodgkin lymphoma: results from the prospective, multicenter PETAL and OPTIMAL60 trials. *European journal of nuclear medicine and molecular imaging* 2021; 48: 3550 – 3559
 - 57 *Honda T, Maruyama D, Kurihara H et al.* 288O Role of FDG-PET/CT and gastrointestinal endoscopy in the staging of diffuse large B-cell lymphoma (DLBCL). *Annals of Oncology* 2015; 26: ix85
 - 58 *Capobianco N, Meignan M, Cottreau A-S et al.* Deep-Learning 18F-FDG Uptake Classification Enables Total Metabolic Tumor Volume Estimation in Diffuse Large B-Cell Lymphoma. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* 2021; 62: 30 – 36
 - 59 *Yamanaka S, Miyagawa M, Sugawara Y et al.* The prognostic significance of whole-body and spleen MTV (metabolic tumor volume) scanning for patients with diffuse large B cell lymphoma. *Int J Clin Oncol* 2021; 26: 225 – 232

- 60 *Vercellino L, Cottreau A-S, Casasnovas O et al.* High total metabolic tumor volume at baseline predicts survival independent of response to therapy. *Blood* 2020; 135: 1396 – 1405
- 61 *Liang J-H, Zhang Y-P, Xia J et al.* Prognostic Value of Baseline and Interim Total Metabolic Tumor Volume and Total Lesion Glycolysis Measured on 18F-FDG PET-CT in Patients with Follicular Lymphoma. *Cancer Res Treat* 2019; 51: 1479 – 1487
- 62 *Metser U, Dubebout J, Baetz T et al.* 18 F-FDG PET/CT in the staging and management of indolent lymphoma: A prospective multicenter PET registry study. *Cancer* 2017; 123: 2860 – 2866
- 63 *Song M-K, Chung J-S, Lee J-J et al.* Metabolic tumor volume by positron emission tomography/computed tomography as a clinical parameter to determine therapeutic modality for early stage Hodgkin's lymphoma. *Cancer science* 2013; 104: 1656 – 1661
- 64 *Frood R, Burton C, Tsoumpas C et al.* Baseline PET/CT imaging parameters for prediction of treatment outcome in Hodgkin and diffuse large B cell lymphoma: a systematic review. *European journal of nuclear medicine and molecular imaging* 2021; 48: 3198 – 3220
- 65 *O JH, Wahl RL.* PERCIST in Perspective. *Nucl Med Mol Imaging* 2018; 52: 1 – 4
- 66 *Eude F, Toledano MN, Vera P et al.* Reproducibility of Baseline Tumour Metabolic Volume Measurements in Diffuse Large B-Cell Lymphoma: Is There a Superior Method? *Metabolites* 2021; 11
- 67 *Martín-Saladich Q, Reynés-Llompart G, Sabaté-Llobera A et al.* Comparison of different automatic methods for the delineation of the total metabolic tumor volume in I-II stage Hodgkin Lymphoma. *Sci Rep* 2020; 10: 12590
- 68 *Meignan M, Sasanelli M, Casasnovas RO et al.* Metabolic tumour volumes measured at staging in lymphoma: methodological evaluation on phantom experiments and patients. *Eur J Nucl Med Mol Imaging* 2014; 41: 1113 – 1122
- 69 *Thurner L, Pfreundschuh† M.* Therapie der aggressiven B-Zell-Lymphome. *Trillium Krebsmedizin* 2015: 140 – 147
- 70 *Pfreundschuh† M.* Improvement of Outcome and Reduction of Toxicity in Elderly Patients with CD20+ Aggressive B-Cell Lymphoma by an Optimised Schedule of the Monoclonal Antibody Rituximab, Substitution of Conventional by Liposomal Vincristine, and FDG-PET based Reduction of Therapy in Combination with Vitamin D Substitution; 2020

- 71 *Boellaard R, O'Doherty MJ, Weber WA et al.* FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *European journal of nuclear medicine and molecular imaging* 2010; 37: 181 – 200
- 72 *Ilyas H, Mikhaeel NG, Dunn JT et al.* Defining the optimal method for measuring baseline metabolic tumour volume in diffuse large B cell lymphoma. *Eur J Nucl Med Mol Imaging* 2018; 45: 1142 – 1154
- 73 *Kwee TC, Cheng G, Lam MGEH et al.* SUVmax of 2.5 should not be embraced as a magic threshold for separating benign from malignant lesions. *Eur J Nucl Med Mol Imaging* 2013; 40: 1475 – 1477
- 74 *Ronald Boellaard P, Nanda C. Krak MD, Otto S. Hoekstra P et al.* Effects of Noise, Image Resolution, and ROI Effects of Noise, Image Resolution, and ROI Definition on the Accuracy of Standard Uptake Values: A Simulation Study. *J Nucl Med* 2004 2004: 1519 – 1527
- 75 *Sasanelli M, Meignan M, Haioun C et al.* Pretherapy metabolic tumour volume is an independent predictor of outcome in patients with diffuse large B-cell lymphoma. *Eur J Nucl Med Mol Imaging* 2014; 41: 2017 – 2022
- 76 *El-Galaly TC, Villa D, Alzahrani M et al.* Outcome prediction by extranodal involvement, IPI, R-IPI, and NCCN-IPI in the PET/CT and rituximab era: A Danish-Canadian study of 443 patients with diffuse-large B-cell lymphoma. *American journal of hematology* 2015; 90: 1041 – 1046
- 77 *Jelicic J, Stauffer Larsen T, Bukumiric Z et al.* The clinical applicability of current prognostic models in follicular lymphoma: A systematic review. *Critical reviews in oncology/hematology* 2021; 164: 103418
- 78 *Cottreau A-S, Lanic H, Mareschal S et al.* Molecular Profile and FDG-PET/CT Total Metabolic Tumor Volume Improve Risk Classification at Diagnosis for Patients with Diffuse Large B-Cell Lymphoma. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2016; 22: 3801 – 3809
- 79 *Boellaard R.* Standards for PET image acquisition and quantitative data analysis. *J Nucl Med* 2009; 50 Suppl 1: 11S-20S
- 80 *Keramida G, Peters AM.* FDG PET/CT of the non-malignant liver in an increasingly obese world population. *Clinical physiology and functional imaging* 2020; 40: 304 – 319
- 81 *Keramida G, Roldao Pereira L, Kaya G et al.* Hepatic and splenic 18 F-FDG blood clearance rates (Ki) in hepatic steatosis and diabetes mellitus. *Clinical physiology and functional imaging* 2020; 40: 99 – 105

- 82 *Reinholdt KB, Dias AH, Hoff CM et al.* Incidental FDG-Avid Focuses in Palatine Tonsils on PET/CT. *The Laryngoscope* 2022; 132: 2370 – 2378
- 83 *Kirkaite A, Hanschke SW, Truß MC, Lorenzen J.* Non-Hodgkin-Lymphom der Prostata: German Medical Science GMS Publishing House; 2019
- 84 *Jadvar H.* Is There Use for FDG-PET in Prostate Cancer? *Seminars in Nuclear Medicine* 2016; 46: 502 – 506
- 85 *Gundrum JD, Mathiason MA, Moore DB et al.* Primary testicular diffuse large B-cell lymphoma: a population-based study on the incidence, natural history, and survival comparison with primary nodal counterpart before and after the introduction of rituximab. *JCO* 2009; 27: 5227 – 5232
- 86 *Harper L, Simmons C, Nathan M et al.* Establishing Normal Testicular 18F-FDG PET/CT SUVs. *AJR. American journal of roentgenology* 2021; 216: 1634 – 1640
- 87 *Fueger BJ, Yeom K, Czernin J et al.* Comparison of CT, PET, and PET/CT for staging of patients with indolent non-Hodgkin's lymphoma. *Molecular imaging and biology* 2009; 11: 269 – 274
- 88 *Rosenbaum SJ, Lind T, Antoch G et al.* False-positive FDG PET uptake--the role of PET/CT. *Eur Radiol* 2006; 16: 1054 – 1065
- 89 *Morita Y, Nogami M, Sakaguchi K et al.* Enhanced Release of Glucose Into the Intraluminal Space of the Intestine Associated With Metformin Treatment as Revealed by 18FFluorodeoxyglucose PET-MRI. *Dia Care* 2020; 43: 1796 – 1802
- 90 *Tan GJS, Berlangieri SU, Lee ST et al.* FDG PET/CT in the liver: lesions mimicking malignancies. *Abdom Imaging* 2014; 39: 187 – 195

8. Abkürzungsverzeichnis

aaIPI	Age adjusted International Prognostic Index
ALT	Alanin-Aminotransferase
AST	Aspartat-Aminotransferase
BCL	B-Zell-Lymphom
CAR	Chimärer Antigenrezeptor
CD	Cluster of Differentiation
CHOP	Cyclophosphamid, Hydroxydaunorubicin, Oncovin (Vincristin), Prednison
CLL	Chronisch lymphatische Leukämie
CRF	Case Report File
CT	Computertomographie
DICOM	Digital Imaging and Communications in Medicine
DLBCL	Diffuses, Großzelliges B-Zell-Lymphom
DSHNHL	Deutsche Studiengruppe Hochmaligne Non-Hodgkin-Lymphome
EBV	Epstein-Barr-Virus
ECOG	Eastern Cooperative Oncology Group
EF	Ejektionsfraktion
FDG	F18-Fluorodeoxyglucose
FeV1	Forciertes Expiratorisches Volumen in 1 s
GvH-D	Graft-versus-Host-Disease
HL	Hodgkin-Lymphom
IMPI	Internationaler metabolischer prognostischer Index
IPI	Internationaler prognostischer Index
IQR	Interquartilsabstand
MTV	Metabolisches Tumervolumen

NHL	Non-Hodgkin-Lymphom
NYHA	New York Heart Association
OS	Gesamtüberleben
PERCIST	PET Response Criteria in Solid Tumors
PET	Positronenemissionstomographie
PFS	Progressionsfreies Überleben
R-CHOP	Rituximab + CHOP
R-IPi	Revised International Prognostic Index
RKI	Robert Koch Institut
ROI	Region of Interest
SCT	Stammzelltransplantation
SUV	Standardized Uptake Value
TLG	Total Lesion Glycolysis
TMTV	Total Metabolic Tumor Volume
VOI	Volume of Interest

9. Originaldokument des CRF

CRF-Bogen, welcher standardmäßig begleitend zur Auswertung verwendet wurde.

MTV CRF for Trial:	☐ OPTIMAL > 60 ☐ Other: _____
Patient ID	☐ DE- _____ - 2009-1. ☐ Other: _____
Date of PET Scan	_____.20_____. Note: 1 time point = 1 CRFs (also for addit. region)
Scanner Vendor and Model	☐ Siemens ☐ GE ☐ Philips Model: _____
Patient data	Weight _____ kg Height _____ cm Glucose _____ mg/dl
Applied Activity and scan delay p.i.	_____. MBq. _____ min p.i.
Tracer: ☐ FDG ☐ FET ☐ Ga-68-PSMA	☐ F-18-AIF-PSMA-11 ☐ Other: _____
PET reconstruction	☐ PSF ☐ TOF ☐ _____ Iterations ☐ Subsets _____ ☐ _____ ☐ post recon Gauss filter _____ mm
Primary PET image properties	Dim.: ☐ 128 ² ☐ 168 ² ☐ 200 ² ☐ _____ ² Pixel Size: _____ mm Slice thickness: _____ mm
Transformed PET image properties ☐ No transformation used here	Dim.: ☐ 128 ² ☐ 168 ² ☐ 200 ² ☐ _____ ² ☐ Conv. FWHM _____ mm Pixel Size: _____ mm Slice thickness: _____ mm
CT image properties Recon-Kernel: ☐ AC ☐ Soft ☐ Hard ☐ no CT available ☐ not needed	Reconstruction: ☐ FBP ☐ Iterative ☐ iMAR or similar Dim.: ☐ 512 ² ☐ _____ x _____ Slice thickness: _____ mm Pixel Size: _____ mm Slice increment: _____ mm
MR image properties ☐ no MRI available ☐ not needed	Sequence: ☐ T1w ☐ ceT1w ☐ ceMPRAGE ☐ T2w ☐ FLAIR Dim.: ☐ 512 ² ☐ _____ x _____ ☐ Other: _____ Pixel Size: _____ mm Slice thickness: _____ mm
Blood Pool (BP) Reference ROI ☐ Blood Pool not needed ☐ Spill-over (tumor->BP) ☐ no Sp-o	V _____ ml ☐ 0.157 ml ☐ thread ☐ sphere ☐ block ☐ cylinder in ☐ Aorta ☐ Other: _____ defined ☐ manually ☐ autom. ☐ autom., but corrected
Blood Pool SUV	Min _____ Mean _____ $\pm$ SD _____ Max _____
SUR measurements possible	☐ Yes ☐ No, ☐ impossible, separate scans of BP & tumor
Reference organ ROI Organ: ☐ Liver ☐ Normal brain ☐ Other:	V _____ ml ☐ 14.1 ml ☐ sphere ☐ block ☐ indiv. contour defined ☐ manually ☐ autom. ☐ autom., but corrected Spill-over from tumor to reference region: ☐ Yes ☐ No
Reference organ SLIV ☐ not done	Min _____ Mean _____ $\pm$ SD _____ Max _____
Reference organ SUR ☐ not done	Min _____ Mean _____ $\pm$ SD _____ Max _____
Liver derived parameters possible	☐ no need ☐ yes ☐ not, due to ☐ liver involved ☐ _____
Analysis tool used Version:	☐ syngo.via ☐ ROVER ☐ LifeX ☐ PARS ☐ Other: _____ ☐ HERMES TumorFinder ☐ MetabTool ☐ FIJI Beth Israel
Delineation method Used units: ☐ SUV ☐ SUR ☐ qPET ☐ Bq/cm ☐ Cnts ☐ Other: _____ CT-selection: ☐ not used ☐ HU Min _____ ☐ no ☐ HU Max _____ ☐ no Time for eval. (e.g. taken by screenshots) Start: ____:____:____ End: ____:____:____ Seconds: _____ ☐ not measured	Selection: ☐ completely manual ☐ masked manually ☐ absol. cut-off ☐ 2.5 ☐ 4.0 ☐ Other: _____ ☐ SUV ☐ > BP ☐ PERCIST: $SUV_{peak} \geq 2 * Mean(BP) + 2 * SD(BP) =$ _____ ☐ PERCIST: $SUV_{peak} \geq 1.5 * Mean(liver) + 2 * SD(liver) =$ _____ ☐ qPSMA: $\geq 4.3 / Mean(liver) * [Mean(liver) + SD(liver)] =$ _____ Contouring: ☐ manual ☐ 41% ☐ 50% ☐ _____ % of local max. ☐ absolute: Min: _____ ☐ 0.0 Max: ☐ _____ ☐ No max. ☐ Ratio to reference organ $\geq 1.6 (= SUV_{peak} \geq 1.6)$ ☐ Adaptive threshold (39% Max, FWHM 8 mm)
Minimum accepted ROI-volume	☐ All ☐ ≥ 0.1 ml ☐ ≥ 1 ml ☐ ≥ 2 ml ☐ Other: $\geq$ _____ ml
Number of ROIs finally accepted MTV is over- (+)(+)(+)(-) under-estimated	# of ROIs: _____ All manifestations scanned: ☐ Yes ☐ No ☐ Algorithm failed ☐ Tumor removed ☐ No tumor present
Total MTV and TLG	MTV _____ ml, TLG _____ ☐ TLG to be calc.
SUV in Total Tumor	Min _____ Mean _____ $\pm$ SD _____ Max _____
SUR in Total Tumor ☐ not done	Min _____ Mean _____ $\pm$ SD _____ Max _____
Comment:	_____

Date: _____.20_____. Reader: (printed letters) _____ Signature: _____

NUK-UKR CRF MTV V1.0f DH 2020-08-17

10. Danksagung

Mein besonderer Dank gilt Prof. Dirk Hellwig, der mich während meiner Doktorarbeit mit fachlicher Expertise stetig unterstützt hat und ein Bearbeiten dieses Themas ermöglicht hat. Ohne seine Anleitung und sein Engagement, anstehende Probleme in der Organisation bzw. in der IT schnellstmöglich zu lösen, wäre eine Abgabe in diesem Zeitfenster nicht möglich gewesen.

Ein weiterer großer Dank gebührt meinen Eltern, die mir mit ihrer unermüdlichen Unterstützung und ihrem Glauben an mich das Studium und diesen akademischen Weg ermöglicht haben.

Vielen Dank für alles.

11. Lebenslauf

Vor- und Zuname: Lukas Pichler

Geburtsdatum: 27.01.1998

Staatsbürgerschaft: Österreich

2004-2008: Volksschule; Kopplerstraße 68, 5321 Koppl, Bundesland Salzburg, Österreich

2008-2016: Gymnasium; BRG Salzburg, Akademiestraße 19, 5020 Salzburg
Schwerpunkt: Naturwissenschaften und Labor.

10/2016 - 6/2017: Ableistung des Zivildiensts beim Roten Kreuz Salzburg

10/2017-11/2024: Studium der Humanmedizin an der Universität Regensburg.

9/2019: Abschluss Vorklinik und erstes Staatsexamen.

ab 4/2021: Beginn der Arbeit an der Dissertation.

4/2023: Abschluss des zweiten Staatsexamens.

11/2023 – 10/2024: Absolvierung des Praktischen Jahres.

11/2024: Abschluss des dritten Staatsexamens