

Medizinische Fakultät der Martin-Luther-Universität Halle-Wittenberg

**Übereinstimmung verschiedener Messverfahren zur
mediastinalen Tumorzvolumenbestimmung beim
Hodgkin-Lymphom im Kindes- und Jugendalter**

Dissertation
zur Erlangung des akademischen Grades
Doktor der Medizin (Dr. med.)

vorgelegt
der Medizinischen Fakultät
der Martin-Luther-Universität Halle-Wittenberg

von Susanne Taufer, geb. Rülke
geboren am 14.05.1987 in Sangerhausen

Gutachter:

1. PD Dr. med. Hubert Gufler
2. Prof. Dr. med. Maciej Pech, Magdeburg
3. Apl. Prof. Dr. med. Alexey Surov, Leipzig

05.02.2019

19.02.2020

Meiner Familie.

REFERAT

Hintergrund und Zielsetzung. Ein großer Mediastinaltumor (sog. Bulk) gilt als etablierter Risikofaktor beim Hodgkin-Lymphom im Kindes- und Jugendalter und kann sehr exakt mit Hilfe gängiger Segmentierungsverfahren abgeschätzt werden. Für die onkologische Therapieverlaufsevaluierung gebräuchliche Richtlinien zur Beurteilung der Tumorausdehnung wie RECIST- (Response Evaluation Criteria In Solid Tumors) und Cheson-Kriterien bzw. die Ellipsoid-Methode der EuroNet-PHL-Studie beinhalten hingegen Messungen auf Grundlage des bzw. der größten, senkrecht zueinanderstehenden Durchmesser. Ziel dieser Untersuchung ist daher der Vergleich dieser metrischen Methoden mit dem Referenzstandard Tumorummetrie hinsichtlich Übereinstimmung und Reproduzierbarkeit sowie Beurteilung der Tumorgrößenreduktion.

Methoden. CT-Daten von 202 konsekutiven Patienten der EuroNet-PHL-C1-Studie mit mediastinalem Befall zwischen 2009 und 2010 wurden evaluiert. Nach Ausschluss von 123 Patienten konnten die Tumorummolumina von 79 Patienten initial und zum Zeitpunkt des ersten Restagings dreidimensional mit Hilfe der Segmentierungsmethode und der Ellipsoid-Methode der EuroNet-PHL-Studie sowie zweidimensional auf Grundlage der Cheson-Kriterien und eindimensional auf Grundlage der RECIST-Kriterien bestimmt werden, nachdem letztere in ein Volumen transformiert wurden. Für die Übereinstimmung der Methoden erfolgte die Berechnung der mittleren prozentualen Abweichung und der RSME (root mean square error), für den Interobserver-Vergleich der mittleren relativen Interobserver-Differenz und der Variationskoeffizienten. Bland-Altman-Plots dienten der graphischen Darstellung der Zusammenhänge.

Ergebnisse. Hinsichtlich der Übereinstimmung der Methoden bei der Bestimmung des Tumorummolumens wich die Ellipsoid-Methode im Mittel um ca. -17 %, die transformierte Cheson-Methode um ca. 24 % und die transformierte RECIST-Methode um ca. 1200 % von der Referenzmethode ab. Die Tumorgrößenreduktion wurde von allen Methoden im Durchschnitt nur leicht überschätzt (Ellipsoid: 4,5 %, Cheson: 5,7 %, RECIST: 3 %). Hinsichtlich der Interobserver-Variabilität wies die Segmentierungs-Methode die beste Reproduzierbarkeit auf (VK: 9,4 %), gefolgt von der Ellipsoid- (VK: 16,7 %), der Cheson- (VK: 17,8 %) und der RECIST-Methode (VK: 18,4 %).

Schlussfolgerung. Die manuelle Segmentierungsmethode ist am genauesten und am besten reproduzierbar, jedoch aufgrund des hohen Zeitaufwands und der hohen Anforderungen an die Bildqualität für den Einsatz in großen multizentrischen Studien zum Hodgkin-Lymphom nicht geeignet. Eine akzeptable Alternative für die Bestimmung des mediastinalen Volumens stellt lediglich die Ellipsoid-Methode dar. Die RECIST-Methode ist für die Abschätzung des mediastinalen Tumorummolumens beim Hodgkin-Lymphom ungeeignet, hinsichtlich der Bestimmung des Therapieansprechens jedoch akzeptabel. Alle Alternativmethoden weisen eine hinreichende Reproduzierbarkeit auf.

Tauer, Susanne: Übereinstimmung verschiedener Messverfahren zur mediastinalen Tumorummolumenbestimmung beim Hodgkin-Lymphom im Kindes- und Jugendalter, Halle (Saale), Univ., Med. Fak., Diss., 64 Seiten, 2018

I. INHALTSVERZEICHNIS

II. ABKÜRZUNGSVERZEICHNIS

III. ABBILDUNGSVERZEICHNIS

IV. TABELLENVERZEICHNIS

1.	EINLEITUNG.....	1
1.1	Morbus Hodgkin im Kindes- und Jugendalter	2
1.1.1	Epidemiologie und Prognose.....	2
1.1.2	Pathogenese und Pathologie	2
1.1.3	Krankheitsstadien und Prognosefaktoren	3
1.2	Methoden zur Bestimmung der TumorgroÙe	5
1.2.1	WHO-Kriterien.....	6
1.2.2	Cotswolds-Kriterien.....	6
1.2.3	RECIST-Kriterien.....	7
1.2.4	International-Working-Group-/Cheson-Kriterien.....	8
1.2.5	Dreidimensionale Segmentierungsverfahren	8
1.3	Bestimmung der mediastinalen Tumormasse beim Hodgkin-Lymphom.....	9
1.4	EuroNet-PHL-C1-Studie.....	10
1.4.1	Studienziel und Teilnehmer	10
1.4.2	Begutachtung der diagnostischen Ergebnisse.....	11
1.4.3	Diagnostik.....	11
1.4.4	Therapiegruppen und Therapieplan.....	14
2.	ZIELSTELLUNG	16
3.	MATERIAL UND METHODIK.....	17
3.1	Patientenkollektiv	17
3.1.1	Einschlusskriterien.....	17
3.1.2	Mediastinaler Tumorbefall	20
3.2	Bestimmung der TumorgroÙe	20
3.2.1	Dreidimensionale TumorgroÙenbestimmung mit Segmentierungssoftware.....	20
3.2.2	Dreidimensionale TumorgroÙenbestimmung nach Ellipsoid-Methode.....	22
3.2.3	Zweidimensionale TumorgroÙenbestimmung nach WHO-/Cotswolds- bzw. Cheson-Kriterien.....	23
3.2.4	Eindimensionale TumorgroÙenbestimmung nach RECIST-Kriterien.....	24
3.3	Umwandlung in Volumina	24
3.4	Beurteilung des Therapieansprechens (Response)	25

3.5	Reproduzierbarkeit der Messungen	25
3.6	Statistik	26
4.	ERGEBNISSE.....	28
4.1	Patientencharakteristika.....	28
4.2	Kenngößen der gemessenen Tumervolumina	29
4.2.1	Anzahl der Messungen	29
4.2.2	Tumervolumen zum Zeitpunkt des Stagings	30
4.2.3	Tumervolumen zum Zeitpunkt des Restagings	30
4.2.4	Tumorgrößenreduktion	32
4.3	Übereinstimmung der Methoden hinsichtlich des Tumervolumens	32
4.3.1	Übereinstimmung Segmentierungs- und Ellipsoid-Methode	32
4.3.2	Übereinstimmung Segmentierungs- und Cheson-Methode	33
4.3.3	Übereinstimmung Segmentierungs- und RECIST-Methode	33
4.4	Übereinstimmung hinsichtlich der Bewertung des Therapieansprechens	35
4.5	Interobservervariabilität.....	36
4.5.1	Kenngößen der Volumenmessungen des 2. Untersuchers	36
4.5.2	Interobservervariabilität der Messungen	37
4.6	Zeitmessungen für die OsiriX-Volumetrie	39
5.	DISKUSSION.....	40
5.1	Übereinstimmung der Methoden hinsichtlich des Tumervolumens und des Therapieansprechens.....	40
5.2	Abweichungen zwischen Segmentierungs- und Ellipsoid-Methode	44
5.3	Reproduzierbarkeit der Methoden	45
5.4	Der Einsatz der Tumervolumetrie in Therapieoptimierungsstudien	47
5.5	Ausblick: Die Verlaufsbeurteilung des Hodgkin-Lymphoms im Kindesalter	52
6.	ZUSAMMENFASSUNG	53
7.	LITERATURVERZEICHNIS.....	55
8.	THESEN	64
V.	SELBSTSTÄNDIGKEITSERKLÄRUNG	
VI.	ERKLÄRUNG ÜBER FRÜHERE PROMOTIONSVERSUCHE	
VII.	DANKSAGUNG	

II. ABKÜRZUNGSVERZEICHNIS

1D	Eindimensional
2D	Zweidimensional
3D	Dreidimensional
A	Area (Flächeninhalt)
A.	Arteria (Arterie)
Abb.	Abbildung(en)
BSG	Blutsenkungsgeschwindigkeit
BWK	Brustwirbelkörper
bzw.	beziehungsweise
c	center (hier: Fenstermitte)
ca.	circa
CI _{95 %}	95 %-Konfidenzintervall
cm	Zentimeter
cm ²	Quadratcentimeter
cm ³	Kubikcentimeter
COPDAC	Chemotherapeutika: Cyclophosphamid, Vincristin, Prednison, Dacarbazin
COPP	Chemotherapeutika: Cyclophosphamid, Vincristin, Prednison, Pro-carbazin
CR	Complete Response/Remission (Komplette Remission)
CRu	Complete Response/Remission unconfirmed (Nicht bestätigte Komplette Remission)
CT	Computertomographie
d	Durchmesser
d. h.	das heißt
DICOM	Digital Imaging and Communications in Medicine
E	extranodal (außerhalb der Lymphknoten)
EBV	Epstein-Barr-Virus
EuroNet-PHL-C1/-C2	European Network on Pediatric Hodgkin's Lymphoma - Classical Hodgkin-1/-2
et al.	et alii (und andere)

FDG	Fluordesoxyglukose
GPOH-HD	Gesellschaft für Pädiatrische Onkologie und Hämatologie-Hodgkin's Disease (Hodgkin-Studiengruppe der Gesellschaft für Pädiatrische Onkologie und Hämatologie)
GR	Größenreduktion
HE	Hounsfield-Einheiten
HIV	Humanes Immundefizienz-Virus
IWG	International Working Group (Internationale Arbeitsgruppe, hier: für Erarbeitung von Richtlinien für das Therapieansprechen von Lymphomen)
KM	Kontrastmittel
L- und H-Zellen	lymphozytäre und histiozytäre Zellen
LR	Lymphknotenregionen
M _A	Messergebnis Alternativmethode
M _R	Messergebnis Referenzmethode
MAC OS X	Macintosh Operating System 10
ml	Milliliter
mm	Millimeter
MRT	Magnetresonanztomographie
NC	No Change (Keine Veränderung $\hat{=}$ Stabile Erkrankung)
OEPA	Chemotherapeutika: Vincristin, Etoposid, Prednison, Adriamycin
p.a.-Projektion	posterior-anteriore Projektion
PACS	Picture Archiving and Communication System
PD	Progressive Disease (Krankheitsprogression)
PET	Positronen-Emissions-Tomographie
PPLLSG	Polish Pediatric Leukemia/Lymphoma Study Group (Polnische Studiengruppe für pädiatrische Hämatologie)
PR	Partial Response/Remission (Partielle Remission)
PRO	Progression
r	Radius
R	Randomisierung
RD	Relative Differenz
RECIL	Responde Evaluation Criteria In Lymphoma

RECIST	Response Evaluation Criteria In Solid Tumors
RID	Relative Interobserver-Differenz
RMSE	root mean square error (Wurzel des mittleren quadratischen Fehlers)
ROI	region of interest
RT	Radiotherapie
SD	Stable Disease (Stabile Erkrankung)
SEOP	Sociedad Española de Oncología pediátricas (Spanische Gesellschaft für pädiatrische Onkologie)
SFCE	Société Française de Lutte contre les Cancers et Leucémies de l'Enfant et de l'Adolescent (Französische Gesellschaft für Onkologie und Hämatologie bei Kindern und Jugendlichen)
sog.	sogenannt
SPD	sum of product diameters (Summe der Produkte der orthogonalen Durchmesser)
UKCCSG	United Kingdom Children's Cancer Study Group (Kinderkrebs-Studiengruppe des Vereinigten Königreichs)
US1	Untersucher 1
US2	Untersucher 2
u. v. m.	und vieles mehr
Tab.	Tabelle
TG	Therapiegruppe
V	Volumen
v. a.	vor allem
VK	Variationskoeffizient
w	width (hier: Fensterbreite)
WHO	World Health Organization
z. B.	zum Beispiel
ZNS	Zentrales Nervensystem

III. ABBILDUNGSVERZEICHNIS

Abb. 1:	Therapieplan der EuroNET-PHL-C1-Studie	15
Abb. 2:	Kontrastmittelanreicherung	18
Abb. 3:	Pulsationsartefakte.....	19
Abb. 4:	Benutzeroberfläche der Bildbearbeitungssoftware OsiriX.....	20
Abb. 5:	Manuelle Festlegung der Tumorgrenzen mit OsiriX.....	21
Abb. 6:	3D-Darstellung einer mediastinalen Raumforderung	22
Abb. 7:	Bestimmung der Tumordurchmesser in der axialen Ebene.....	23
Abb. 8:	Verteilung der Tumorumina	31
Abb. 9:	Absolute Übereinstimmung der Tumorumina.....	34
Abb. 10:	Mittlere prozentuale Abweichung	34
Abb. 11:	Abweichungen der Responseergebnisse aller Patienten.....	35
Abb. 12:	Abweichungen der Responseergebnisse der Patienten mit einer Tumoruminreduktion <80 %	36
Abb. 13:	Absolute Interobserver-Differenzen	38
Abb. 14:	Relative Interobserver-Differenzen	39

IV. TABELLENVERZEICHNIS

Tab. 1:	WHO-Einteilung des Morbus Hodgkin nach Histologie.....	3
Tab. 2:	Cotswolds-Modifikation der Ann-Arbor-Klassifikation	4
Tab. 3:	Bildgebende Diagnostik im Rahmen der EuroNet-PHL-C1-Studie.....	12
Tab. 4:	Lokale Remissionskategorien der EuroNet-PHL-C1-Studie.....	13
Tab. 5:	Therapiegruppen.....	14
Tab. 6:	Kriterien zur Beurteilung der Bildqualität.....	19
Tab. 7:	Ausschlusskriterien und Anzahl der vom Ausschluss betroffenen Patienten	28
Tab. 8:	Patientencharakteristika.....	29
Tab. 9:	Kenngößen der jeweiligen Tumorumfänge.....	31
Tab. 10:	Kenngößen der Tumorumfangreduktion (Response)	32
Tab. 11:	Kenngößen der Übereinstimmung bei der Bestimmung des Tumorumfanges	33
Tab. 12:	Kenngößen der Interobserver-Untersuchungen.....	38

1. EINLEITUNG

Zur Behandlung maligner Tumorerkrankungen stehen dem Onkologen heutzutage multimodale Therapiekonzepte zur Verfügung, welche operative Verfahren, medikamentöse Therapien insbesondere mit Zytostatika und Antikörpern sowie die Strahlentherapie einschließen. Um die Wirkung bzw. den Erfolg eines onkologischen Behandlungskonzepts sinnvoll und objektiv einzuschätzen, werden in Studien neben verschiedenen Parametern wie der Überlebenszeit (Gesamtüberleben bzw. progressionsfreies Überleben) oder patientenrelevanten Daten (z. B. Lebensqualität) auch Tumorparameter herangezogen. Dazu wiederum gehören die mit Hilfe nuklearmedizinischer Verfahren (v. a. PET) bestimmte Stoffwechselaktivität, laborchemisch nachgewiesene Tumormarker, aber auch die mittels bildgebender Verfahren messbare Tumorgrößen- ausdehnung, welche sowohl beim Hodgkin- als auch bei Non-Hodgkin-Lymphomen eine prognostische Bedeutung haben (Cheson et al., 2014). Mit Hilfe der Bildgebung ist es nicht nur möglich, die exakte Ausdehnung des Primärtumors zu berechnen, sondern auch die Größe und Anzahl von Lymphknoten- sowie Fernmetastasen zu bestimmen. Insbesondere in Phase-II-Studien zur Überprüfung der Wirksamkeit eines neuen Medikaments, jedoch auch in Therapieoptimierungsstudien, dient die Bestimmung der Tumorgröße als Grundlage für die Beurteilung der Effektivität einer Therapie. Für das Hodgkin-Lymphom spielt die Tumorgrößenbestimmung mittels bildgebender Verfahren zusätzlich insofern eine Rolle, als dass das Vorliegen einer „bulky disease“ bzw. eines großen Mediastinaltumors einen anerkannten Risikofaktor darstellt (Mauch et al., 1978; Lee et al., 1980; Willett et al., 1988; Barrington et al., 2014), der mit einer Prognoseverschlechterung einhergeht, sodass Patienten entsprechend intensiver behandelt werden müssen.

Verschiedene Richtlinien, die den Untersuchern objektive Kriterien zur Bewertung einer Tumorthherapie liefern und die Vergleichbarkeit klinischer Studien ermöglichen sollen, wurden seit einigen Jahrzehnten etabliert. Dazu gehören beispielsweise die WHO-Kriterien (World Health Organization, 1979) und die RECIST-Kriterien (Therasse et al., 2000; Eisenhauer et al., 2009), welche zuletzt auch für Lymphome modifiziert wurden (Younes et al., 2017). Für das Hodgkin-Lymphom im Speziellen führte man mit der Cotswolds-Modifikation der Ann-Arbor-Klassifikation auch erstmals Kriterien ein, die Empfehlungen für die Beurteilung des Therapieansprechens beinhalteten (Lister et al., 1989). Eine Erweiterung und Verbesserung dieser stellen die Therapieansprechkriterien für maligne Lymphome einer internationalen Arbeitsgruppe aus Onkologen, Radiologen und Pathologen dar, die sogenannten International-Working-Group (IWG)- bzw. nach dem Erstautor auch Cheson-Kriterien (Cheson et al., 1999; 2007; 2014).

All diese Richtlinien wurden bisher insbesondere in klinischen Studien mit erwachsenen Probanden getestet. Für Kinder und Jugendliche existieren keine speziellen Empfehlungen bezüglich der Aus- und Bewertung klinischer Studien. Vielmehr sind diese abhängig vom jeweiligen Studienprotokoll und Therapiezentrum.

1.1 Morbus Hodgkin im Kindes- und Jugendalter

1.1.1 Epidemiologie und Prognose

Die malignen Lymphome machen in Deutschland ca. 14,3 % aller Tumorerkrankungen im Kindes- und Jugendalter bis zum 18. Lebensjahr aus und stehen damit nach den Leukämien (30,4 %) und den ZNS-Tumoren (23,6 %) bezogen auf die relative Häufigkeit an dritter Stelle. Dennoch ist das Hodgkin-Lymphom mit einer relativen Häufigkeit von 4,5 % in Bezug auf alle malignen Tumorleiden im Kindesalter eine seltene Erkrankung. Jährlich erkranken ca. 100 Kinder und Jugendliche bis einschließlich zum 14. Lebensjahr bzw. ca. 150 bis einschließlich zum 18. Lebensjahr neu am Morbus Hodgkin. Das mediane Alter bei Erkrankungsbeginn liegt bei 12 Jahren und 7 Monaten. Kleinkinder erkranken sehr selten. Die Inzidenzrate nimmt bis zum 15. Lebensjahr stetig zu, wobei Jungen insgesamt etwas häufiger betroffen sind als Mädchen. Mit einer 5-Jahres-Überlebensrate von ca. 99 % und einer 10-Jahres-Überlebensrate von ca. 98 % gehört das Hodgkin-Lymphom zu den malignen Erkrankungen mit den besten Überlebenswahrscheinlichkeiten im Kindesalter. Die Mortalitätsrate hat sich durch kontinuierliche Weiterentwicklung der Therapie ständig verringert und ist 15 Jahre nach Diagnosestellung mit 0,4 im Vergleich zu allen Tumorerkrankungen insgesamt (34,4) sehr niedrig. Jedoch treten bei ca. 13,2 % aller Hodgkin-Patienten innerhalb von 30 Jahren nach der Diagnosestellung Zweitmalignome auf, was im Vergleich zu allen anderen Malignomen (6,6 %) relativ viel ist (Kaatsch et al., 2016).

1.1.2 Pathogenese und Pathologie

Der Morbus Hodgkin gehört ebenso wie die Non-Hodgkin-Lymphome zu den malignen Lymphomen, also den bösartigen Erkrankungen des lymphatischen Systems. Erstmals wurde die Erkrankung 1832 von dem britischen Arzt und Pathologen Thomas Hodgkin beschrieben (Hodgkin, 1832). Aus bisher noch unklarer Ursache kommt es zu einer Expansion entarteter monoklonaler B-Lymphozyten vor allem in lymphatischen Geweben (Kuppers et al., 1999). Durch lymphogene und hämatogene Streuung sowie die Ausbreitung per continuitatem sind jedoch auch extralymphatische Gewebe, wie die Lunge, die Leber, die Haut, seröse Häute oder das Knochenmark betroffen. Wahrscheinlich ist, dass ein Zusammenspiel verschiedener Faktoren, wie Umweltfaktoren, genetische und ethnische Faktoren zum Ausbruch der Erkrankung führt. Als Risikofaktoren gelten virale Infektionskrankheiten, wie eine EBV- oder HIV-Infektion (Biggar et al., 2000; Swerdlow, 2003; Young und Murray, 2003), sowie verschiedene Krankheiten mit Immundefekten oder immunsuppressive Therapien (Filipovich et al., 1992; Knowles, 1999; Suarez et al., 2015).

Nach histologischen Gesichtspunkten unterteilt die WHO das Hodgkin-Lymphom in folgende Gruppen:

Tab. 1: WHO-Einteilung des Morbus Hodgkin nach Histologie
Nach (Harris et al., 2000)

Einteilung nach Histologie	
1. Klassisches Hodgkin-Lymphom	<ul style="list-style-type: none">• Nodulär-sklerosierender Typ• Gemischtzelliger Typ• Lymphozytenreicher Typ• Lymphozytenarmer Typ
2. Noduläres lymphozytenprädominantes Hodgkin-Lymphom	

Das klassische Hodgkin-Lymphom tritt mit ca. 95 % weitaus häufiger auf als das noduläre lymphozytenprädominante mit ca. 5 % (Harris et al., 2000). Im Kindes- und Jugendalter wird der lymphozytenprädominante Typ als eigenständige Krankheitsentität gewertet, sodass hierfür spezifische Therapiestudien und -pläne existieren (Mauz-Körholz et al., 2007; 2015). Die malignen Zellen beim klassischen Typ des Morbus Hodgkin sind die mehrkernigen und nach ihren Entdeckern benannten Reed-Sternberg-Riesenzellen, welche aus den einkernigen Hodgkin-Zellen (entartete B-Lymphozyten) fusionieren. Für den lymphozytenprädominanten Typ sind die sogenannten L- und H-Zellen (lymphozytäre und histiozytäre Zellen) typisch.

1.1.3 Krankheitsstadien und Prognosefaktoren

Wie die meisten malignen Tumorentitäten galt auch der Morbus Hodgkin zunächst als unheilbare Erkrankung. Seit der Mitte des 20. Jahrhunderts jedoch konnte insbesondere mit der Anwendung der Strahlentherapie die Prognose der Patienten stetig verbessert werden (Kaplan und Rosenberg, 1966). Bereits zu diesem Zeitpunkt erfolgte eine von der Tumorausdehnung abhängige Therapieplanung. Während der kurative Therapieansatz zunächst vorwiegend die Bestrahlung umfasste und Chemotherapeutika nur zur palliativen Behandlung eingesetzt wurden, existieren heutzutage stadienorientierte Therapiekonzepte mit beiden Methoden, die dazu führen, dass die 5-Jahres-Überlebensrate bei Erwachsenen etwa 85 % und bei Kindern etwa 99 % beträgt (Kaatsch et al., 2016; Robert Koch-Institut und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V., 2017).

Ein wichtiges Kriterium für die Therapieplanung ist daher die Ausbreitung des Tumors im Körper, welche sich in der Einteilung der Krankheitsstadien widerspiegelt. Diese Stadieneinteilung

erfolgt heutzutage nach der Ann-Arbor-Klassifikation bzw. nach der Cotswolds-Modifikation der Ann-Arbor-Klassifikation:

Tab. 2: Cotswolds-Modifikation der Ann-Arbor-Klassifikation
 Modifiziert nach (Lister et al., 1989)

Stadium	Definition
I	Befall einer Lymphknotenregion oder: Vorliegen eines lokalisierten extranodalen Herdes (IE)
II	Befall von zwei oder mehr Lymphknotenregionen auf einer Seite des Zwerchfells oder: Vorliegen lokalisierter extranodaler Herde mit Befall einer oder mehrerer Lymphknotenregionen auf einer Seite des Zwerchfells (IIE)
III	Befall von zwei oder mehr Lymphknotenregionen beidseits des Zwerchfells oder: Befall lokalisierter extranodaler Herde und Lymphknoten beidseits des Zwerchfells (IIIE)
IV	Disseminierter Befall eines oder mehrerer extralymphatischer Organe mit oder ohne Lymphknotenbefall

E = Extranodal

Zum lymphatischen System werden Lymphknoten, Milz, Thymus, Waldeyer-Rachenring, Blinddarm und Peyer-Plaques gezählt.

Neben der Tumorausdehnung existieren diverse Faktoren, deren prognostische Wertigkeit für das Hodgkin-Lymphom untersucht wurden bzw. belegt sind. Dazu gehören Alter, Geschlecht, das Vorliegen von B-Symptomen, die Blutsenkungsgeschwindigkeit, die Zahl der weißen Blutzellen, der Serumalbumin- und Hämoglobinspiegel usw. (Hasenclever und Diehl, 1998; Cuccaro et al., 2014).

Im Rahmen von klinischen Studien der Gesellschaft für Pädiatrische Onkologie und Hämatologie (GPOH) spielen für die Zuordnung der Patienten zu den Therapiegruppen jedoch v. a. folgende Faktoren eine Rolle: der Extranodalbefall, das Vorliegen von B-Symptomen, die Blutsenkungsgeschwindigkeit und der Nachweis einer großen zusammenhängenden Tumormasse (Yiallourou, 2016).

Ein extranodaler Befall liegt dann vor, wenn sich die Erkrankung von einem Lymphknoten bzw. einer Lymphknotenregion in das unmittelbar angrenzende Gewebe oder Organ ausbreitet (per continuitatem, z. B. in die Pleura, das Pericard oder die Lunge), und betrifft nur die Stadien I bis III. Ein Befall von Organen, die sich nicht in unmittelbarer Nachbarschaft zu einer

Lymphknotenregion befinden, wird als Stadium IV gewertet, da die Ausbreitung in diesem Fall hämatogen erfolgt. Ein Leberbefall wird dem Stadium IV zugeordnet (Yiallouros, 2016).

Neben dem Extranodalbefall wird auch das Vorhandensein von B-Symptomen als Risikofaktor gewertet. Dazu gehören: ein unklarer Gewichtsverlust von mehr als 10 % innerhalb der letzten 6 Monate und/oder andauerndes oder wiederholt auftretendes Fieber (über 38°C) unklarer Ursache und/oder starker Nachtschweiß. Beim Auftreten von B-Symptomen wird das Krankheitsstadium zusätzlich durch den Buchstaben „B“ gekennzeichnet, beim Fehlen dieser Symptome zusätzlich durch den Buchstaben „A“ (Lister et al., 1989).

Die Blutsenkungsgeschwindigkeit (BSG) kann bei vielen entzündlichen Erkrankungen, aber auch bei Tumorerkrankungen erhöht sein. Beim Morbus Hodgkin gilt ein BSG-Wert ab 30 mm als Risikofaktor (Yiallouros, 2016).

Auch das Vorliegen einer großen zusammenhängenden Tumormasse („bulky disease“) stellt aufgrund der höheren Rezidivwahrscheinlichkeit einen international anerkannten Risikofaktor dar, der eine intensiviertere Therapie zur Folge hat (Mauch et al., 1978; Lee et al., 1980; Barrington et al., 2014). Eine „bulky disease“ außerhalb des Thorax lag nach den Cotswolds-Kriterien vor, wenn die größte zusammenhängende Tumormasse eine Ausdehnung von mehr als 10 cm aufwies. Ein mediastinales Lymphknotenkonglomerat wurde zudem als Bulk bezeichnet, wenn die maximale Weite des Mediastinums in einer p.a.-Projektion des Röntgen-Thorax mindestens 1/3 des Gesamthoraxdurchmessers betrug (gemessen auf Höhe BWK 5/6). Dem Krankheitsstadium wurde in diesem Fall der Buchstabe „X“ angehängt (Lister et al., 1989). Im aktuellen Konsensusbericht zum Staging und Response-Assessment von Lymphomen (Lugano-Klassifikation) wird die prognostische Bedeutung einer großen mediastinalen Tumormasse betont, gleichzeitig jedoch empfohlen, den größten Tumordurchmesser CT-gestützt zu bestimmen, was die zusätzliche Anfertigung von Röntgen-Aufnahmen überflüssig macht. Auch die Kennzeichnung mit dem Buchstaben „X“ ist nicht länger notwendig (Cheson et al., 2014).

1.2 Methoden zur Bestimmung der Tumorgöße

Schon seit den 50er bzw. 60er Jahren des 20. Jahrhunderts wurde versucht, objektive Kriterien für die Therapie bösartiger Erkrankungen zu definieren, die das Ansprechen auf eine systemisch-onkologische Therapie beurteilen. Nach *Zubrod et al.* galt das Ansprechen auf eine Therapie als positiv, wenn sich entweder die gesamte messbare Tumormasse verringert hatte, keine Läsion an Größe zugenommen hatte und keine neuen Läsionen entstanden waren oder die behandelnden Ärzte unter Berücksichtigung von subjektiven Ansprechkriterien und unerwünschten Wirkungen zu dem Schluss kamen, dass die Behandlung für den Patienten von Nutzen war (Zubrod et al., 1960).

Mit der Weiterentwicklung der bildgebenden Verfahren einhergehend haben sich insbesondere die radiologisch-morphologischen Veränderungen der Tumormasse als objektive Kriterien für die Beurteilbarkeit des Therapieansprechens etabliert.

1.2.1 WHO-Kriterien

In dem Versuch, die bis dahin relativ unübersichtliche und uneinheitliche Literatur zur Bewertung von Krebstherapien zu vereinfachen und damit auch für verschiedene Untersucher vergleichbar zu machen, verfasste die WHO 1979 ein Handbuch zur Auswertung von Ergebnissen von Tumortherapien (World Health Organization, 1979). Danach wurde von allen messbaren Tumorkläsionen annäherungsweise eine Fläche bestimmt, indem man das Produkt aus dem größten Längs- und dem größten Querdurchmesser der Läsion bildete. Beim Vorliegen mehrerer Läsionen berechnete man die Summe aller Produkte der größten orthogonalen Durchmesser der Läsionen („sum of product diameters“ - SPD). Die WHO-Methode war insbesondere für die konventionelle Röntgentechnik, aber auch für die frühen Computertomographen ausgelegt. Nach der Therapie erfolgte eine erneute Bestimmung der Größe der residualen Läsionen. Das Gesamtansprechen aller messbaren Tumorkläsionen wurde durch diejenige Läsion festgelegt, die das schlechteste Therapieansprechen aufwies.

1.2.2 Cotswolds-Kriterien

1988 fand in Cotswolds, England, eine Konferenz statt, die zum Ziel hatte, die 1971 eingeführte Ann-Arbor-Klassifikation für das Hodgkin-Lymphom in Anbetracht neuer technologischer Entwicklungen und bereits gewonnener Erfahrungen in der Anwendung dieser Klassifikation zu überarbeiten. Die wesentlichen Modifikationen beinhalteten die Einführung der CT-Untersuchung als bildgebendes Verfahren im Rahmen des Stagings, die Einführung der Definition Bulk als einen prognostischen Faktor sowie die Erweiterung um Kriterien, die das Therapieansprechen bzw. den Remissionsstatus beschreiben (Lister et al., 1989).

Die Erarbeitung dieser Empfehlungen erfolgte in Anlehnung an die 1979 eingeführten WHO-Kriterien, d. h. für die Verlaufsbeurteilung der Lymphome sollten alle messbaren Läsionen durch die Berechnung der Summe des Produkts der größten senkrecht zueinanderstehenden Tumordurchmesser erfasst werden. Eine zusätzliche Kategorie, nämlich „Complete Remission (unconfirmed/uncertain)“, wurde für Patienten hinzugefügt, deren Remissionsstatus unklar war. Dies galt insbesondere für Fälle, in denen die Patienten trotz Fehlen klinischer Zeichen einer Erkrankung radiologische Auffälligkeiten aufwiesen, bei welchen man nicht sicher sagen konnte, ob sie persistierende Tumorreste oder Effekte der Therapie (z. B. Fibrose) darstellten.

1.2.3 RECIST-Kriterien

Obwohl die WHO-Kriterien weithin akzeptiert und angewendet wurden, ergaben sich im Laufe der Zeit einige Probleme: Die Kriterien beinhalteten weder eine Mindestgröße für die zu messenden Läsionen noch begrenzten sie deren Anzahl. Darüber hinaus variierte je nach Studienprotokoll die Grundlage der Zuordnung zur Kategorie „Progressive Disease“: Einige Untersucher bezogen sich auf die Größenzunahme lediglich einer Läsion, andere auf die gesamte Tumormasse, was die Vergleichbarkeit verschiedener Studien reduzierte (Therasse et al., 2000).

Mit dem Ziel, die beim WHO-Schema aufgetretenen Defizite zu verbessern und sich den neuen technologischen Gegebenheiten der Computer- und Magnetresonanztomographie anzupassen, erfolgte schließlich im Jahr 2000 die Einführung der RECIST-Kriterien, d. h. der Response Evaluation Criteria In Solid Tumors (Therasse et al., 2000). Eine erste Überarbeitung (RECIST 1.1) fand im Jahr 2008 statt (Eisenhauer et al., 2009). Der wesentliche Unterschied zu den WHO-Kriterien bestand darin, dass die Größe der Läsionen nun nicht länger mit Hilfe einer zweidimensionalen Methode berechnet wurde, sondern eindimensional mit Hilfe des größten Durchmessers. Die Begründung dafür ist, dass Änderungen im Durchmesser eines annähernd kugelförmigen Tumors besser mit der durch eine Chemotherapie hervorgerufenen Änderung der Zellzahl korreliert als zweidimensionale Messungen (James et al., 1999). Die RECIST-Richtlinien basieren daher auf der Annahme, die Form eines Tumors ähnele einer Kugel.

Die soliden Tumore werden in messbare und nicht messbare Läsionen unterteilt. Messbare Läsionen müssen in der CT eine Mindestgröße von 10 mm aufweisen. Läsionen, die kleiner als 10 mm messen, sowie Knochenläsionen, Aszites, Pleuraergüsse, inflammatorische Brustkrankungen oder Hautläsionen stellen nicht messbare Läsionen dar. Für die Beurteilung des Therapieansprechens legte man weiterhin fest, dass anhand des größten Durchmessers sogenannte Zielläsionen bestimmt werden, welche die Gesamttumorlast repräsentieren sollen. Man begrenzte deren Anzahl (nach den überarbeiteten Richtlinien von 2008) auf maximal fünf insgesamt und maximal zwei pro betroffenes Organ. Schließlich ist die Summe der längsten Durchmesser der einzelnen Zielläsionen zu bilden und im Verlauf zu kontrollieren. Alle anderen Läsionen stellen Nicht-Zielläsionen dar. Diese müssen nicht ausgemessen werden, sind jedoch bei der Verlaufsbeurteilung ebenfalls hinsichtlich der Größenänderung bzw. dem Nachweis neuer Läsionen zu berücksichtigen. Das Gesamtansprechen ergibt sich dann aus einem Zusammenspiel der Response-Bewertung der Zielläsionen, der Nicht-Zielläsionen und eventuell neu aufgetretener Läsionen.

In der überarbeiteten Version wurde auch die Definition von Lymphknoten als Ziel- bzw. Nicht-Zielläsionen angepasst. Demnach gelten Lymphknoten mit einem Kurzdurchmesser von weniger als 10 mm als nicht betroffen und mit einem Kurzdurchmesser von 15 mm als sicher pathologisch, sodass diese mögliche Zielläsionen darstellen. Lymphknoten mit einem Kurzdurchmesser zwischen 10 und 15 mm sind Nicht-Zielläsionen (Schwartz et al., 2009).

2017 wurden mit den sogenannten RECIL-Kriterien spezielle Empfehlungen für die Verlaufsbeurteilung maligner Lymphome veröffentlicht, die sich an den bekannten RECIST-Kriterien orientieren und insbesondere eine eindimensionale Messung des längsten Durchmessers dreier Zielläsionen (Lymphome) beinhaltet (Younes et al., 2017).

1.2.4 International-Working-Group-/Cheson-Kriterien

1999 entwickelte eine internationale Arbeitsgruppe aus Klinikern, Radiologen und Pathologen (International Working Group bzw. IWG) Richtlinien für das Therapieansprechen von Non-Hodgkin-Lymphomen bei Erwachsenen, die nach dem Erstautor der entsprechenden Publikation auch „Cheson-Kriterien“ genannt wurden (Cheson et al., 1999). Diese wurden später auch von den Hodgkin-Lymphom-Studiengruppen übernommen. Eine überarbeitete Fassung entstand im Jahr 2007 (Cheson et al., 2007). Ähnlich den WHO-Richtlinien und den Cotswolds-Kriterien schlagen die Autoren zur Größenbestimmung die Berechnung der Summe des Produkts der größten orthogonalen Durchmesser der Läsionen vor, wobei sich die Autoren auf sechs dominante Läsionen beschränkten. Diese sollten repräsentativ für die gesamte Tumormasse sein. Eine entscheidende Veränderung stellte jedoch die Einbeziehung der funktionellen Bildgebung im Sinne einer FDG-PET-Untersuchung dar, woraus eine sehr komplexe Response-Beurteilung resultierte. Eine weitere Überarbeitung der Kriterien spiegelt sich in der 2014 veröffentlichten Lugano-Klassifikation wieder, also den Empfehlungen für die Evaluierung, das Staging und die Verlaufsbeurteilung von Hodgkin- und Non-Hodgkin-Lymphomen (Cheson et al., 2014). Darin wird insbesondere die Bedeutung von FDG-PET-Untersuchungen für Staging und Restaging als Standard betont. Außerdem wird weiterhin auf die prognostische Bedeutung einer großen zusammenhängenden Tumormasse, also eines Bulk, hingewiesen.

1.2.5 Dreidimensionale Segmentierungsverfahren

Das stetige Voranschreiten der technischen Entwicklung mit Spiral- und Mehrzeilencomputertomographen, modernen Magnetresonanztomographen und nicht zuletzt auch der funktionellen Bildgebung ermöglicht es uns heutzutage einen Tumor nicht nur genauestens zu lokalisieren, sondern auch ihn als dreidimensionales Objekt darzustellen, sein Volumen berechnen zu lassen, eine virtuelle Therapieplanung durchzuführen u. v. m.

Grundlage für die moderne digitale Bildanalyse sind sogenannte Segmentierungsverfahren. Die Segmentierung ist eine der ersten Schritte im Rahmen der digitalen Bildverarbeitung. Sie dient der Zuordnung verschiedener Bildpunkte zu einem Objekt oder einer Region anhand von Homogenitätskriterien, mit dem Ziel, unterschiedliche und relevante Bildbestandteile voneinander getrennt darzustellen (Lehmann et al., 1997). Homogen bedeutet hier, die Bildelemente sind in ihren

Eigenschaften hinsichtlich eines Kriteriums ähnlich. In der Fachliteratur unterscheidet man methodisch punkt- bzw. pixelorientierte, kantenorientierten, regionenorientierte, texturorientierte und modellbasierte Segmentierungsverfahren (Handels, 2009). Grundlage der meisten in der medizinischen Bildgebung verwendeten Segmentierungsalgorithmen, auf deren Erläuterung an dieser Stelle aufgrund der Komplexität des Themas verzichtet werden soll, ist die Zuordnung der einzelnen Bildelemente aufgrund von Grauwertinformationen (Lehmann et al., 1997).

Wichtige Anwendungsgebiete solcher Segmentierungsverfahren sind z. B. die Planung von Behandlungen, wie die präoperative Planung einer Tumoresektion oder die Erstellung eines Bestrahlungsplanes sowie die Darstellung der Koronargefäße oder einzelner Ventrikel des Herzens zur Analyse der Myokardfunktion in der Kardiologie. Auch Grundlage dreidimensionaler Volumenmessverfahren sind verschiedene Segmentierungsalgorithmen. Hierbei stehen dem Untersucher grundsätzlich zwei Verfahren zur Verfügung: die manuelle und die (semi-)auto-matische Segmentierung (Gavrielides et al., 2009). Manuelle Segmentierungsalgorithmen verwenden die sogenannten Regions of Interest (ROIs), um die Grenzen des relevanten Objekts festzulegen. Dabei muss der Untersucher selbstständig und in jeder Schicht mit Hilfe von bestimmten Werkzeugen die Begrenzungspunkte des Objekts festlegen, welche dann mit Hilfe der Computersoftware verbunden werden. Für die Volumenberechnung werden die Voxel innerhalb des Objekts gezählt und anschließend mit dem Volumen eines einzelnen Voxels multipliziert (Tingelhoff et al., 2007; Fabel et al., 2008).

(Halb-)Automatische Segmentierungsverfahren wiederum verwenden verschiedene Techniken. Eine häufig angewandte Methode ist die „Growing-Region“-Funktion. Dabei muss der Untersucher das auszumessende Objekt markieren, z. B. mit einem sogenannten Saatpunkt („seed-point“). Die Computersoftware erweitert die Markierung dann mit Hilfe des Growing-Region-Algorithmus anhand von Grauwertdifferenzen zum Nachbarpixel auf das gesamte Objekt (Handels, 2009). Unzufriedenstellende Ergebnisse können im Anschluss manuell korrigiert werden.

1.3 Bestimmung der mediastinalen Tumormasse beim Hodgkin-Lymphom

Wie bereits unter Abschnitt 1.1.3 beschrieben, sind große mediastinale Lymphknotenkonglomerate im Rahmen des Hodgkin-Lymphoms von prognostischer Bedeutung, da sie einen Risikofaktor für ein Tumorrezidiv darstellen. Um das Ausmaß der mediastinalen Beteiligung zu ermitteln, wird nach den Cotswolds-Kriterien der im Röntgen-Thorax bestimmte Durchmesser des Mediastinums ins Verhältnis zum Gesamthoraxdurchmesser gesetzt, wobei ein Verhältnis von mindestens 1:3 für eine große zusammenhängende Tumormasse spricht. Patienten mit einer „bulky disease“ werden aufgrund der erhöhten Rezidivwahrscheinlichkeit intensiver therapiert.

Je nach Studiengruppe wird bzw. wurde jedoch eine Vielzahl abweichender Definitionen, insbesondere für die mediastinale Tumormasse, verwendet (Hopper et al., 1991). Mit dem technischen

Fortschritt und der Anwendung der Computertomographie sind zudem Untersuchungen angestellt worden, die prüfen sollten, ob die Bestimmung des Verhältnisses von Mediastinaldurchmesser zum Thoraxdurchmesser im Röntgen-Thorax noch zeitgemäß ist. *Kriz et al.* und *Bradley et al.* wiesen nach, dass zwischen den Modalitäten Röntgen und CT für die Bestimmung des Verhältnisses eine sehr gute Korrelation besteht (Bradley et al., 1999; Kriz et al., 2012). Für die Therapieplanung spielt jedoch auch eine Rolle, ob die volumetrische Bestimmung der „wahren“ mediastinalen Tumormasse hinsichtlich der Prognose und Rezidivwahrscheinlichkeit zusätzliche Informationen liefern kann. *Willet et al.* beispielweise untersuchten 53 Hodgkin-Patienten im Stadium IA bis IIB bezüglich der Korrelation zwischen dem Verhältnis aus Mediastinaldiameter und Thoraxdiameter und dem computertomographisch bestimmten mediastinalen Tumolvolumen (Willet et al., 1988). Obwohl eine gewisse Korrelation der Parameter bestand, konnte zusätzlich festgestellt werden, dass Patienten, die ein Verhältnis von 1:3 oder kleiner und damit im Sinne der Cotswolds-Kriterien keine „bulky disease“ aufwiesen, insbesondere dann ein Tumorrezidiv entwickelten, wenn das initial bestimmte mediastinale Tumolvolumen 200 ml übertraf. Eine Konsequenz ist, dass Patienten je nach angewandter Definition bzw. Art und Weise der Tumorgrößenbestimmung gegebenenfalls einer anderen Therapie zugeführt werden.

1.4 EuroNet-PHL-C1-Studie

1.4.1 Studienziel und Teilnehmer

Das „European Network on Pediatric Hodgkin’s Lymphoma“ (EuroNet-PHL) führte im Zeitraum von Januar 2007 bis Oktober 2012 eine multizentrische Therapieoptimierungsstudie zum klassischen Hodgkin-Lymphom im Kindes- und Jugendalter durch (EuroNet-PHL-C1-Studie). Ein wesentliches Ziel dieser Studie bestand darin, unter Einbeziehung der FDG-PET im Rahmen von Staging und Restaging zu untersuchen, ob bei Patienten mit einem adäquaten Ansprechen auf eine Bestrahlung verzichtet werden kann. Dadurch erhoffte man sich, die unerwünschten Therapieeffekte, insbesondere das Auftreten von Zweitmalignomen, zu minimieren (Körholz et al., 2006). Die teilnehmenden Studiengruppen wurden von bereits vorbestehenden nationalen und länderübergreifenden Studiengruppen gebildet. Dazu gehörte die Studiengruppe für das Hodgkin-Lymphom der Deutsche Gesellschaft für pädiatrische Onkologie und Hämatologie (GPOH-HD), welche kideronkologische Zentren aus den Ländern Deutschland, Österreich, Schweiz, Dänemark, Schweden, Norwegen, der Tschechischen Republik und der Slowakei einschließt, sowie die „Children's Cancer and Leukemia Group“ Großbritanniens (UKCCSG), welche die Länder England, Wales, Irland und Schottland umfasst. Ebenso beteiligten sich Studiengruppen aus Spanien (SEOP - Sociedad Española de Oncología pediátricas), Frankreich (SFCE - Société Française de

Lutte contre les Cancers et Leucémies de l'Enfant et de l'Adolescent) und Polen (PPLLSG - Polish Pediatric Leukemia/Lymphoma Study Group).

1.4.2 Begutachtung der diagnostischen Ergebnisse

Für die GPOH-HD, die UKCCSG und die SEOP erfolgte eine zentrale Begutachtung, d. h. die Beurteilung von Staging und Therapieansprechen wurde von einer Gruppe aus Referenzonkologen, -radiologen, -nuklearmedizinern und -strahlentherapeuten der Universitäten Halle-Wittenberg und Leipzig vorgenommen. Diese bearbeiteten die zugesandten Patienten-Daten und gaben die zusammengeführten Ergebnisse mit Therapieempfehlungen an die entsprechenden kinderonkologischen Zentren zurück. Eine eigenständige Beurteilung der Untersuchungen wurde lediglich in Frankreich und Polen durchgeführt, jedoch anhand von Empfehlungen des Studienprotokolls.

1.4.3 Diagnostik

Die Diagnosesicherung erfolgte histologisch durch die Biopsie einer betroffenen Lymphknotenregion oder anderer primär betroffener Organe. In länderspezifischen Referenz-Pathologien musste im Folgenden die Diagnose des klassischen Morbus Hodgkin bestätigt werden.

Neben der histologischen Diagnosesicherung und verschiedenen klinischen und labortechnischen Untersuchungen beinhaltete das Studienprotokoll auch bildgebende diagnostische Maßnahmen, zum einen zur Klassifizierung des Stadiums und Einordnung der entsprechenden Therapiegruppe und zum anderen zur Beurteilung des Therapieansprechens (siehe Tab. 3). Im Rahmen der Staginguntersuchungen vor Beginn der Therapie sollte in jedem Fall eine FDG-PET-Untersuchung erfolgen. Weiterhin wurde vorgegeben, eine kontrastmittelgestützte Schnittbildgebung von der Schädelbasis bis zur Symphyse durchzuführen, wobei Hals-, Abdomen- und Beckenuntersuchungen entweder mit CT oder MRT erfolgen konnten. Für den Thorax war eine CT-Untersuchung ebenso obligat wie eine sonographische Untersuchung des Abdomens. Die Anfertigung einer Röntgen-Thorax-Übersichtsaufnahme erfolgte zum Ausmessen des Verhältnisses von maximalem Durchmesser der mediastinalen Raumforderung und maximalem Thoraxinnendurchmesser auf Höhe BWK 5 bzw. 6 und damit zur Beurteilung des Ausmaßes der mediastinalen Beteiligung. Zum Zeitpunkt des ersten Restagings, also 14 Tage nach Beendigung der ersten zwei Chemotherapiezyklen, gab das Studienprotokoll eine erneute FDG-PET-Untersuchung vor. Darüber hinaus sollten alle initial betroffenen Lymphknotenregionen erneut mit Hilfe von CT oder MRT untersucht werden. Bei nachgewiesenem Lungenbefall war eine low-dose-CT-Untersuchung des Thorax obligatorisch. Nach vier Chemotherapiezyklen bei Therapiegruppe 2 bzw. nach sechs Chemotherapiezyklen bei Therapiegruppe 3 erfolgte schließlich das zweite Restaging,

ebenfalls mit Untersuchung der initial betroffenen Regionen bzw. Organe, jedoch ohne FDG-PET-Untersuchung.

Tab. 3: Bildgebende Diagnostik im Rahmen der EuroNet-PHL-C1-Studie
Modifiziert nach (Körholz et al., 2006)

	Initiales Staging	Erstes Restaging	Zweites Restaging	Follow-up
Zeitpunkt	Vor Beginn der Therapie	14 Tage nach Ende der ersten 2 Chemotherapiezyklen	14 Tage nach Ende der weiteren 4 (TG 2) bzw. 6 (TG 3) Chemotherapiezyklen	Beginn: 6 Wochen nach Abschluss der kompletten Therapie
FDG-PET	+	+	-	Nur bei bestätigtem Rezidiv
Schnittbildgebung (mit KM)	Schädelbasis bis Symphyse: 1. CT Thorax, MRT Hals, Abdomen, Becken <u>oder</u> : 2. CT Hals, Thorax, Abdomen, Becken	Der initial betroffenen LR: 1. CT Thorax, MRT Hals, Abdomen, Becken <u>oder</u> : 2. CT Hals, Thorax, Abdomen, Becken	Der initial betroffenen LR: 1. CT Thorax, MRT Hals, Abdomen, Becken <u>oder</u> : 2. CT Hals, Thorax, Abdomen, Becken	Der initial betroffenen LR: 1. MRT <u>und</u> : 2. CT-Thorax (beim initialen Vorliegen eines Lungenbefalls)
Röntgen	Thorax	-	-	-
Sono	Abdomen	Abdomen	Abdomen	Abdomen

CT = Computertomographie, FDG-PET = Fluordesoxyglucose-Positronenemissionstomographie, KM = Kontrastmittel, LR = Lymphknotenregionen, MRT = Magnetresonanztomographie, Sono = Sonographie, TG = Therapiegruppe

Um das Therapieansprechen bewerten zu können, wurden alle betroffenen Lymphknotenregionen beim Staging und Restaging ausgemessen und deren Volumen bestimmt.

Ein Lymphknoten galt als betroffen, wenn sein längster Durchmesser 2 cm überschritt. Lymphknoten, deren längste Durchmesser weniger als 1 cm betragen, wurden als nicht betroffen betrachtet, ungeachtet des FDG-PET-Ergebnisses. Lymphknoten zwischen 1 und 2 cm im längsten Durchmesser wurden als betroffen betrachtet, wenn die PET-Untersuchung an dieser Stelle positiv bzw. als nicht betroffen betrachtet, wenn sie negativ ausfiel.

Davon ausgehend, dass die Form der Raumforderungen annäherungsweise einem Ellipsoid entspricht, erfolgte die Volumenberechnung anhand folgender Formel:

$$V = \frac{a * b * c}{2}$$

Anschließend wurden die jeweiligen betroffenen Lymphknotenregionen in Anhängigkeit vom residuellen Tumorzvolumen den entsprechenden lokalen Remissions-Kategorien zugeordnet:

Tab. 4: Lokale Remissionskategorien der EuroNet-PHL-C1-Studie

Modifiziert nach (Körholz et al., 2006)

Remissions-Kategorie	Definition
CR	Residuales Tumorzvolumen ≤ 5 % des Ausgangsvolumens und ≤ 2 ml
CRu	Residuales Tumorzvolumen ≤ 25 % des Ausgangsvolumens oder ≤ 2 ml
PR	Residuales Tumorzvolumen ≤ 50 % des Ausgangsvolumens oder ≤ 5 ml
NC	Weder CR, CRu, PR noch PRO
PRO	Residuales Tumorzvolumen ≥ 125 % des Ausgangsvolumens

CR = Complete Remission, CRu = Complete Remission unconfirmed, NC = No Change, PR = Partial Remission, PRO = Progression

Auch der extranodale Befall (Pleura, Pericard usw.) wurde entsprechend der Cotswolds-Kriterien erfasst und der Befall von Leber und Knochenmark entsprechend als Stadium IV gewertet. Für die Beurteilung des Therapieansprechens extranodaler Herde wurden folgende drei Kriterien

unterschieden: Nicht mehr nachweisbarer Befall, weiterhin nachweisbarer Befall und zunehmender Befall. Das Gesamtansprechen (Overall Response) ergab sich aus den Ergebnissen der Response-Beurteilung der schlechtesten nodalen und extranodalen Läsionen.

Für die Zuordnung zu den entsprechenden Response-Gruppen „Adäquates Ansprechen“ und „Inadäquates Ansprechen“ wurde das Gesamtansprechen mit den Ergebnissen der PET-Untersuchung zusammengefasst. Patienten mit PET-positiven Befunden, die sich nicht in kompletter Remission befanden, Patienten mit unklaren PET-Ergebnissen und den Remissionskategorien „Complete Remission unconfirmed“ und „Partial Remission“ sowie alle Patienten der Kategorie „No change“ wurden der Response-Gruppe mit inadäquatem Ansprechen zugeordnet und mussten sich im Gegensatz zu Patienten mit adäquatem Ansprechen nach Ende der Chemotherapie einer zusätzlichen Bestrahlung unterziehen.

1.4.4 Therapiegruppen und Therapieplan

Die Einordnung des Tumorstadiums erfolgte nach der Cotswolds-Modifikation der Ann-Arbor-Klassifikation (siehe Tab. 2, Seite 4). Daraus ergab sich schließlich die Einteilung in die Therapiegruppen:

Tab. 5: Therapiegruppen
(Körholz et al., 2006)

Therapiegruppe	Ann-Arbor-Stadium
1	I A/B und II A
2	I _E A/B, II _E A, II B oder III A
3	II _E B, III _E A/B, III B oder IV A/B

A = ohne B-Symptome, B = mit B-Symptomen, E = Extranodal

Die Therapie erfolgte schließlich nach folgendem Schema:

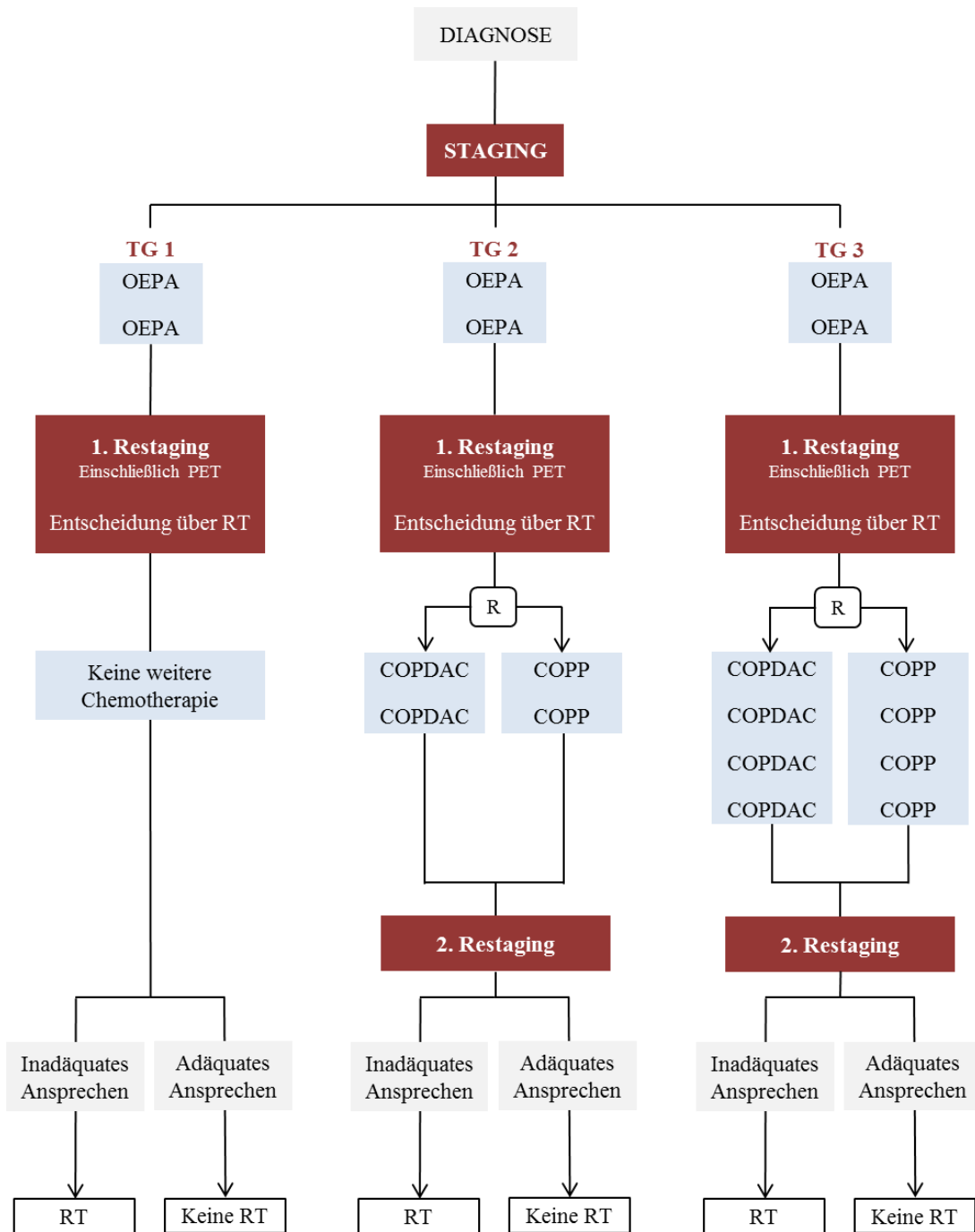


Abb. 1: Therapieplan der EuroNET-PHL-C1-Studie
(Körholz et al., 2006)

COPDAC = Cyclophosphamid, Vincristin, Prednison, Dacarbazin; COPP = Cyclophosphamid, Vincristin, Prednison, Procarbazin; OEPA = Vincristin, Etoposid, Prednison, Adriamycin; PET = Positronen-Emissions-Tomographie; R = Randomisierung; RT = Radiotherapie; TG = Therapiegruppe

2. ZIELSTELLUNG

Ziel dieser Arbeit soll es zum einen sein, durch die retrospektive Auswertung der CT-Daten von Patienten der EuroNet-PHL-C1-Studie verschiedene ein-, zwei- und dreidimensionale Methoden aus den bereits beschriebenen Klassifikationen zur Beurteilung von Tumorthapien hinsichtlich ihrer Übereinstimmung bei der Bestimmung der größten zusammenhängen mediastinalen Tumormasse („bulky disease“) im Rahmen des Hodgkin-Lymphoms im Kindes- und Jugendalter zu vergleichen. Dabei soll eine manuell volumetrische Segmentierungsmethode als Referenzmethode dienen und mit der Ellipsoid-Methode der EuroNet-PHL-C1-Studie sowie der eindimensionalen Berechnung anhand der RECIST-Kriterien und der zweidimensionalen Berechnung anhand der WHO- bzw. Cotswolds- und Cheson-Kriterien verglichen werden.

In einem zweiten Schritt werden alle Methoden hinsichtlich der Beurteilung des Therapieansprechens untersucht und durch Wiederholung der Messungen durch einen zweiten Untersucher bei einem Teil des Patientenguts die Interobservervariabilität und damit die Reproduzierbarkeit aller Methoden bestimmt.

3. MATERIAL UND METHODIK

3.1 Patientenkollektiv

Für diese retrospektive Arbeit wurden 202 Kinder und Jugendliche bis zum vollendeten 18. Lebensjahr untersucht, die neu an einem Morbus Hodgkin mit mediastinaler Beteiligung erkrankten und im Zeitraum von Oktober 2009 bis Juni 2010 aufeinanderfolgend in die EuroNet-Pediatric Hodgkin's Lymphoma-C1-Studie aufgenommen wurden.

Für die Durchführung der Studie liegt ein positives Votum der Ethikkommission der Martin-Luther-Universität Halle-Wittenberg vor.

3.1.1 Einschlusskriterien

Da die Volumetrie verschiedenen Untersuchungen zufolge sowohl beim Hodgkin-Lymphom als auch bei anderen Tumorentitäten sehr exakte Ergebnisse hinsichtlich des „wahren Tumolvolumens“ liefert (Nawaratne et al., 1997; Sohaib et al., 2000; Mukherji et al., 2005; Preda et al., 2007; Prionas et al., 2010), soll sie im Rahmen dieser Arbeit als Referenzmethode für die Bestimmung des Tumolvolumens dienen.

Die Genauigkeit der Resultate der dreidimensionalen Volumenmethode in Bezug auf das „wahre Volumen“ ist jedoch abhängig von der Qualität der vorliegenden Untersuchung, da verschiedenste, v. a. untersuchungs- und patientenbezogene Faktoren die Präzision der Volumetrie beeinträchtigen können (McNitt-Gray et al., 2009). Damit die Methode diese Anforderungen erfüllt, wurden folgende Einschlusskriterien bestimmt:

a) Vorhandensein einer kontrastmittelgestützten CT-Untersuchung:

Voraussetzung für die Bestimmung des Tumolvolumens bzw. die Volumenreduktion war das Vorhandensein einer kontrastmittelgestützten CT-Thorax-Untersuchung sowohl zum Zeitpunkt des Stagings als auch des ersten Restagings der Patienten.

Dass die CT-Volumetrie geeignet ist, das Tumolvolumen beim Hodgkin-Lymphom möglichst exakt abzuschätzen, wurde bereits belegt (Gobbi et al., 2001; Torricelli et al., 2004; Weßling et al., 2014). Kinder, von denen nur Untersuchungen ohne Kontrastmittel oder bei denen keine CT- bzw. nur MRT-Untersuchungen vorlagen, wurden demnach ausgeschlossen.

b) Optimale Kontrastmittelanreicherung:

Für eine optimale Kontrastmittelanreicherung wurde festgelegt, dass die Differenz der Dichtewerte von Muskulatur und Aorta descendens mindestens 40 Hounsfield-Einheiten betragen muss. Dafür wurde jeweils eine größtmögliche „region of interest“ (ROI) in beide Bereiche platziert (siehe Abb. 2). Eine Differenz kleiner 40 HE bedeutete den Ausschluss des Patienten, da ohne

ausreichend intravaskuläres Kontrastmittel insbesondere die Gefäße nicht sicher von den mediastinalen Lymphknoten zu unterscheiden sind. Dies erschwert die Abgrenzbarkeit des Tumors und reduziert damit die Messgenauigkeit.

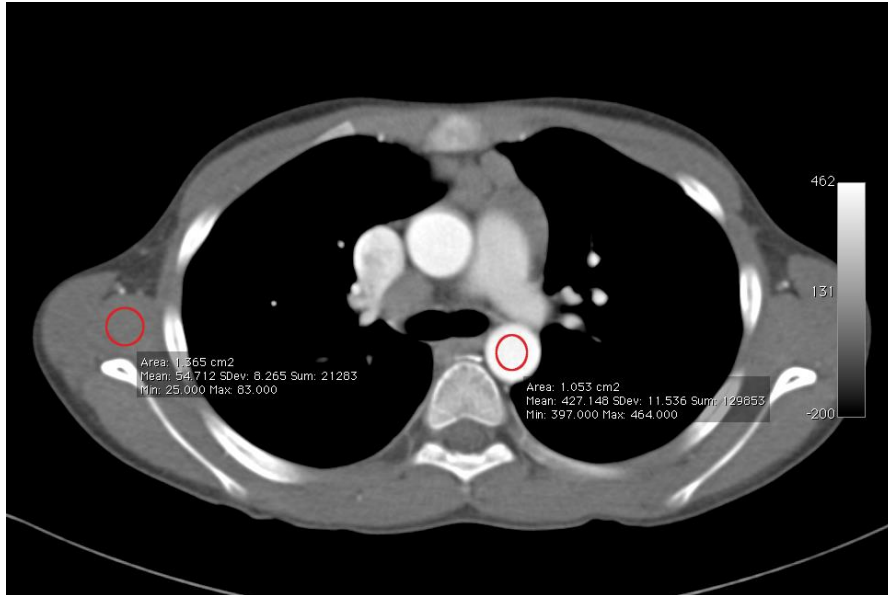


Abb. 2: Kontrastmittelanreicherung

Die Abbildung zeigt die in der Aorta descendens und in der Muskulatur platzierten ROIs zur Beurteilung der Dichteunterschiede.

c) Rekonstruierte Schichtdicke ≤ 5 mm:

Einige Autoren haben bereits den Einfluss der Schichtdicke auf die Durchführung der Volumetrie untersucht und festgestellt, dass der Messfehler mit zunehmender Schichtdicke zunimmt (Petrou et al., 2007; Buerke et al., 2010b). Ursachen dafür sind v. a. im zunehmenden Auftreten von Partial-Volumen-Effekten zu sehen (Reeves et al., 2006).

Daher wurden diejenigen Bilddaten mit einer Schichtdicke von mehr als 5 mm nicht berücksichtigt.

d) Optimale Bildqualität:

Um das Volumen möglichst exakt zu bestimmen, ist es erforderlich, dass sich die Grenzen des Tumors und der umgebenden Gewebe scharf abzeichnen. Ebenso sollten möglichst wenig Atem-, Pulsations- oder andere Artefakte vorliegen, denn diese beeinträchtigen durch unscharfe Konturen und Wiederholung bzw. Verdopplung anatomischer Strukturen die genaue Abgrenzbarkeit des Tumors (siehe Abb. 3).



Abb. 3: Pulsationsartefakte

Die Konturen der Gefäße und der mediastinalen Raumforderung sind unscharf, was zu einer größeren Variabilität bei der Festlegung der Markierungspunkte und damit zu Messungenauigkeiten führt.

Bezugnehmend auf bereits vorliegende Arbeiten zur Bildqualität bei CT-Untersuchungen (Sagara et al., 2010) sind daher folgende Kriterien bestimmt worden, die anhand einer fünf- bzw. dreistufigen Skala beurteilt wurden: Schärfe, Artefakte und Diagnostische Akzeptanz (siehe Tab. 6). Die Note 1 führte zum Ausschluss von dieser Untersuchung.

Die Beurteilung der Bildqualität wurde gemeinsam mit einem Radiologen mit langjähriger Berufserfahrung vorgenommen.

Tab. 6: Kriterien zur Beurteilung der Bildqualität

Modifiziert nach (Sagara et al., 2010)

Skala	Schärfe	Artefakte	Diagnostische Akzeptanz
1	Nicht akzeptabel (Verschwommene/Unscharfe Tumorbegrenzung)	Vorhanden, beeinflussen die Bildinterpretation	Nicht akzeptabel
2	Schlechter als der Durchschnitt (Artefakte der großen Gefäße außer Lungenarterie und Aortenwurzel)	Vorhanden, beeinflussen nicht die Bildinterpretation	Suboptimal
3	Durchschnitt (Ausgeprägte Artefakte an Aortenwurzel und/oder Lungenarterie)	Fehlend	Durchschnitt
4	Besser als der Durchschnitt (Nicht 5 oder 3)	-	Besser als Durchschnitt
5	Sehr genaue Abbildung (Keine oder geringe Artefakte)	-	Ausgezeichnet

3.1.2 Mediastinaler Tumorbefall

Die Auswertung für den Vergleich der verschiedenen Messmethoden zur Tumorgößenbestimmung beschränkte sich in dieser Arbeit auf das Mediastinum. Ein Befall des Mediastinums ist beim Hodgkin-Lymphom sehr häufig. Darüber hinaus gilt ein großer Mediastinaltumor weithin als Risikofaktor, der mit einer Prognoseverschlechterung einhergeht (Mauch et al., 1978; Lee et al., 1980; Willett et al., 1988; Barrington et al., 2014).

Als obere mediastinale Grenze wurde für diese Untersuchung der Unterrand der Schilddrüse festgelegt, da dieser auch bei Patienten mit einem kontinuierlichen Befall aus dem Mediastinum bis in den Hals immer eindeutig zu definieren ist. Die untere Begrenzung wurde durch das Zwerchfell gebildet. Der Untersuchungsbereich umfasste daher folgende Lymphknotenregionen: die bronchopulmonalen Lymphknoten im Bereich der Lungenhili, die zentralen mediastinalen Lymphknoten (im Bereich der großen Gefäße, der Trachea und ihrer Bifurkation und der A. thoracica interna) sowie die Lymphknoten im Bereich der supradiaphragmatischen Recessus.

3.2 Bestimmung der Tumorgöße

3.2.1 Dreidimensionale Tumorgößenbestimmung mit Segmentierungssoftware

Für die volumetrische Messung des Tumorumfanges verwendeten wir die kostenlose Software OsiriX in der Version 5.6 der Schweizer Firma Pixmeo, die sich auf medizinische Bildverarbeitungssoftware spezialisiert hat. OsiriX ist ein Bildbearbeitungssystem für DICOM-Bilddaten. Es dient der Darstellung, Archivierung und Bearbeitung von Bilddaten, die durch medizinische Bildgebung wie CT, MRT, PET, Ultraschall usw. erzeugt werden (siehe Abb. 4).

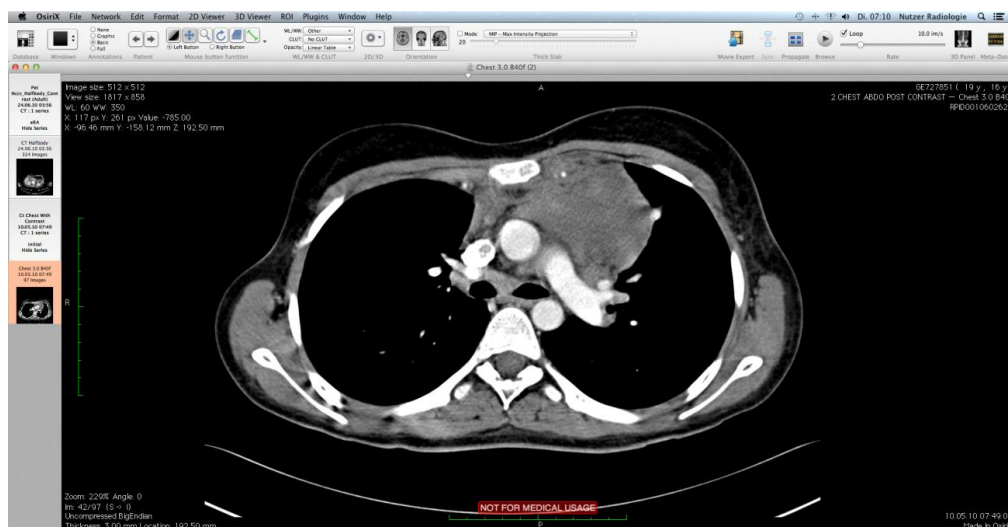


Abb. 4: Benutzeroberfläche der Bildbearbeitungssoftware OsiriX

OsiriX wird laut Herstellerangaben weltweit in vielen Kliniken eingesetzt.

Als Arbeitscomputer diente ein Apple iMac. Zunächst erfolgte der Transfer der anonymisierten CT-Bilddaten vom PACS-Server der Universitätsklinik für Diagnostische Radiologie Halle in die OsiriX-Datenbank des Apple iMac.

Alle Messungen wurden im Weichteilfenster mit der Fenstermitte c (center) = 40 und der Fensterbreite w (width) = 350 durchgeführt. Unter dem Programmpunkt „ROI –Werkzeuge“ wählte man das Tool „Geschlossenes Polygon“. Mit Hilfe dessen konnte nun in allen Schichten, in denen sich das Lymphom nachweisen ließ, die genaue Tumorbegrenzung manuell nachgezeichnet werden. Dafür setzte man einzelne Markierungspunkte entlang der Grenze des Objekts (siehe Abb. 5). Es musste dabei darauf geachtet werden, insbesondere Gefäße, die sich eventuell schlechter kontrastierten, nicht mit in die Messung einzubeziehen. Alternativ ersetzte OsiriX die Tumorbegrenzung zwischen zwei manuell segmentierten Schichten automatisch mittels der Funktion „Fehlende ROIs ergänzen“. Diese konnten im Anschluss noch manuell korrigiert werden.

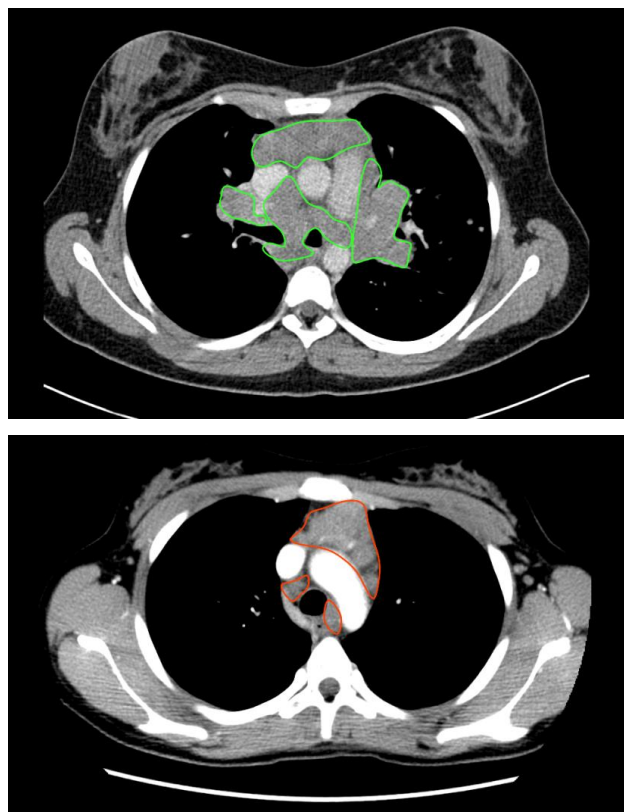


Abb. 5 a und b: Manuelle Festlegung der Tumorgrenzen mit OsiriX

Die Abbildungen zeigen das Ergebnis der manuellen Segmentierung: Die Raumforderung wird mit Hilfe von Markierungspunkten manuell umrandet. Beim Befall multipler Lymphknotenregionen (a) oder beim Vorliegen komplexer Lymphknotenkonglomerate (b) kann das Gesamtvolumen aus mehreren Einzelvolumina gebildet werden.

Nachdem der Tumor in jeder Schicht umrandet und allen ROIs einer Serie ein Name gegeben wurde, erhielt man über eine gleichnamige Funktion das Volumen der Raumforderung. Dabei wurde die manuell eingegrenzte Fläche mit der Schichtdicke multipliziert und die Summe aller Volumina dieser Schichten gebildet. Zusätzlich erfolgte noch eine 3D-Darstellung der Raumforderung (siehe Abb. 6). In gleicher Weise wurden die Messungen für das Restaging vorgenommen.

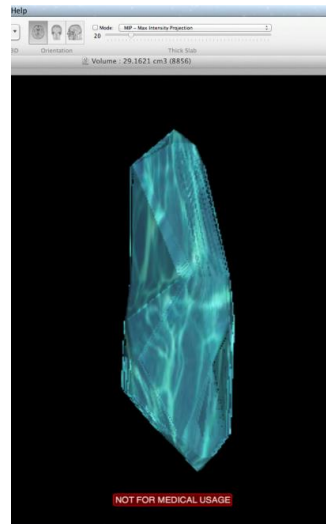


Abb. 6: 3D-Darstellung einer mediastinalen Raumforderung

Aufgrund der Größe und Form der mediastinalen Raumforderungen war es in vielen Fällen nicht möglich das Volumen durch nur eine einzige Messung zu erfassen. Mehrere Einzelmessungen wurden in diesen Fällen zu einem Gesamtvolumen addiert.

3.2.2 Dreidimensionale Tumorgößenbestimmung nach Ellipsoid-Methode

In der EuroNet-PHL-C1-Studie erfolgte die Berechnung des Tumolvolumens nach folgender Formel:

$$V = \frac{a * b * c}{2}$$

Grundlage dafür ist die Annahme, dass Lymphknoten in ihrer Form einem Ellipsoid ähneln, wobei die Variablen a, b und c den drei senkrecht zueinanderstehenden Hauptachsen (Durchmessern) des Ellipsoids entsprechen. Diese Formel stellt eine Vereinfachung der Volumenberechnung eines Ellipsoids dar, in der die Konstante $\pi/6$ unter der Annahme, dass π näherungsweise 3 entspricht, auf $\frac{1}{2}$ gekürzt wird.

In derjenigen axialen Ebene, in der sich der Tumor am größten darstellte, erfolgte die Messung des größten Längsdurchmessers (a) sowie des senkrecht dazu stehenden größten queren

Durchmessers (b) mit Hilfe des digitalen Lineals (siehe Abb. 7). Beim Vorliegen koronarer oder sagittaler Rekonstruktionen konnte in diesen schließlich die längste Ausdehnung in kraniokaudaler Richtung (c) ausgemessen werden. Für den Fall, dass keine Rekonstruktionen vorlagen, errechnete man die kraniokaudale Längenausdehnung, indem man die Anzahl der axialen Schichten, in denen der Tumor abgebildet war, mit der rekonstruierten Schichtdicke multiplizierte.



Abb. 7: Bestimmung der Tumordurchmesser in der axialen Ebene

Die so erhaltenen Messwerte setzte man schließlich in die oben genannte Formel zur Berechnung des Volumens ein. Analog zur Vorgehensweise bei der OsiriX-Methode wurden gegebenenfalls mehrere Einzelmessungen vorgenommen und zu einem Gesamtvolumen addiert. Zum Zeitpunkt des Restagings, also 10 - 14 Tage nach Ende der ersten zwei Chemotherapiezyklen, erfolgte in gleicher Weise die erneute Bestimmung der Tumorgröße.

3.2.3 Zweidimensionale Tumorgrößenbestimmung nach WHO-/Cotswolds- bzw. Cheson-Kriterien

Die WHO-Kriterien sowie die speziell für das Hodgkin-Lymphom ausgearbeiteten Cotswolds- bzw. Cheson-Kriterien sehen eine zweidimensionale Berechnung vor. Als Grundlage für die Beurteilung des Therapieansprechens erfolgt die Bildung der Summe der Produkte der größten Längs- und Querdurchmesser der Läsion (a und b), d. h. man bestimmt eine Tumorfläche. Da für die EuroNet-PHL-C1-Studie bereits die längsten Durchmesser der Tumore in allen drei Raumbenen bestimmt wurden, konnten die entsprechenden Werte diesen Datensätzen entnommen werden.

Für jede Einzelläsion erfolgte schließlich die Bildung des Produkts beider Durchmesser. Aus diesen Ergebnissen wiederum wurde eine Gesamtsumme für die komplette mediastinale Raumforderung gebildet.

Im Folgenden wird, da die Vorgehensweise zur Berechnung des Tumolvolumens dieser Methoden identisch ist, nur noch von der Cheson-Methode gesprochen.

3.2.4 Eindimensionale Tumorgößenbestimmung nach RECIST-Kriterien

Die RECIST-Kriterien empfehlen, den größten Durchmesser in derjenigen Ebene zu bestimmen, in der die Bilder erstellt wurden. Für die CT-Untersuchungen heißt das dementsprechend in der axialen Ebene. Auch die längsten Durchmesser konnten demzufolge dem Datensatz der EuroNet-PHL-Studie entnommen werden. Die größten Durchmesser aller Einzelläsionen wurden schließlich zu einem Gesamtergebnis aufsummiert.

3.3 Umwandlung in Volumina

Um die Ergebnisse der ein- und zweidimensionalen Methoden mit denen der dreidimensionalen Methoden hinsichtlich ihrer Übereinstimmung vergleichen zu können, wurden die Durchmesser und Flächen in ein Volumen umgerechnet. Ausgehend von der Annahme, die Tumorform ähnelt einer Kugel, wurden folgende Umrechnungsformeln aufgestellt:

(1) Umwandlung des Durchmessers in ein Volumen:

Die Berechnung des Volumens einer Kugel erfolgt nach folgender Formel:

$$V = \frac{4}{3}\pi r^3$$

Da für die RECIST-Kriterien die Summe der (längsten) Durchmesser der Läsionen verwendet wird, kann man mit Hilfe dieses Durchmessers auch das Volumen der entsprechenden (Tumor-)Kugel berechnen. Für das Tumolvolumen gilt, wenn $d = 2r$:

$$V = \frac{1}{6}\pi d^3$$

(2) Umwandlung der Fläche in ein Volumen:

Die Fläche eines Kreises wird wie folgt berechnet:

$$A = \pi r^2 \quad \text{bzw.}$$

$$A = \frac{\pi}{4}d^2$$

Der Durchmesser dieses Kreises beträgt dann:

$$d = 2\sqrt{\frac{A}{\pi}}$$

Setzt man diesen Durchmesser schließlich in die Formel zur Berechnung des Volumens ein, so erhält man:

$$V = \frac{1}{6} \pi \left(2 \sqrt{\frac{A}{\pi}} \right)^3 \quad \text{bzw.}$$

$$V = \frac{4}{3\sqrt{\pi}} A^{\frac{3}{2}}$$

3.4 Beurteilung des Therapieansprechens (Response)

Wie bereits im Abschnitt 1.3 dargestellt, wurden je nach angewandter Richtlinie zur Beurteilung des Therapieansprechens unterschiedliche Ansprech-Kriterien definiert. Für die vorliegende Arbeit wird das Ansprechen des Tumors lediglich mit Hilfe des Surrogat-Endpunkts „Tumorzvolumenreduktion im Bereich des Mediastinums“ erfasst. Weitere Kriterien, die in klinischen Studien für die Bewertung des Therapieansprechens ebenso eine Rolle spielen (z. B. Ergebnisse von PET- oder Labor-Untersuchungen), wurden für diese Auswertung nicht berücksichtigt.

Für jede Methode erfolgte daher nach Abschluss der Messungen die Berechnung der entsprechenden Größenreduktion (GR) des Volumens (GR₃), der Summe der Durchmesser (GR₁) und der Summe der Flächen (GR₂) in Prozent.

Dafür wurde folgende Formel verwendet:

$$GR_{1-3} = \left(1 - \frac{\text{Tumorzvolumen bzw. Summe zum Zeitpunkt des Restagings}}{\text{Tumorzvolumen bzw. Summe zum Zeitpunkt des initialen Stagings}} \right) \times 100$$

Um die Ergebnisse der Volumenreduktion mit denen der Durchmesser- und Flächenreduktion vergleichen zu können, erfolgte die Umwandlung der 1D- und 2D-Werte in die entsprechende Volumenresponse. Unter Voraussetzung, dass der Tumor näherungsweise einer Kugel entspricht, ergaben sich folgende Umwandlungsformeln:

(1) Umwandlung GR₁ in GR₃: $GR_3 = 1 - (1 - GR_1)^3$

(2) Umwandlung GR₂ in GR₃: $GR_3 = 1 - \sqrt{(1 - GR_2)^3}$

3.5 Reproduzierbarkeit der Messungen

Um eine Aussage darüber treffen zu können, wie exakt die beschriebenen Messtechniken sind und ob verschiedene Untersucher zu ähnlichen Ergebnissen kommen, führten für eine zufällige Auswahl von 40 Patienten zwei Untersucher die Messungen durch. Die Ellipsoid-Methode wurde

im Rahmen der zentralen Auswertung der Bilddaten aller eingeschlossenen Patienten von einem Oberarzt für Diagnostische Radiologie durchgeführt. Daher sind auch die Werte für die RECIST- und die WHO-Messungen dessen Datensatz entnommen. Die Messung des Tumorumfangs mit Hilfe der Bildbearbeitungssoftware OsiriX erfolgte für alle Patienten durch eine zum Zeitpunkt der Datenerhebung sich im letzten Ausbildungsjahr befindliche Medizinstudentin mit wenig radiologischen Kenntnissen. Vor Beginn der eigentlichen Messungen führte diese an ca. 20 Patienten Probemessungen durch. Die Wiederholung der Messungen an den 40 zufällig ausgewählten Patienten erfolgte durch einen weiteren Oberarzt für Diagnostische Radiologie für alle Methoden. Alle drei Untersucher arbeiteten verblindet gegenüber den Ergebnissen der anderen Untersucher. Lediglich vor Beginn der Messungen erfolgte eine gemeinsame Absprache hinsichtlich des Thymusbefalls der Patienten im Rahmen ihrer Erkrankung und damit der Einbeziehung oder des Ausschlusses des Thymus in die bzw. von der Messung.

3.6 Statistik

Die statistischen Analysen und Abbildungen wurden mit Hilfe von RStudio, einer Open-Source-Software und Benutzeroberfläche für die für statistische Berechnungen und Auswertungen verwendete Programmiersprache R in der Version v0.97 erstellt. Weiterhin fand Microsoft Excel 2010 Verwendung für die Analyse und Auswertung der Daten.

Die Beurteilung der Übereinstimmung der Methoden hinsichtlich der gemessenen Volumina und der Bewertung des Therapieansprechens erfolgte nach dem Ansatz von Bland und Altman (Bland und Altman, 1986; 1996; 1999). Dafür wurden für jedes Wertepaar die absoluten Differenzen in cm³ und deren Mittelwerte berechnet. Darüber hinaus erfolgte die Bestimmung der relativen Differenzen (RD) jedes Wertepaars wie folgt:

$$RD = \frac{M_A - M_R}{M_R} * 100 \%$$

M_A stellt dabei das Messergebnis der jeweiligen Alternativmethode und M_R das Messergebnis der Referenzmethode dar. Von allen RDs konnte schließlich der Mittelwert und damit die mittlere prozentuale Abweichung der Methode vom Referenzstandard, also die systematische Verzerrung bestimmt werden. Zusätzlich wurden die Standardabweichungen und die 95 %-Übereinstimmungsgrenzen berechnet. Bland-Altman-Diagramme dienten zur graphischen Darstellung der Zusammenhänge.

Zur Bewertung der Präzision (Schätzgenauigkeit) der entsprechenden Alternativmethoden erfolgte zusätzlich die Berechnung der Wurzel des mittleren quadratischen Fehlers (RMSE) mit Angabe des 95 %-Konfidenzintervalls:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{prozentualer Fehler}^2)}$$

Dabei stellt $\frac{1}{n} \sum_{i=1}^n (\text{prozentualer Fehler})^2$ den Mittelwert der quadratischen prozentualen Fehler dar. Je größer der RMSE ist, desto geringer ist die Prognosegüte der entsprechenden Methode. Für die Veranschaulichung der Ergebnisse der Interobserver-Analyse wurden ebenfalls Bland-Altman-Diagramme gewählt. Dafür erfolgte jeweils die Berechnung der absoluten Interobserver-Differenzen in cm^3 bzw. cm^2 und cm und deren Mittelwerte. Darüber hinaus wurde für jedes Wertepaar die relative Interobserver-Differenz (RID) wie folgt bestimmt:

$$RID = \frac{US_1 - US_2}{\frac{1}{n} \sum_{i=1}^n (US_1, US_2)} * 100 \%$$

US_1 stellt dabei die jeweilige Messung des 1. Untersuchers, US_2 die Messung des 2. Untersuchers und $\frac{1}{n} \sum_{i=1}^n (US_1, US_2)$ den Mittelwert der Ergebnisse beider Untersucher dar. Aus allen Einzelwerten konnte schließlich die mittlere relative Interobserver-Differenz gebildet werden. Alle Angaben erfolgten mit Standardabweichungen und 95 %-Übereinstimmungsgrenzen. In den Diagrammen wurden die Differenzen der Messwertpaare gegen deren Mittelwerte aufgetragen. Zusätzlich erfolgte die Berechnung des Variationskoeffizienten (VK) nach Bland und Altman (Bland, 2006). Dafür wurde folgende Formel verwendet:

$$VK = \sqrt{\frac{1}{n} \sum_{i=1}^n \left\{ \left(\frac{US_1 - US_2}{US_1 + US_2} \right)^2 \right\}}$$

Der Variationskoeffizient bietet den Vorteil, dass er dimensionslos ist und dem Vergleich der Streuung verschiedener Messreihen dient. Ein niedriger Variationskoeffizient deutet auf eine hohe Reproduzierbarkeit der Methode hin. Zusätzlich wurden die 95 %-Konfidenzintervalle bestimmt.

4. ERGEBNISSE

4.1 Patientencharakteristika

Von den 202 zunächst ausgewählten konsekutiven Patienten der EuroNet-PHL-C1-Studie verblieben nach Berücksichtigung der unter Abschnitt 3.1.1 beschriebenen Ein- und Ausschlusskriterien noch 79 Patienten für den Vergleich der Methoden zur Bestimmung der Tumorgröße bzw. des Tumolvolumens. Die Tab. 7 gibt die Anzahl der vom Ausschluss betroffenen Patienten in Abhängigkeit vom Ausschlusskriterium wieder:

Tab. 7: Ausschlusskriterien und Anzahl der vom Ausschluss betroffenen Patienten

Ausschlusskriterium	Anzahl
Fehlende CT-Untersuchung bzw. fehlendes Kontrastmittel	81
Rekonstruierte Schichtdicke > 5 mm	5
Inadäquate Kontrastmittelanreicherung	11
Inakzeptable Bildqualität	6
Sonstige Ursachen, z. B.:	
unvollständige Abbildung des Tumors oder technische Probleme bei der Messung	20

CT = Computertomographie

Dem Patientenkollektiv gehörten 43 Mädchen (54,4 %) und 36 Jungen (45,6 %) an. Das mittlere Alter der Kinder betrug $13,9 \pm 2,9$ Jahre. Weitere wichtige Daten zu den untersuchten Patienten können der Tab. 8 entnommen werden.

Tab. 8: Patientencharakteristika

	Anzahl	Prozent
Gesamtzahl	79	
Geschlecht		
Weiblich	43	54,4
Männlich	36	45,6
Alter (Jahren)		
Mittelwert	13,9 ($\pm 2,9$)	
Median	15	
Spannweite	4-17	
Herkunft		
Deutschland	30	38
Bulgarien	1	1,3
Dänemark	1	1,3
Großbritannien	19	24,1
Irland	2	2,5
Norwegen	2	2,5
Österreich	3	3,8
Schweden	7	8,9
Schweiz	1	1,3
Slowakei	1	1,3
Spanien	4	5,1
Tschechien	8	10,1
Tumorstadium		
I	0	0
II	37	46,8
III	20	25,3
IV	22	27,8
Therapiegruppe		
1	21	26,6
2	21	26,6
3	37	46,8

4.2 Kenngrößen der gemessenen Tumervolumina

4.2.1 Anzahl der Messungen

Für den Vergleich aller vier Methoden hinsichtlich der Übereinstimmung der gemessenen Volumina und der Übereinstimmung der Beurteilung des Therapieansprechens standen uns unter Berücksichtigung der Ein- und Ausschlusskriterien die Bilddaten von 79 Patienten sowohl zum Zeitpunkt des Stagings als auch des ersten Restagings zur Verfügung. Insgesamt konnten somit jeweils 158 Volumenmessungen bzw. 79 Tumolvolumenreduktionen miteinander verglichen werden.

Für die Bestimmung der Interobservervariabilität wurden von beiden Untersuchern die Ergebnisse der Staging- und der Restaging-Messungen von jeweils 40 zufällig ausgewählten Patienten herangezogen, d. h. 80 Volumenergebnisse konnten verglichen werden.

4.2.2 Tumolvolumen zum Zeitpunkt des Stagings

Für die Segmentierungsmethode mit Hilfe der OsiriX-Software ergab sich initial ein mittleres Volumen von $218,4 \text{ cm}^3 \pm 207,1 \text{ cm}^3$ (Median: $163,7 \text{ cm}^3$, Spannweite: $4,3 \text{ cm}^3 - 1056,2 \text{ cm}^3$). Ein ähnliches Ergebnis lieferte auch die in der Hodgkin-Studie verwendete Ellipsoid-Methode mit einem durchschnittlichen Volumen von $211,2 \text{ cm}^3 \pm 220,1 \text{ cm}^3$ (Median: 138 cm^3 , Spannweite: $2 \text{ cm}^3 - 1016 \text{ cm}^3$). Für die Summe der Flächen aller Einzelläsionen (Cheson-Methode) ergab sich eine mittlere Fläche von $49,8 \text{ cm}^2 \pm 33,5 \text{ cm}^2$ (Median: $38,5 \text{ cm}^2$, Spannweite: $1,5 \text{ cm}^2 - 139,1 \text{ cm}^2$), nach Transformation in das entsprechende Tumolvolumen von $308 \text{ cm}^3 \pm 295 \text{ cm}^3$ im Mittel (Median: 138 cm^3 , Spannweite: $1,4 \text{ cm}^3 - 1233,7 \text{ cm}^3$). Die mittlere Summe der Durchmesser aller Einzelläsionen (RECIST-Methode) betrug $14,7 \text{ cm} \pm 7,3 \text{ cm}$ (Median: $14,4 \text{ cm}$, Spannweite: $1,7 \text{ cm} - 34 \text{ cm}$). Nach der Umwandlung in die Volumina ergab sich hier ein mittleres Volumen von $2992,1 \text{ cm}^3 \pm 3834,7 \text{ cm}^3$ (Median: $1563,6 \text{ cm}^3$, Spannweite: $2,6 \text{ cm}^3 - 20579,5 \text{ cm}^3$).

4.2.3 Tumolvolumen zum Zeitpunkt des Restagings

Nach zwei Zyklen Chemotherapie betrug das mit Hilfe der Segmentierungsmethode bestimmte Tumolvolumen durchschnittlich noch $50 \text{ cm}^3 \pm 56,6 \text{ cm}^3$ (Median: 34 cm^3 , Spannweite: $0,6 \text{ cm}^3 - 286,3 \text{ cm}^3$). Mit Hilfe der Ellipsoid-Methode wurde ein mittleres Volumen von $34 \text{ cm}^3 \pm 39,1 \text{ cm}^3$ (Median: 21 cm^3 , Spannweite: $0,3 \text{ cm}^3 - 86 \text{ cm}^3$) errechnet. Die mittlere Fläche (Cheson-Methode) betrug noch $13,7 \text{ cm}^2 \pm 10,2 \text{ cm}^2$ (Median: $11,8 \text{ cm}^2$, Spannweite: $0,5 \text{ cm}^2 - 44,3 \text{ cm}^2$), was nach Umwandlung im Durchschnitt einem Volumen von $45,7 \text{ cm}^3 \pm 49 \text{ cm}^3$ (Median: $30,3 \text{ cm}^3$, Spannweite: $0,3 \text{ cm}^3 - 221,6 \text{ cm}^3$) entsprach. Die mittlere Summe der Durchmesser (RECIST-Methode) belief sich zum Zeitpunkt des Restagings auf im Mittel $7,9 \text{ cm} \pm 4,5 \text{ cm}$ (Median: $7,2 \text{ cm}$, Spannweite: $0,9 \text{ cm} - 23,1 \text{ cm}$) bzw. das entsprechende Volumen auf $553,2 \text{ cm}^3 \pm 992,2 \text{ cm}^3$ (Median: $195,4 \text{ cm}^3$, Spannweite $0,4 \text{ cm}^3 - 64565,1 \text{ cm}^3$). Die Tab. 9 gibt einen Überblick über die Kenngrößen aller 158, d. h. sowohl zum Staging als auch zum Restaging gemessenen bzw. berechneten Tumolvolumina, welche die Grundlage dieser Untersuchung bildeten. Die Abb.8 a-d veranschaulicht die Verteilung der Tumolvolumina.

Tab. 9: Kenngrößen der jeweiligen Tumervolumina

	Mittelwert ± Standardabweichung	Median	Minimum	Maximum
Segmentierungs-Methode in cm ³	134,2 ± 173,6	66,4	0,6	1056,2
Ellipsoid-Methode in cm ³	122,6 ± 181,2	44,8	0,3	1016
Cheson-Methode in cm ³	176,9 ± 248,8	74,2	0,3	1233,7
RECIST-Methode in cm ³	1772,7 ± 3054,8	485,3	0,4	20579,5

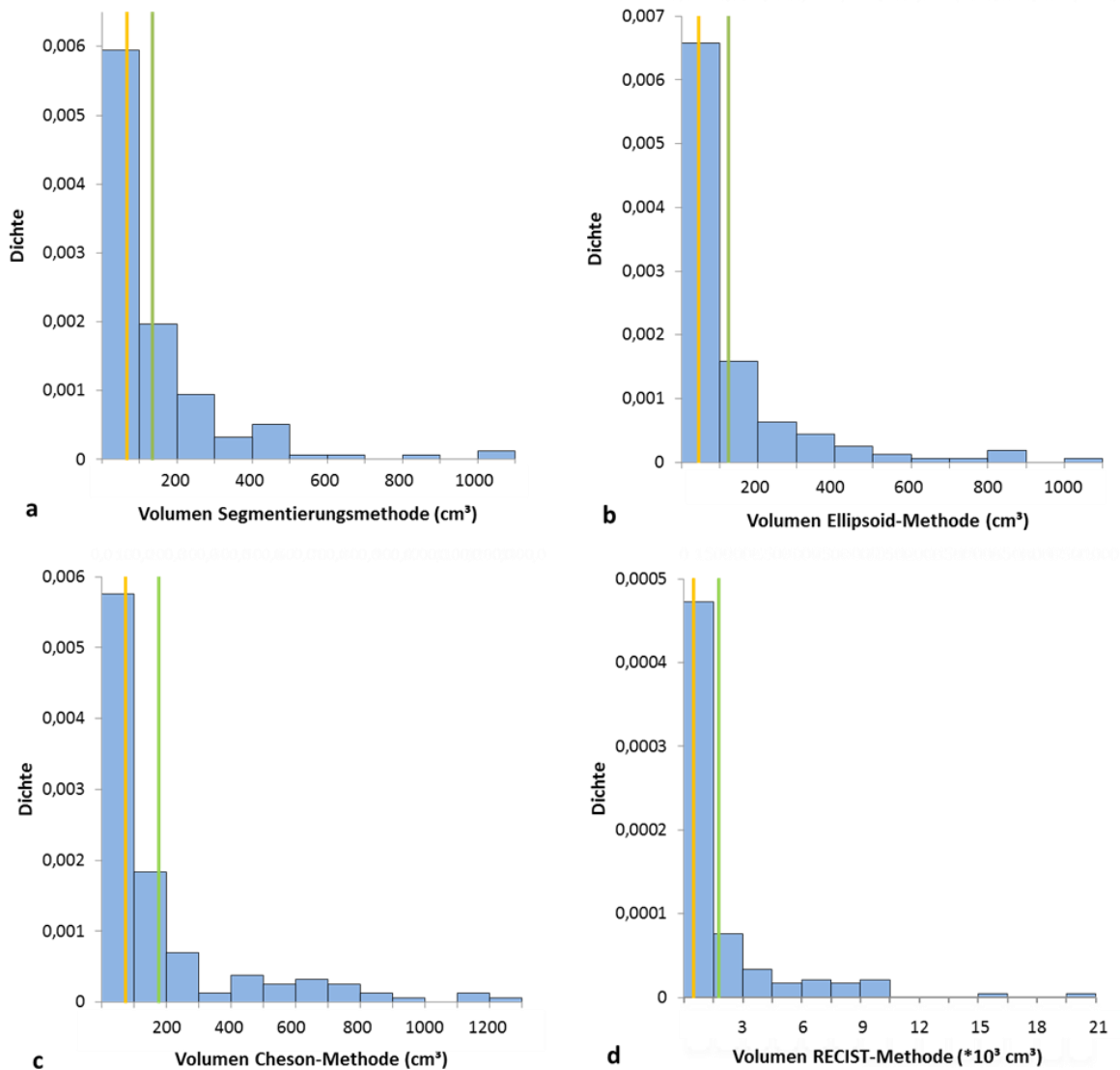


Abb. 8 a-d: Verteilung der Tumervolumina

Die Abbildung veranschaulicht mit Hilfe von Dichtehistogrammen die Verteilung der mittels Segmentierungs- und Ellipsoid-Methode bestimmten (a und b) und der auf Grundlage der Cheson- und RECIST-Methode extrapolierten Tumervolumina (c und d).

(Grüne Linie = Mittelwert; Gelbe Linie = Median)

4.2.4 Tumorgrößenreduktion

Die berechnete prozentuale Reduktion der Tumorumfänge betrug für die Segmentierungsmethode durchschnittlich $79,1 \% \pm 10,6 \%$. Die errechneten Mittelwerte für das Therapieansprechen mittels Ellipsoid-Methode lagen mit $83,5 \% \pm 9,3 \%$ geringfügig höher. Die Umwandlung der Response-Ergebnisse der Cheson-Methode in die entsprechende Volumen-Response führte zu einer mittleren prozentualen Reduktion von $84,8 \% \pm 9,6 \%$, die der RECIST-Methode in die Volumen-Response zu einer prozentualen Reduktion von $82 \% \pm 11,9 \%$ im Mittel. Die Tab. 10 gibt die Kenngrößen der jeweiligen Tumorumfangreduktion wieder.

Tab. 10: Kenngrößen der Tumorumfangreduktion (Response)

	Mittelwert \pm Standardabweichung	Median	Minimum	Maximum
Segmentierungs-Methode in %	79,1 \pm 10,6	81,2	34,1	94,5
Ellipsoid-Methode in %	83,5 \pm 9,3	85	40,8	94
Cheson-Methode in %	84,8 \pm 9,6	85,5	26,9	98,2
RECIST-Methode in %	82 \pm 11,9	84	39,4	99,8

4.3 Übereinstimmung der Methoden hinsichtlich des Tumorumfangs

Die Tab. 11 gibt einen Überblick über die Ergebnisse der Übereinstimmungsuntersuchung aller verwendeten Methoden hinsichtlich der gemessenen Tumorumfänge. Die Abb. 9 a-c und Abb. 10 a-c veranschaulichen die absoluten und relativen Abweichungen der Methoden anhand von Bland-Altman-Diagrammen.

4.3.1 Übereinstimmung Segmentierungs- und Ellipsoid-Methode

Vergleicht man die Ergebnisse der 158 Volumengmessungen zwischen der Ellipsoid-Methode der EuroNet-PHL-Studie und dem Referenzstandard (Segmentierungsmethode), so ergab sich eine mittlere absolute Differenz von $-11,6 \text{ cm}^3 \pm 46,5 \text{ cm}^3$ (siehe Abb. 9 a). Die mittlere prozentuale Differenz belief sich auf $-16,6 \% \pm 31,5 \%$. Das „wahre Volumen“ wurde durch die Ellipsoid-Methode also um durchschnittlich 17 % (= systematische Verzerrung) unterschätzt (siehe Abb. 10 a). Die Wurzel des mittleren quadratischen Fehlers (RMSE) als Parameter für die Schätzgenauigkeit der Methode betrug $35,6 \%$ ($\text{CI}_{95\%}$: $31,3 \%$ - $39,5 \%$).

4.3.2 Übereinstimmung Segmentierungs- und Cheson-Methode

Beim Vergleich der Volumina der Cheson-Methode und der Segmentierungsmethode ergab sich eine mittlere Differenz von $43,1 \text{ cm}^3 \pm 117,6 \text{ cm}^3$ (siehe Abb. 9 b). Die mittlere prozentuale Abweichung betrug $23,8 \% \pm 75,7 \%$ (siehe Abb. 10 b). Das Tumolvolumen der Segmentierungsmethode wurde also auf diese Weise um durchschnittlich 24 % überschätzt. Der RMSE lag sogar bei 79,4 % (CI_{95 %}: 41,4 % - 104,4 %).

4.3.3 Übereinstimmung Segmentierungs- und RECIST-Methode

Nach Transformation der Summen der Durchmesser aller Läsionen in die entsprechenden Volumina ergab sich hinsichtlich der Übereinstimmung mit den Volumina der Segmentierungsmethode eine mittlere absolute Abweichung von $1649 \text{ cm}^3 \pm 2976 \text{ cm}^3$ (siehe Abb. 9 c). Die mittlere prozentuale Differenz lag bei $1195,2 \% \pm 1482,9 \%$ (siehe Abb. 10 c). Die Referenzmethode wurde also deutlich überschätzt. Der RMSE betrug 1904 % (CI_{95 %}: 1506,4 % - 2232,9 %).

Tab. 11: Kenngrößen der Übereinstimmung bei der Bestimmung des Tumolvolumens

	Mittelwert	Standard- abweichung	95 %- Übereinstimmungsbereich	
			Untere Grenze	Obere Grenze
Segmentierung - Ellipsoid				
absolute Differenz (cm ³)	-11,6	46,5	-102,7	79,5
relative Differenz (%)	-16,6	31,5	-78,3	45,1
Segmentierung - Cheson				
absolute Differenz (cm ³)	43,1	117,6	-187,4	273,6
relative Differenz (%)	23,8	75,7	-124,6	172,2
Segmentierung - RECIST				
absolute Differenz (cm ³)	1649	2976	-4184	7482
relative Differenz (%)	1195,2	1482,9	-1711,3	4101,7

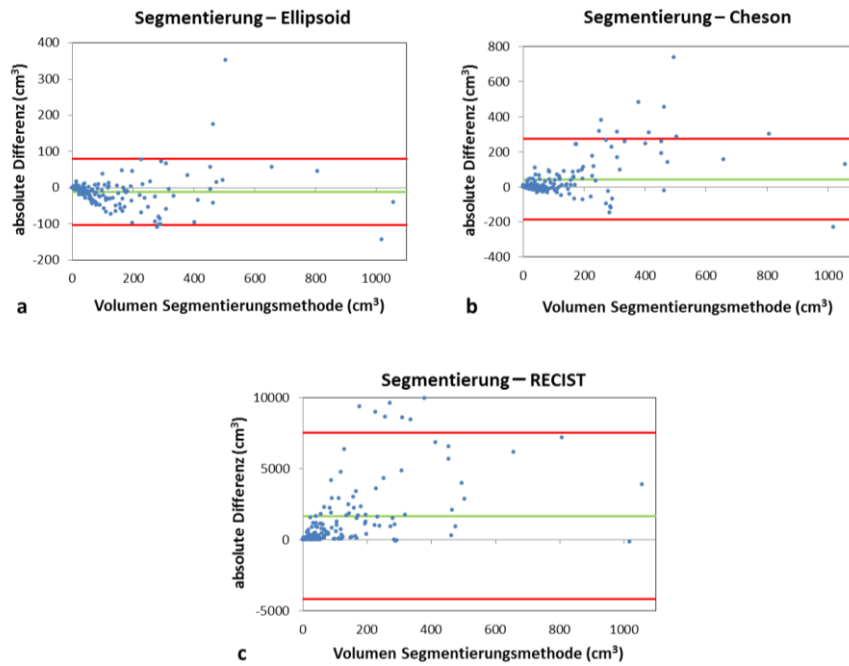


Abb. 9 a-c: Absolute Übereinstimmung der Tumorumfänge

Die Abbildungen zeigen Bland-Altman-Diagramme für die Übereinstimmung der Methoden hinsichtlich der Volumina. Dabei wurde die absolute Differenz der Messwerte gegen das entsprechende Volumen des Referenzstandards (Segmentierungsmethode) aufgetragen. (Grüne Linie = mittlere absolute Abweichung der Methode; rote Linien: 95 %-Übereinstimmungsgrenzen)

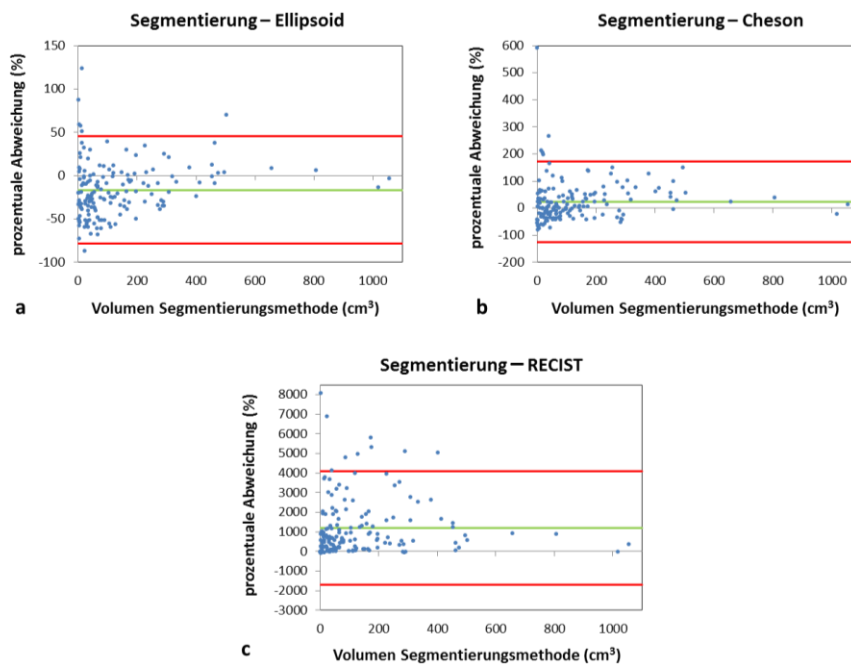


Abb. 10 a-c: Mittlere prozentuale Abweichung

Die Abbildungen zeigen Bland-Altman-Diagramme für die prozentuale Abweichung der Methoden. Dafür wurde die prozentuale Abweichung vom Referenzvolumen für jede Einzelmessung der jeweiligen Alternativ-Methoden gegen das Volumen des Referenzstandards (Segmentierungsmethode) aufgetragen. (Grüne Linie = mittlere prozentuale Abweichung der Methoden; rote Linien = 95 %-Übereinstimmungsgrenzen)

4.4 Übereinstimmung hinsichtlich der Bewertung des Therapieansprechens

Beim Vergleich der Response-Ergebnisse der Ellipsoid-Methode der EuroNet-PHL-Studie mit denen der Segmentierungsmethode ergab sich eine mittlere Differenz von $4,5 \% \pm 8 \%$ ($CI_{95 \%}$: $2,7 \% - 6,2 \%$), das Therapieansprechen wurde also durch die Ellipsoid-Methode um ca. $4,5 \%$ überschätzt.

Ähnliche Ergebnisse fanden sich auch in Bezug auf die Cheson-Methode mit einer mittleren Differenz von ca. $5,7 \% \pm 7,2 \%$ ($CI_{95 \%}$: $4,1 \% - 7,3 \%$). Eine durchschnittliche Differenz von $3 \% \pm 11,6 \%$ ($CI_{95 \%}$: $0,4 \% - 5,5 \%$) ergab sich beim Vergleich mit der RECIST-Methode.

Die Abb. 11 veranschaulicht die Abweichungen der Volumenreduktionsergebnisse der jeweiligen Alternativmethode gegenüber den Ergebnissen der Referenzmethode als kumulative Verteilungsfunktion. Ca. 71 % der Volumenreduktionsergebnisse der Ellipsoid-Methode wichen um weniger als 10 % von den Ergebnissen des Referenzstandards ab. Für die auf das Volumen transformierte Cheson-Methode betrug der Anteil mit einer prozentualen Abweichung von weniger als 10 % ca. 70 % und für die umgewandelte RECIST-Methode ca. 62 %.

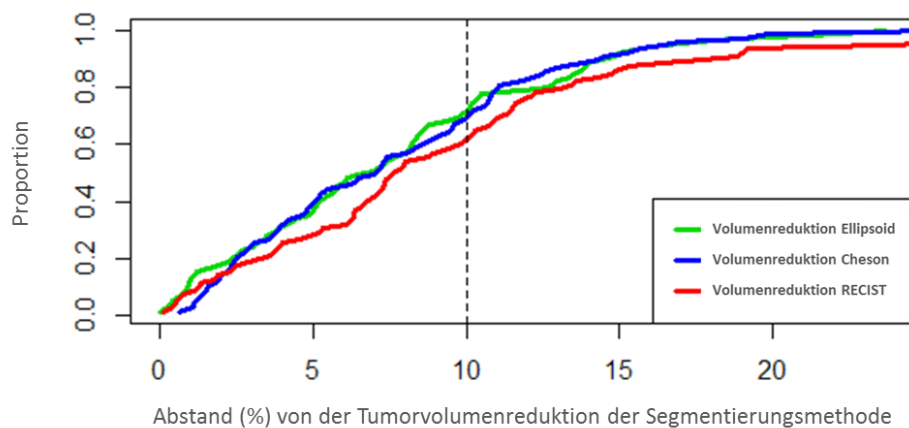


Abb. 11: Abweichungen der Responseergebnisse aller Patienten

Die Abbildung gibt als kumulative Verteilungsfunktion den prozentualen Abstand der Ergebnisse der Alternativmethoden von den Volumenreduktionsergebnissen der Segmentierungsmethode wieder.

Da beim Hodgkin-Lymphom jedoch der Großteil der Patienten ein sehr gutes Therapieansprechen mit einer deutlichen Tumorvolumenreduktion zeigt, sind Abweichungen der Methoden von weniger als 10 % erwartungsgemäß selten. Größere Abweichungen zeigen sich vor allem bei denjenigen Patienten, die ein weniger gutes Therapieansprechen (Tumorvolumenreduktion $< 80 \%$) aufwiesen (siehe Abb. 12). Hierbei wichen nur noch 59 % der Volumenreduktionsergebnisse der Ellipsoid-Methode und 46 % der Volumenreduktionsergebnisse der beiden übrigen Methoden um weniger als 10 % von den Ergebnissen des Referenzstandards ab.

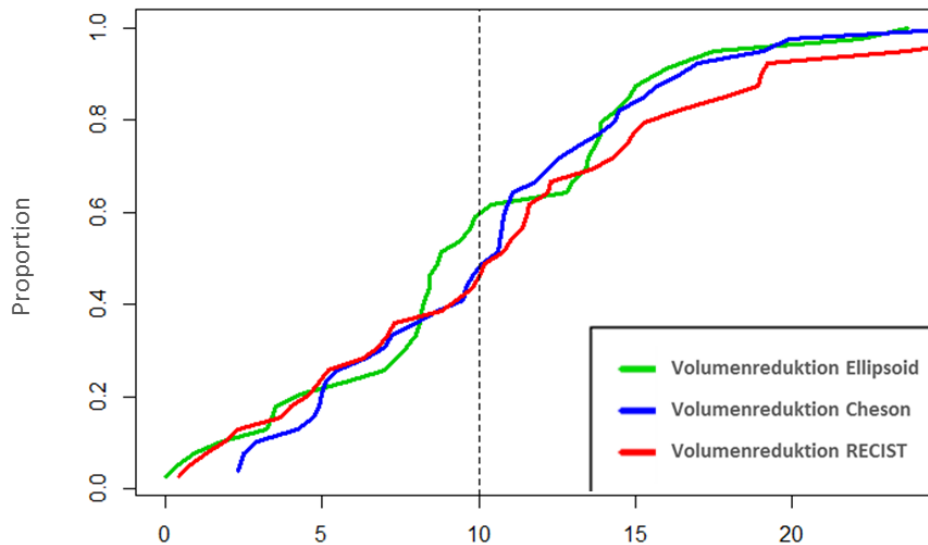


Abb. 12: Abweichungen der Responseergebnisse der Patienten mit einer Tumorumfangsreduktion <80 %

4.5 Interobservervariabilität

4.5.1 Kenngrößen der Volumenmessungen des 2. Untersuchers

Für den 2. Untersucher ergab sich initial für die 40 zufälligen Messungen ein mittleres, mit Hilfe der Segmentierungsmethode bestimmtes Volumen von $202,9 \text{ cm}^3 \pm 184,6 \text{ cm}^3$ (Median: 174 cm^3 , Spannweite: $3,8 \text{ cm}^3 - 1075 \text{ cm}^3$), für die Ellipsoid-Methode von $194,2 \text{ cm}^3 \pm 191,1 \text{ cm}^3$ (Median: $158,3 \text{ cm}^3$, Spannweite: $4,2 \text{ cm}^3 - 1099 \text{ cm}^3$), für die umgewandelte Cheson-Methode von $278,4 \text{ cm}^3 \pm 252,5 \text{ cm}^3$ (Median: $195,3 \text{ cm}^3$, Spannweite: $2,8 \text{ cm}^3 - 957,9 \text{ cm}^3$) und für die transformierte RECIST-Methode von $3003,7 \text{ cm}^3 \pm 3699,2 \text{ cm}^3$ (Median: $1646,4 \text{ cm}^3$, Spannweite: $5,6 \text{ cm}^3 - 15598,5 \text{ cm}^3$).

Nach zwei Zyklen Chemotherapie betrug das Segmentierungs-Volumen zum Zeitpunkt des Restagings noch durchschnittlich $40,3 \text{ cm}^3 \pm 36,8 \text{ cm}^3$ (Median: $32,3 \text{ cm}^3$, Spannweite: $0,5 \text{ cm}^3 - 180 \text{ cm}^3$), das Ellipsoid-Volumen $34,1 \text{ cm}^3 \pm 33,6 \text{ cm}^3$ (Median: $23,5 \text{ cm}^3$, Spannweite: $0,2 \text{ cm}^3 - 142 \text{ cm}^3$), das Cheson-Volumen $42,8 \text{ cm}^3 \pm 35,5 \text{ cm}^3$ (Median: $34,3 \text{ cm}^3$, Spannweite: $0,3 \text{ cm}^3 - 150,7 \text{ cm}^3$) und das RECIST-Volumen $418,4 \text{ cm}^3 \pm 507,2 \text{ cm}^3$ (Median: 221 cm^3 , Spannweite: $0,5 \text{ cm}^3 - 2352,1 \text{ cm}^3$).

Die errechnete prozentuale Tumormassensreduktion betrug für die Segmentierungsmethode im Mittel $79,7 \% \pm 11,4 \%$ (Median: $82,9 \%$, Spannweite: $34,6 \% - 94,3 \%$), für die Ellipsoid-Methode durchschnittlich $81,6 \% \pm 11,5 \%$ (Median: $84,1 \%$, Spannweite: $41,8 \% - 95,2 \%$), für die Cheson-Methode im Mittel $80,5 \% \pm 11,5 \%$ (Median: $81,1 \%$, Spannweite: $53,8 \% - 96,5 \%$) und für die RECIST-Methode im Durchschnitt $78,7 \% \pm 18,1 \%$ (Median: $82,7 \%$, Spannweite: $19,2 \% - 99,5 \%$).

4.5.2 Interobservervariabilität der Messungen

Die Tab. 12 gibt die wichtigsten Kenngrößen der Untersuchung auf Reproduzierbarkeit der Methoden (Interobserver-Untersuchung) wieder. Die Abb. 13 a-d und Abb. 14 a-d veranschaulichen mit Hilfe von Bland-Altman-Diagrammen die absoluten und relativen Interobserver-Abweichungen.

Vergleicht man die Ergebnisse der jeweils 40 Messungen der Segmentierungsmethode beider Untersucher, so ergab sich eine mittlere absolute Interobserver-Differenz von $4,3 \text{ cm}^3 \pm 15,8 \text{ cm}^3$ (siehe Abb. 13 a). Die mittlere relative Differenz betrug $7 \% \pm 17,5 \%$ (siehe Abb. 14 a). Der Variationskoeffizient für diese Messungen belief sich auf $9,4 \%$ ($\text{CI}_{95\%}$: $5,4 \% - 12,2 \%$).

Für den Vergleich der jeweils 40 Messungen nach der Ellipsoid-Methode der EuroNet-PHL-Studie wurde eine mittlere absolute Interobserver-Differenz von $2,5 \text{ cm}^3 \pm 49,4 \text{ cm}^3$ berechnet (siehe Abb. 13 b). Darüber hinaus bestand eine durchschnittliche relative Differenz von $5,4 \% \pm 33 \%$ (siehe Abb. 14 b) mit einem Variationskoeffizienten von $16,7 \%$ ($\text{CI}_{95\%}$: $13,2 \% - 19,6 \%$).

Im Hinblick auf die Übereinstimmung der jeweiligen Messungen beider Untersucher zur Cheson-Methode ergab sich eine mittlere absolute Differenz von $0,7 \text{ cm}^2 \pm 7,8 \text{ cm}^2$ (siehe Abb. 13 c). Die mittlere prozentuale Interobserver-Differenz betrug $5,4 \% \pm 35,2 \%$ (siehe Abb. 14 c). Der Variationskoeffizienten belief sich auf $17,8 \%$ ($\text{CI}_{95\%}$: $13,7 \% - 21,1 \%$).

Der Vergleich der Messungen im Rahmen der RECIST-Methode beider Untersucher lieferte eine mittlere absolute Abweichung von $1 \text{ cm} \pm 3,9 \text{ cm}$ (siehe Abb. 13 d). Die mittlere relative Differenz belief sich auf $8,7 \% \pm 35,7 \%$ (siehe Abb. 14 d) mit einem Variationskoeffizienten von $18,4 \%$ ($\text{CI}_{95\%}$: $14,9 \% - 21,3 \%$).

Tab. 12: Kenngrößen der Interobserver-Untersuchungen

	Mittelwert	Standard- abweichung	95 %- Übereinstimmungsbereich	
			Untere Grenze	Obere Grenze
Segmentierungsmethode				
absolute Differenz (cm ³)	4,3	15,8	-26,7	35,3
relative Differenz (%)	7	17,5	-27,3	41,3
Ellipsoid-Methode				
absolute Differenz (cm ³)	2,5	49,4	-94,3	99,3
relative Differenz (%)	5,4	33	-59,3	70,1
Cheson-Methode				
absolute Differenz (cm ²)	0,7	7,8	-14,6	16
relative Differenz (%)	5,4	35,2	-63,6	74,4
RECIST-Methode				
absolute Differenz (cm)	1	3,9	-6,6	8,6
relative Differenz (%)	8,7	35,7	-61,3	78,7

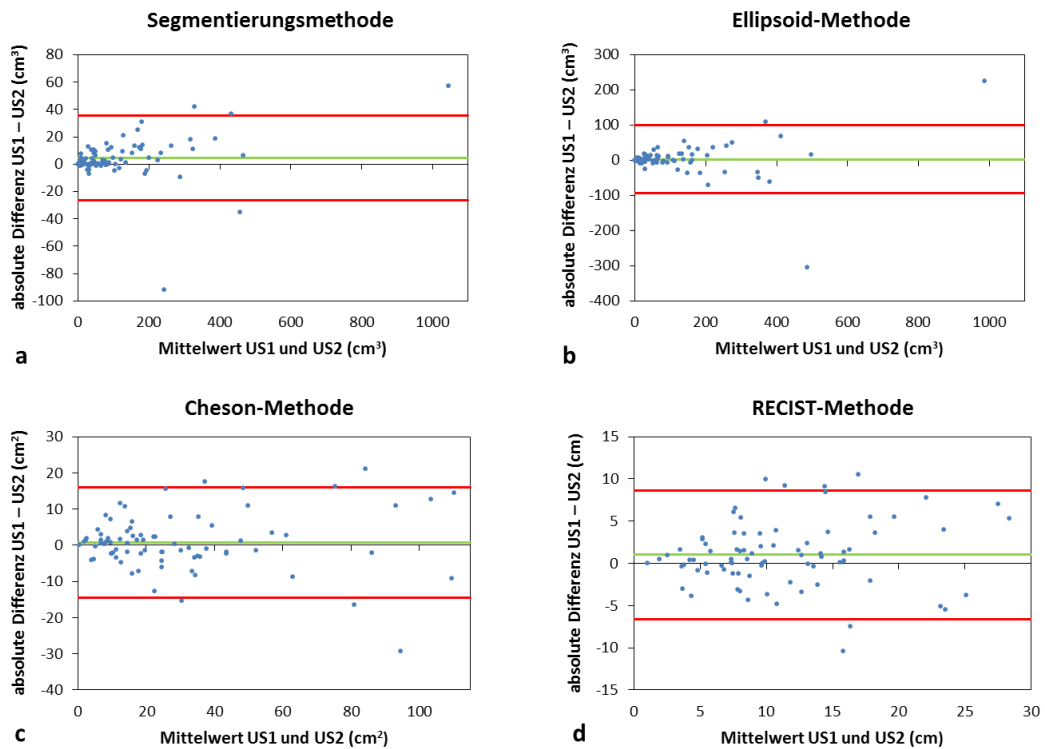


Abb. 13 a-d: Absolute Interobserver-Differenzen

Die Abbildungen veranschaulichen die absolute Abweichung der methodenspezifischen Messergebnisse verschiedener Untersucher. Dafür wurden die Differenzen der Messwertpaare der jeweiligen Untersucher gegen den entsprechenden Mittelwert aufgetragen.

(Grüne Linie = mittlere absolute Interobserver-Differenzen; rote Linien: 95 %-Übereinstimmungsgrenzen; US1= Untersucher 1; US2 = Untersucher 2)

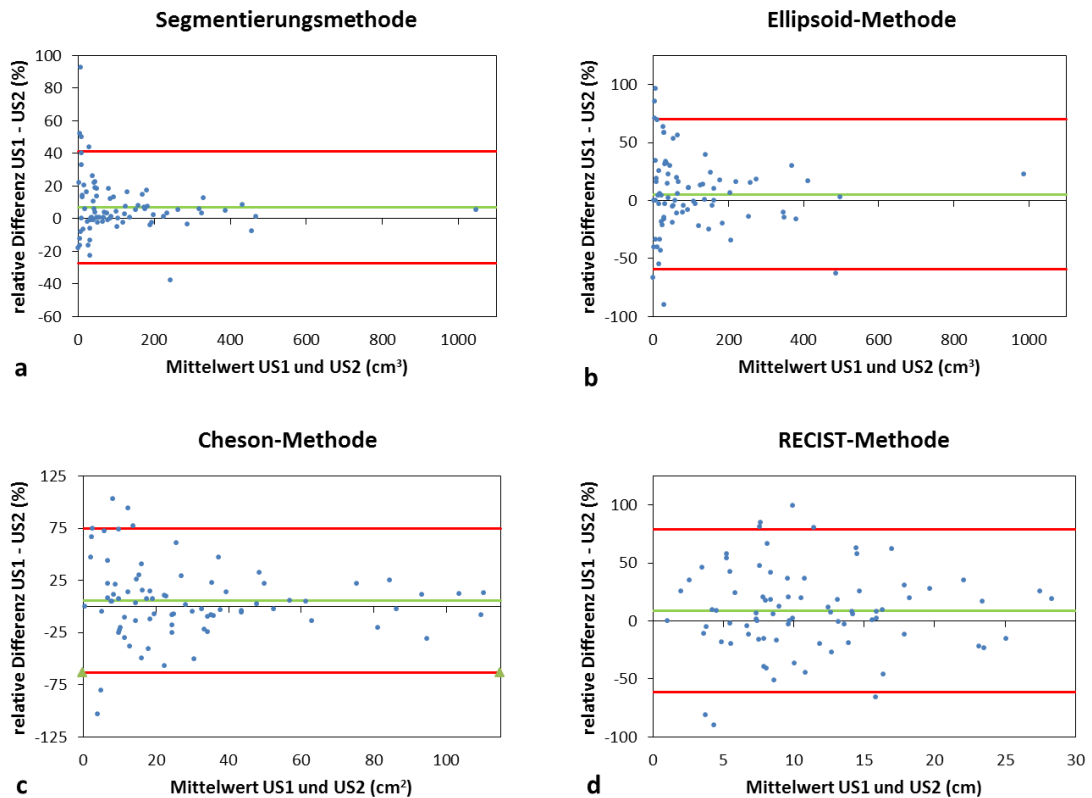


Abb. 14 a-d: Relative Interobserver-Differenzen

Die Abbildungen zeigen die relative Abweichung der methodenspezifischen Messergebnisse verschiedener Untersucher. Dafür wurde die relative Interobserver-Differenz gegen den Mittelwert der Ergebnisse der Untersucher aufgetragen. (Grüne Linie = mittlere absolute Interobserver-Differenzen; rote Linien: 95 %-Übereinstimmungsgrenzen; US1= Untersucher 1; US2 = Untersucher 2)

4.6 Zeitmessungen für die OsiriX-Volumetrie

Für die Segmentierungsmessung der Tumervolumina mit Hilfe der OsiriX-Software wurde im Mittel eine Zeitdauer von 9 Minuten und 53 Sekunden (Median: 7 Minuten und 48 Sekunden) mit einer Spannweite von 12 Sekunden bis 39 Minuten und 58 Sekunden benötigt.

5. DISKUSSION

5.1 Übereinstimmung der Methoden hinsichtlich des Tumolvolumens und des Therapieansprechens

Dass eine große zusammenhängende Tumormasse und insbesondere ein mediastinaler Bulk einen Prognosefaktor beim Hodgkin-Lymphom darstellen, wurde bereits vielfach belegt (Mauch et al., 1978; Lee et al., 1980; Barrington et al., 2014). Konsekutiv werden pädiatrische Patienten in multinationalen Studien wie der EuroNet-PHL-Studie beim Vorliegen großer mediastinaler Tumorumfänge den fortgeschrittenen Tumorstadien zugeordnet und erhalten eine intensiviertere Therapie (Yiallouros, 2016).

Wie bereits unter Abschnitt 1.2 (Seite 5) dargelegt, existieren diverse Richtlinien und Methoden für die Therapieverlaufsbeurteilung und Größenbestimmung von Tumoren. Den Untersuchungsergebnissen dieser Analyse liegt die Annahme zu Grunde, dass die dreidimensionale Segmentierungsmethode das Tumorumfang am exaktesten bestimmen kann (Nawaratne et al., 1997; Sohaib et al., 2000; Preda et al., 2007; Prionas et al., 2010), sodass diese hier als Referenzmethode diente. Vergleicht man die Volumina, welche auf Grundlage der Cheson-Methode bestimmt wurden, mit denen der Referenzmethode, so ergibt sich eine mittlere prozentuale Abweichung von ca. 24 % (CI_{95 %}: 11,7 % - 35,9 %). Eine enorme Überschätzung von im Durchschnitt ca. 1200 % (CI_{95 %}: 958,5 % - 1431,9 %) resultiert, wenn man diejenigen Volumina mit der Referenzmethode vergleicht, die aus der Summe der Durchmesser der Einzelläsionen errechnet wurden. Diese Einschätzung wird durch die Betrachtung des RMSE belegt. Die Wurzel des mittleren quadratischen Fehlers als Indikator für die Prognosegüte der alternativen Messmethoden im Vergleich zum Referenzstandard beträgt für die Cheson-Methode fast 80 % (CI_{95 %}: 41,4 % - 104,4 %) und die RECIST-Methode sogar ca. 1900 % (CI_{95 %}: 1506 % - 2232 %). Im Hinblick auf die komplexe dreidimensionale Ausdehnung mediastinaler Raumforderungen beim Hodgkin-Lymphom sind diese teils deutlichen Abweichungen jedoch nicht verwunderlich. Wie im Abschnitt 1.2.3 (Seite 7) dargestellt, liegt der ein- und zweidimensionalen Methode die Annahme zu Grunde, die Tumorform ähnele einer Kugel (James et al., 1999). Auch in den aktuellsten Empfehlungen einer internationalen Arbeitsgruppe zur Verlaufsbeurteilung maligner Lymphome (RECIL-Kriterien) wird auf dieser Annahme basierend eine eindimensionale Bestimmung dreier repräsentativer Zielläsionen zur Abschätzung der Tumorlast vorgeschlagen (Younes et al., 2017). Die Ursachen für die in dieser Untersuchung erhaltenen Abweichungen der Cheson- und RECIST- (bzw. RECIL-) Methode vom Referenzstandard liegen darin begründet, dass sich die Tumorform der mediastinalen Lymphome mitunter jedoch stark von der einer Kugel unterscheidet. Der Befall der Lymphknoten durch den Morbus Hodgkin im Mediastinum betrifft meist mehrere Lymphknotengruppen bzw. -regionen und nicht selten entstehen Lymphknotenkonglomerate mit sehr komplexer Form. Die

Verwendung lediglich des größten Durchmessers oder der größten Fläche zur Abschätzung des mediastinalen Tumolvolumens muss zwangsläufig zu einer solch ausgeprägten Überschätzung führen, da die Lymphome bzw. Lymphomkonglomerate nicht in allen drei Raumebenen den gleichen Durchmesser aufweisen. Schon *Spears* legte dar, dass die Abschätzung der Tumorgöße durch die Bestimmung lediglich des Durchmessers zu ungenauen Ergebnissen führt, wenn die Tumurlänge mehr als das Zweifache der Tumorbreite beträgt (*Spears*, 1984).

In der EuroNet-PHL-Studie wird zur Beurteilung des Therapieansprechens das Tumolvolumen mit Hilfe einer Ellipsoid-Formel abgeschätzt, wobei die drei senkrecht zueinanderstehenden größten Durchmesser der Raumforderung die Hauptachsen des Ellipsoids darstellen. Der Vergleich der Referenzmethode mit Hilfe der Segmentierungssoftware OsiriX und der Ellipsoid-Methode hinsichtlich der Übereinstimmung der Volumina zeigte, dass die in der Studie verwendete Mess-technik das „wahre“ Tumolvolumen um durchschnittlich 17 % (CI₉₅ %: -21,6 % - 11,6 %) unterschätzte, der RSME betrug 36 % (CI₉₅ %: 31,3 % - 39,5 %). Auch hier ist davon auszugehen, dass die erhaltenen Abweichungen vorwiegend aufgrund der Annahme resultieren, die teils sehr komplexe Form der Lymphome entspräche einem Ellipsoid. Jedoch ist diese Methode etwas genauer in der Abschätzung des mediastinalen Tumolvolumens beim Morbus Hodgkin als die Cheson-Methode.

Neben den methodenbedingten Ursachen für die Verzerrung der Messergebnisse der Alternativmethoden im Vergleich zur Referenzmethode können hierfür auch untersuchungs- und patientenspezifische Faktoren eine Rolle spielen. Viele Autoren beschäftigten sich bereits mit den Ursachen für Verzerrung (Bias) und Varianz bei der Beurteilung des Therapieansprechens auf Grundlage der Tumorgößenänderung (*Ko et al.*, 2003; *Winer-Muram et al.*, 2003; *Kuhnigk et al.*, 2006; *Das et al.*, 2007; *Petrou et al.*, 2007; *Way et al.*, 2008; *Prionas et al.*, 2010). Prinzipiell können die Messergebnisse demnach durch alle Faktoren, die mit dem zur Bildakquisition verwendeten CT-Scanner in Zusammenhang stehen, beeinflusst werden. Dazu gehören insbesondere: der Gerätehersteller, das Geräte-Modell, die Anzahl der Detektorreihen, das Detektormaterial, das Röntgenspektrum, das Filtermaterial sowie die Rekonstruktionsalgorithmen. Darüber hinaus können unterschiedlich eingestellte technische Parameter wie Röhrenspannung, Röhrenstrom, die Kollimation und der Pitch-Faktor ebenso zur Abweichung der Messergebnisse beitragen wie differente Rekonstruktionsparameter (rekonstruierte Schichtdicke, Rekonstruktionskern, rekonstruiertes „field of view“). Schließlich ist die Beeinflussung der Messergebnisse auch durch patientenbezogene (Atmung, Pulsation, Muskelzuckungen usw.) sowie bildanalysebezogene Faktoren (verwendete Segmentierungsmethode, -algorithmus und Software-Version des Bildverarbeitungsprogramms) vorstellbar. Dabei muss jedoch diskutiert werden, in welchem Ausmaß diese Parameter die Präzision der Untersuchungsergebnisse beeinträchtigen und ob sie tatsächlich eine klinische Relevanz besitzen. Im Kontext dieser Untersuchung ist diesbezüglich am ehesten ein Einfluss der rekonstruierten Schichtdicke der Computertomographien denkbar, insbesondere bei kleineren

mediastinalen Raumforderungen bzw. Lymphomresiduen. *Buerke et al.* beispielweise untersuchten den Einfluss der Schichtdicke auf die Ergebnisse einer semiautomatischen Segmentierung zur Bestimmung des Volumens thorakaler Lymphknoten bei Patienten mit NSCLC (*Buerke et al., 2010b*). Dabei zeigten sich prozentuale Abweichungen von im Mittel ca. 22 % bei einer Schichtdicke von 3 mm und sogar von 35 % bei einer Schichtdicke von 5mm im Vergleich zum Referenzstandard (Mittelwert der Segmentierungsergebnisse zweier Radiologen bei einer Schichtdicke von 1 mm). Um den Einfluss der Schichtdicke als Fehlerquelle zu minimieren, wurden in der vorliegenden Analyse lediglich Untersuchungen mit einer rekonstruierten Schichtdicke von bis zu 5mm berücksichtigt. Im Mittel betrug die Schichtdicke zum Zeitpunkt des initialen Stagings ca. 3,3 mm und zum Zeitpunkt des Restagings ca. 3,5 mm, sodass auch dies, v. a. durch Partial-Volumen-Effekte, zu den vorhandenen Abweichungen beigetragen haben kann.

Wie die Abb. 10 a-c (Seite 34) veranschaulichen, sind die prozentualen Fehler aller Alternativmethoden v. a. im Bereich kleiner Läsionen besonders groß. Dieser Tatsache liegen wahrscheinlich mehrere Faktoren zu Grunde: Zunächst bewirken schon kleine Ungenauigkeiten bei der Festlegung der Tumorgrenzen bzw. der größten Tumordurchmesser proportional große Messfehler. Darüber hinaus kann die Abgrenzbarkeit kleinerer Lymphknoten im Bereich des Mediastinums aufgrund ähnlicher Dichteverhältnisse von Lymphknoten, angrenzenden Gefäßen und Organen (z. B. Herz, Ösophagus) sowie Bindegewebe erschwert sein. Außerdem sind für sehr kleine Läsionen, wie sie insbesondere zum Zeitpunkt des Restagings auftreten, bei verhältnismäßig breiter Schichtdicke Partial-Volumen-Effekte größer (*Kuhnigk et al., 2006; Petrou et al., 2007; Buerke et al., 2010b*). Die Abb. 8 a-d (Seite 31) zeigen zudem, dass die Dichte der ausgemessenen Tumorknoten im Bereich kleiner Tumorknoten am größten ist. Da Messungenauigkeiten in diesem Bereich jedoch im Verhältnis zur Größe der Läsion zu proportional größeren prozentualen Abweichungen führen, tragen diese auch mehr zum Gesamtfehler bei. Bei einer ausgeglicheneren Verteilung der Tumorknoten würden die Abweichungen der Methoden insgesamt möglicherweise geringer ausfallen.

Betrachtet man hingegen nur die Übereinstimmung der jeweiligen Tumorknotenreduktionen, so liefern alle Methoden ähnliche Ergebnisse. Im Mittel betrug die Abweichung 4,5 % ($CI_{95\%}$: 2,7 % - 6,2 %) für die Ellipsoid-Methode, 5,7 % ($CI_{95\%}$: 4,1 % - 7,3 %) für die umgewandelte Cheson-Methode und 3 % ($CI_{95\%}$: 0,4 % - 5,5 %) für die transformierte RECIST-Methode. Nimmt man die mit Hilfe der OsiriX-Software bestimmte Tumorknotenreduktion als Grundlage, überschätzen alle drei Alternativmethoden die Ansprechrate, wie sie in dieser Untersuchung definiert wurde, im Mittel also nur leicht. Insgesamt ist dies insbesondere aufgrund der Tatsache nicht überraschend, da die meisten pädiatrischen Patienten beim Morbus Hodgkin eine deutliche Tumorknotenreduktion mit Ansprechraten weit über 80 % aufweisen. Bei denjenigen Patienten mit geringerer Tumorknotenreduktion (< 80 %) konnten auch größere Abweichungen der Alternativmethoden festgestellt werden (siehe Abb. 12, Seite 36).

Viele Autoren haben zuvor ebenfalls die Übereinstimmung ein-, zwei- und dreidimensionaler Verfahren zur Tumorgrößenbestimmung untersucht. Bei der Mehrzahl dieser Arbeiten wurde allerdings auf Basis der Tumorgrößenreduktion die Zuordnung zu den jeweiligen Ansprechkategorien geprüft. Ein Teil dieser Untersuchungen kam zu dem Ergebnis, dass alle Methoden hinsichtlich der Response-Kategorien sehr gut übereinstimmen (Sohaib et al., 2000; Shah et al., 2006). Andere Autoren fanden eine gute Übereinstimmung für die RECIST- und Cheson-Methoden (James et al., 1999), jedoch eine schlechtere Übereinstimmung mit der Segmentierungsmethode (James et al., 1999; Warren et al., 2001; Prasad et al., 2002; Galanis et al., 2006).

In der vorliegenden Arbeit wurde jedoch nicht die Übereinstimmung der Zuordnung der Patienten zu den Ansprechkategorien verglichen. Für das Hodgkin-Lymphom im Kindesalter werden heutzutage sehr hohe Heilungsraten erzielt. Im Durchschnitt lag die Tumorgrößenreduktion in dieser Untersuchung je nach Methode im Bereich von 79 % - 84 %. Die überwiegende Mehrheit der untersuchten Patienten wäre daher bei allen Methoden in die Kategorie „Partial Response“ zugeordnet worden. Für eine aussagekräftige Analyse fehlen insbesondere die Vergleiche bezüglich „Stable Disease“ und „Progressive Disease“. Darüber hinaus sind die Grenzen für die jeweiligen Response-Kategorien willkürlich festgesetzt und divergieren mitunter je nach Studienprotokoll. Bezüglich der Krankheitsprogression beispielsweise wurde in einigen Studien für die Segmentierungsmethode eine Tumorzunahme von mindestens 40 % festgelegt (Sohaib et al., 2000; Warren et al., 2001; Galanis et al., 2006; Shah et al., 2006). In anderen Arbeiten wählte man jedoch einen Grenzwert von 73 % (Prasad et al., 2002; Marten et al., 2006), da dieser laut RECIST-Kriterien (bei Annahme eines kugelförmigen Tumors) mit einer 20-prozentigen Zunahme des Durchmessers korreliert (Therasse et al., 2000). Auch die Anzahl der sogenannten Ziel-Läsionen ist abhängig von der verwendeten Richtlinie bzw. dem Protokoll. So empfehlen die RECIST- bzw. RECIL-Kriterien fünf bzw. drei und die Cheson-Kriterien sechs Target-Läsionen, wohingegen die EuroNet-PHL-Studie das gesamte Tumorzunahme für die Follow-up-Beurteilung verwendet. Auch aus diesem Grund wäre die Zuordnung zu den entsprechenden Ansprechkategorien in diesem Rahmen nur vereinfacht möglich, da Daten, wie z. B. der Befall weiterer Lymphknotenregionen zervikal, abdominal bzw. inguinal, nicht berücksichtigt wurden.

Die bevorzugte Vorgehensweise für diese Untersuchung beinhaltete die Umwandlung der ein- und zweidimensionalen Ergebnisse in dreidimensionale Ergebnisse und ermöglichte so den direkten Vergleich der mediastinalen Volumina und der Tumorzunahme, da das mediastinale Tumorzunahme beim Hodgkin-Lymphom im Kindes- und Jugendalter bzw. in der EuroNet-PHL-Studie einen Risikofaktor mit Therapieentscheidungsrelevanz darstellt. Es sind keine Arbeiten bekannt, die eine ähnliche Vorgehensweise beschrieben. Lediglich *Hopper et al.* verglichen eine in das Volumen konvertierte zweidimensionale Methode bei verschiedenen Tumorzunahme mit einer volumetrischen Methode, jedoch unter Verwendung der Volumenformel eines

Quaders. Auch hier fand sich eine deutliche Überschätzung des „echten“ Volumens (Hopper et al., 1996b).

Weßling *et al.* verglichen bei Patienten mit malignen Lymphomen die Übereinstimmung der manuell ausgemessenen größten Durchmesser der Lymphome nach RECIST-Kriterien mit der nach Cheson-Kriterien bestimmten Fläche und dem mittels Segmentierungssoftware (semi-automatisch) bestimmten Volumen. Für eine bessere Vergleichbarkeit wurden diese in die entsprechenden größten Durchmesser (unter Annahme einer Kugelform) heruntergebrochen. Dabei zeigte sich, dass das Lymphknoten-Volumen und die Lymphknoten-Fläche sehr präzise Ergebnisse hinsichtlich des Ansprechens (bzw. der Größenreduktion) lieferten und der manuellen RECIST-Methode überlegen waren, wobei die mittlere relative Änderung aller manuell und semi-automatisch bestimmten Parameter den Referenzstandard darstellte (Weßling et al., 2012).

Die vorliegenden Ergebnisse verdeutlichen also, dass die Ellipsoid-Methode hinsichtlich der Bestimmung der Tumorumfänge akzeptable Ergebnisse liefert, während die Cheson-Methode weniger und die RECIST-Methode gar nicht geeignet sind, das mediastinale Tumorumfang beim Hodgkin-Lymphom im Kindes- und Jugendalter möglichst exakt zu bestimmen. Ein entscheidender Grund dafür ist die Diskrepanz zwischen der Annahme der kugelförmigen Tumorkonfiguration und der tatsächlichen Form der Lymphome, insbesondere im Mediastinum.

Im Gegensatz liefern jedoch alle Alternativmethoden im Vergleich zum Referenzstandard zuverlässige Ergebnisse bei der Bestimmung der Tumorumfangreduktion.

5.2 Abweichungen zwischen Segmentierungs- und Ellipsoid-Methode

Auch Sorensen *et al.* vergleichen die Übereinstimmung einer auf Segmentierungsbasis beruhenden volumetrischen Methode und einer auf den drei längsten Durchmessern der Raumforderung beruhenden Ellipsoid-Methode bei der Bestimmung des Tumorumfangs von Glioblastomen und kamen zu dem Schluss, dass die letztgenannte Messtechnik durchschnittlich zu kleineren Volumina führt (Sorensen et al., 2001).

Da die Ellipsoid-Methode das Standard-Verfahren zur Abschätzung der Tumorausdehnung in der EuroNet-PHL-Studie darstellt, wurde im Konsens der Untersucher noch einmal speziell Augenmerk auf diejenigen Messungen mit den größten absoluten und relativen Abweichungen dieser Methode vom Referenzstandard gelegt. Dabei fanden sich zwölf Messungen mit Abweichungen von mindestens 75ml im Vergleich zur Segmentierungsmethode. Bei neun dieser zwölf Messungen erhielt man kleinere Volumina mit Hilfe der Ellipsoid-Methode. Hierbei konnte während der gemeinsamen Nachprüfung festgestellt werden, dass am ehesten Ungenauigkeiten bei der Festlegung der größten, senkrecht zueinander stehenden Durchmesser in der axialen Ebene die genannten Abweichungen bewirkten, sodass bei Nachmessungen deutlich kleinere Unterschiede resultierten. Bei denjenigen drei Messungen, bei denen sich das Ellipsoid-Volumen größer als das

Referenzvolumen darstellte, zeigten sich in den Bilddaten komplexe Lymphome mit schmalen Ausziehungen in kraniokaudaler Richtung, die zu einer Überschätzung des „wahren“ Volumens führten.

Ferner konnten elf Messungen mit einer relativen Differenz von 60 bis 124 % im Vergleich zur Referenzmethode nachgewiesen werden, davon wiederum eine zum initialen Staging und zehn zum ersten Restaging. In der gemeinsamen Nachprüfung wurde hierbei festgestellt, dass es sich insbesondere bei den abweichenden Messungen zum Restaging um sehr kleine Tumorreste bzw. -volumina handelt. Wie bereits erwähnt, führen jedoch bei kleinen Läsionen schon geringe Unterschiede bei der Festlegung der Grenzen bzw. der Durchmesser zu prozentual hohen Messfehlern. Da Hodgkin-Lymphome jedoch ein gutes Ansprechverhalten zeigen, finden sich gerade zum Restaging viele kleine Tumorreste, deren Messfehler einen großen Anteil am Gesamtfehler aufweisen. Die abweichende Messung zum Zeitpunkt des initialen Staging mit einem ca. 70 % größeren Ellipsoid-Volumen im Vergleich zum Referenzvolumen resultiert am ehesten aus der komplexen Form dieses Lymphoms mit vielen schmalen Ausziehungen. Diese wurden jedoch im Rahmen der Ellipsoid-Methode nicht einzeln ausgemessen, sondern in den jeweiligen längsten Durchmesser mit einbezogen und führten so zu einer derartigen Überschätzung des „wahren“ Volumens.

Um ein möglichst exaktes Volumen bei der Verwendung der Ellipsoid-Methode zu erzielen, ist es also wichtig, dass in allen Ebenen sorgfältig die größten und senkrecht zueinanderstehenden Durchmesser bestimmt werden. Darüber hinaus ist bei der Bearbeitung zu berücksichtigen, dass einzelne Lymphome mit komplexer Form und vielen schmalen Ausläufern, v. a. in der kraniokaudalen Ausrichtung, das echte Tumolvolumen überschätzen können.

5.3 Reproduzierbarkeit der Methoden

Bei der Validierung von Messmethoden spielt neben der Richtigkeit der Daten jedoch auch die Genauigkeit und damit die Reproduzierbarkeit eine entscheidende Rolle. Im Rahmen der Interobserver-Analyse ergab sich für die Segmentierungsmethode zwischen den Ergebnissen der jeweiligen Untersucher eine mittlere prozentuale Differenz von 7 % ($\pm 17,5$ %). Bei der Ellipsoid- und der Cheson-Methode erhielt man eine etwas geringere mittlere Abweichung von jeweils 5,4 % mit jedoch deutlich größerer Streuung der Werte (SD Ellipsoid-Methode: 33 %, SD Cheson-Methode: 35,2 %). Die RECIST-Methode wiederum wies im Mittel eine Interobserver-Differenz von 8,7 % bei ebenfalls großer Streuung (SD: 35,7 %) auf. Auch die errechneten Variationskoeffizienten deuteten auf eine hohe Reproduzierbarkeit der Segmentierungsmethode (9,4 %, $CI_{95\%}$: 5,4 % - 12,2 %) im Vergleich zu den übrigen Methoden hin, welche insgesamt ähnliche Werte aufwiesen (Ellipsoid-Methode: 16,7 %, $CI_{95\%}$: 13,2 % - 19,6 %; Cheson-Methode: 17,8 %, $CI_{95\%}$: 13,7 % - 21,1 %; RECIST-Methode: 18,4 %, $CI_{95\%}$: 14,9 % - 21,3 %).

Zu ähnlichen Ergebnissen kamen bereits andere Autoren. *Sorensen et al.* beispielsweise verglichen eine Ellipsoid-Methode und eine Segmentierungsmethode zur Bestimmung des Tumorumfanges in MRT-Bilddaten bei Glioblastomen und erhielten eine geringere Interobservervariabilität bei der Segmentierungsmethode (Sorensen et al., 2001). *Fabel et al.* fanden bei der Analyse von Lymphknotenmetastasen bei Patienten mit malignen Melanomen ebenso eine geringere Interobservervariabilität für die Segmentierungsmethoden im Vergleich zur eindimensionalen RECIST-Methode (Fabel et al., 2008). Dabei zeigten sich im Vergleich zur vorliegenden Arbeit etwas geringere relative Differenzen mit durchschnittlich 4,7 % für die RECIST-Methode und 4,1 % für die Segmentierungsmethode bei geringerer Streuung. Ursächlich dafür kann einerseits die weniger komplexe Form der Lymphknotenmetastasen im Vergleich zu mediastinalen Lymphomen sein. Andererseits wurden alle Patienten bei *Fabel et al.* an denselben CT-Geräten und mit einheitlichen Akquisitionsparemtern untersucht. Da die Bilddaten der EuroNet-PHL-Studie aus den verschiedensten Hodgkin-Studienzentren Europas entstammen, variieren sie natürlich stark in Hinblick auf Gerätetyp, Schichtdicke usw., was die Messgenauigkeit reduzieren kann. Die sehr gute Reproduzierbarkeit der manuellen Segmentierungsmethoden wurde auch von *Mukherji et al.*, *Preda et al.* und *Dubus et al.* für Hypopharynx- und Lebertumore bestätigt (Mukherji et al., 2005; Preda et al., 2007; Dubus et al., 2011).

Hopper et al. zeigten bei der Analyse verschiedener Tumorentitäten, dass bei der Bestimmung der jeweils größten Tumordurchmesser je nach Tumorlokalisation, -form und -begrenzung eine unterschiedlich hohe Interobservervariabilität resultiert (Hopper et al., 1996a). Während die relative Differenz bei pulmonalen Läsionen nur durchschnittlich 1 - 4 % betrug, erreichte sie bei Raumforderungen im Mediastinum, im Abdomen oder im Becken bis zu 28 %. Eine insgesamt geringere Reproduzierbarkeit lag bei irregulär geformten und unscharf abgrenzbaren Läsionen vor. Ähnliche Ergebnisse bezüglich der Interobservervariabilität der RECIST-Methode im Vergleich zur vorliegenden Arbeit lieferten auch *Muenzel et al.* (Muenzel et al., 2012). Hier betrug die mittlere relative Interobserver-Differenz 7,1 %. Jedoch beschränkte man sich auf scharf abgrenzbare tumoröse Läsionen mit relativ einfacher Form. *Erasmus et al.* untersuchten die Reproduzierbarkeit der RECIST-Methode und einer zweidimensionalen Methode analog der Cheson-Kriterien bei 60 Patienten mit nicht-kleinzelligem Bronchialkarzinom (Erasmus et al., 2003). Während in der vorliegenden Arbeit die Interobservervariabilität der Verfahren auf Grundlage der Bestimmung der größten Tumordurchmesser (RECIST-, Cheson- und Ellipsoid-Methode) ähnlich hoch ist, erhielten *Erasmus et al.* unter Einbeziehung sowohl gleichmäßig als auch irregulär geformter Raumforderungen eine höhere mittlere relative Interobserver-Differenz bei der Cheson-Methode (20,8 %) im Vergleich zur RECIST-Methode (12,3 %). Dies könnte sich damit begründen lassen, dass sich der Messfehler durch eine zusätzliche, vom jeweiligen Untersucher abhängige Messung des zweiten Durchmessers noch vergrößert. Bei *Zhao et al.* wiesen die RECIST- und Cheson-Methode eine ähnlich gute Reproduzierbarkeit auf wie die

Segmentierungsmethode (Zhao et al., 2009). Diese Diskrepanz im Vergleich zu den vorliegenden Untersuchungsergebnissen kann darin begründet sein, dass dabei Lungentumore (NSCLC) untersucht wurden, die in der Regel weniger komplex konfiguriert sind, sich gut vom umgebenden Lungenparenchym abgrenzen und meist singular auftreten, sodass sich Messfehler und Abweichungen reduzieren. Eine geringere Reproduzierbarkeit der manuellen Segmentierungsmethode im Vergleich zu den auf den größten Tumordurchmessern basierenden Messmethoden beschreiben *Shah et al.* bei der Untersuchung von Glioblastomen (Shah et al., 2006). Wie dieser Unterschied im Vergleich zu den hier vorliegenden Ergebnissen zu Stande kommt, ist nicht nachvollziehbar.

Für maligne Lymphome und Lymphknotenmetastasen verglichen *Buerke et al.*, *Puesken et al.* und *Weßling et al.* manuell durchgeführte Messungen des größten Kurz- und Längsachsendurchmessers von Lymphknoten verschiedener Körperregionen (in Anlehnung an die aktuellen RECIST-Kriterien) mit in diesem Fall jedoch semi-automatisch durchgeführten Messungen auf Grundlage einer Segmentierungssoftware und konnten eine signifikant höhere Reproduzierbarkeit für die semi-automatische Segmentierungsmethode nachweisen (Buerke et al., 2010a; Puesken et al., 2010; Weßling et al., 2014).

Für die Segmentierungs-, aber auch die Ellipsoid- und die Cheson-Methode lagen in dieser Untersuchung die größten prozentualen Abweichungen zwischen den zwei Untersuchern im Bereich kleiner Volumina. Hier unterschieden sich die Messergebnisse zum Teil um bis zu 90 %. Die Ursache dafür liegt darin begründet, dass bei kleinen Läsionen schon geringe Unterschiede beim Festlegen der Grenzen der zu umfahrenden Tumorfläche bzw. bei der Festlegung der größten Durchmesser im Verhältnis zur Tumorgröße große prozentuale Abweichungen resultieren. Darüber hinaus kann die Abgrenzbarkeit des residualen Tumorgewebes vom umgebenden Bindegewebe, dem Herzen bzw. Ösophagus oder den großen Gefäßen aufgrund ähnlicher Dichteverhältnisse und je nach Güte der Kontrastierung erschwert sein. Für kleine Tumore, wie man sie aufgrund des guten Ansprechens der Läsionen auf die Chemotherapie beim Hodgkin-Lymphom insbesondere im Rahmen des Restagings findet, bewirken diese Abweichungen jedoch große prozentuale Fehler, sodass deren Messfehler auch überproportional zum Gesamtfehler beitragen.

5.4 Der Einsatz der Tumorvolumetrie in Therapieoptimierungsstudien

Bildverarbeitungssysteme einschließlich unterschiedlichster Segmentierungsverfahren sind heutzutage fester Bestandteil der radiologischen Routinearbeit. Auch die Volumetrie von Raumforderungen gehört dazu, da für die mit Hilfe von Segmentierungssoftware bestimmten Tumorumfängen ein hohes Maß an Genauigkeit hinsichtlich des „wahren“ Tumorumfanges gilt. Daher wird auch darüber diskutiert, welchen Beitrag sie im Rahmen von onkologischen Therapieverlaufsevaluierungen leisten kann.

Es liegen bereits einige Untersuchungen vor, welche die Präzision der Tumorummetrie anhand von Phantomstudien oder postoperativer Tumorummenbestimmung mit Hilfe der Wasserverdrängungsmethode überprüft haben. Die Fehlerquote der CT-basierten Tumorummetrie belief sich dabei auf durchschnittlich weniger als 5 % (Nawaratne et al., 1997; Sohaib et al., 2000; Preda et al., 2007; Prionas et al., 2010). Auch die Genauigkeit der Volumenmessfunktion der Open-Source-Software OsiriX wurde bereits untersucht, exemplarisch verglichen *Van der Vorst et al.* bei Patienten mit Lebermetastasen im Rahmen eines Kolorektalen Karzinoms das mit Hilfe von OsiriX präoperativ berechnete Volumen der kompletten Leber, des Tumors und des voraussichtlichen Resektats mit den entsprechenden Volumenmessungen herkömmlicher radiologischer Software zur Bildbearbeitung (z. B. iNtuition der Firma TeraRecon, USA). Dabei zeigten sich keine signifikanten Unterschiede hinsichtlich der Ergebnisse der Volumenberechnung beider Programme (van der Vorst et al., 2010). *Weissheimer et al.* verglichen die mit Hilfe sechs verschiedener Bildbearbeitungssoftwares, u.a. auch OsiriX, berechneten Volumina des Hypopharynx von Phantomen mit dem wahren Volumen, bestimmt durch Wasserverdrängungsmethode. Auch hier konnte dargestellt werden, dass OsiriX eine hohe Präzision bei der Bestimmung der Volumina aufweist (Weissheimer et al., 2012). Nur sehr wenige Untersucher gelangten zu dem Schluss, dass die CT-gestützte Bestimmung des Tumorummens nicht geeignet ist, das exakte Tumorummen wiederzugeben. *Tann et al.* beispielsweise verglichen das präoperativ mit Hilfe von Segmentierungssoftware bestimmte Volumen renaler Tumore mit dem postoperativen, mit Hilfe der Wasserverdrängungsmethode ermittelten Volumen und erhielten eine durchschnittliche prozentuale Abweichung von 24,5 % (Tann et al., 2001). Ursache dafür mag das infiltrative Wachstum von Nierentumoren und die dadurch schwierige Differenzierbarkeit von Tumorgrenze und umgebendem Gewebe sein.

In der vorliegenden Arbeit wurde aufgrund der mehrheitlich positiven Datenlage bezüglich der Tumorsegmentierung diese Methode als Referenzstandard für die Bestimmung des mediastinalen Tumorummens beim Hodgkin-Lymphom gewählt. Wie bereits erwähnt, zeigten die Alternativmethoden, betrachtet man deren Präzision bei der Messung des „wahren“ Tumorummens, teils deutliche Abweichungen im Vergleich zur Referenzmethode. Dabei erhielt man mit der Ellipsoid-Methode der EuroNet-PHL-Studie mit einem durchschnittlichen prozentualen Fehler von -17 % noch die genauesten Ergebnisse. Die auf Grundlage der Cheson- und der RECIST-Kriterien basierenden Methoden sind zumindest mediastinal weniger bzw. ungeeignet, die exakte Tumorumlast beim Hodgkin-Lymphom im Kindes- und Jugendalter zu bestimmen. Ursache dafür ist die vereinfachende Annahme, dass die Raumforderungen näherungsweise einer Kugel entsprechen. Gerade für Hodgkin-Lymphome im Bereich des Mediastinums trifft diese Annahme jedoch nicht zu, was zu der zum Teil ausgeprägten Überschätzung der Volumina führt. Weiterhin wurde durch diese Arbeit belegt, dass die Segmentierungsmethode die geringste Interobservervariabilität aller Methoden aufweist.

Die vorliegende Untersuchung bestätigt also, was bereits durch andere Autoren dargelegt wurde: Die auf Grundlage von Segmentierungsverfahren bestimmten Tumorumfänge sind sehr exakt und reproduzierbar. Insbesondere sind diese Verfahren für Raumforderungen geeignet, welche nicht kugelförmig, sondern von komplexer Form sind.

Allerdings muss an dieser Stelle die Frage diskutiert werden, welchen zusätzlichen Nutzen die Bestimmung des exakten Tumorumfanges bezüglich der Auswertung von Tumorthapien in klinischen Studien bringt, insbesondere im Hinblick auf die Prognose und Rezidivwahrscheinlichkeit. Aufgrund der Tatsache, dass entsprechende Daten zum Untersuchungszeitpunkt noch nicht vorlagen bzw. veröffentlicht wurden, war es im Rahmen dieser Untersuchung jedoch nicht möglich, Aussagen bezüglich der prognostischen Relevanz des Tumorumfanges im Hinblick auf beispielsweise das Gesamtüberleben oder das progressionsfreie Überleben zu treffen. *Specht et al.* und *Gobbi et al.* haben die prognostische Bedeutung der Tumormasse beim Hodgkin-Lymphom umfangreich geprüft (Specht et al., 1988; Specht und Nissen, 1988; Gobbi et al., 2001; Gobbi et al., 2004). Beispielsweise konnte anhand der Analyse von 351 Kindern und Erwachsenen mit neu diagnostiziertem Hodgkin-Lymphom dargelegt werden, dass die Tumormasse den besten Prädiktor der „Zeit bis zum Therapieversagen“ sowie des „Krankheitsfreien Überlebens“ und nach dem Alter den zweitbesten Prädiktor für das „Gesamtüberleben“ darstellt (Gobbi et al., 2004). Auch in anderen Untersuchungen wurde die prognostische Relevanz der mittels Computertomographie ermittelten Tumormasse beim Morbus Hodgkin (Torricelli et al., 2004) und bei anderen Tumorentitäten (Jagannath et al., 1986; Mukherji et al., 1999; Kraas et al., 2001; Lee et al., 2008; Chung et al., 2009) bestätigt. Die Bedeutung des Tumorumfanges für die Prognose beim Hodgkin-Lymphom ist demnach an anderer Stelle umfassend belegt worden.

Ferner werden Größenänderungen von dreidimensionalen Objekten, in diesem Fall von Tumoren, mit Hilfe der Volumetrie wahrscheinlich eher und sensitiver detektiert als mit der Ausmessung lediglich eines Durchmessers. Denn legt man die Annahme zu Grunde, Tumore seien kugelförmig, entspräche eine Tumormassenzunahme von 40 % im Volumen in etwa einer Progression von 12 % im Durchmesser (Therasse et al., 2000). Insbesondere bei Patienten, welche nicht oder nur schlecht auf eine gegebene Therapie ansprechen, ermöglichen die frühzeitig detektierbaren Tumorumfangänderungen rechtzeitige therapeutische Entscheidungen (Gobbi et al., 2004). Darüber hinaus konnte bereits nachgewiesen werden, dass die Volumetrie den metrischen Methoden (RECIST- und Cheson-Methode) bei der Unterscheidung benignen oder malignen Lymphknoten im Rahmen von malignen Lymphomen aufgrund der Unabhängigkeit der Messung von der räumlichen Anordnung und Konfiguration überlegen ist (Puesken et al., 2010).

Andere Autoren jedoch stellten den zusätzlichen prognostischen Wert der Tumorumfanguetrie in Frage. *Ferrari et al.* beispielsweise verglichen die prognostische Bedeutung des Tumordurchmessers und des Tumorumfanges beim Rhabdomyosarkom und stellten fest, dass die dreidimensionale Erfassung der Tumormasse im Vergleich zur eindimensionalen keine zusätzlichen

Informationen hinsichtlich der Prognose liefert (Ferrari et al., 2010). Ähnlich stellt es sich auch bei anderen Analysen dar (Shah et al., 2006). *Warren et al.* fanden ein längeres medianes progressionsfreies Überleben bei der eindimensionalen Methode im Vergleich zur zwei- und dreidimensionalen Methode, jedoch war der Unterschied nur zur zweidimensionalen Methode signifikant (Warren et al., 2001).

In den aktuellen Empfehlungen zur Verlaufsbeurteilung maligner Lymphome bzw. den Protokollen aktueller Studien zur Behandlung des Hodgkin-Lymphoms wird weiterhin die Bedeutung einer großen mediastinalen Tumormasse für die Prognose betont. Konsekutiv werden Patienten mit einer sogenannten „bulky disease“ den fortgeschrittenen Tumorstadien zugeordnet und entsprechend intensiver therapiert. Ein mediastinaler Bulk liegt nach der Lugano-Klassifikation dann vor, wenn ein Lymphknotenkonglomerat mehr als 10cm bzw. ein Drittel des Thoraxdurchmessers ausmacht und wird demzufolge eindimensional ausgemessen (Cheson et al., 2014). In der Euro-Net-PHL-Studie wiederum, in welcher die Tumorlast volumetrisch erfasst wird (Ellipsoid-Methode), gilt ein Lymphknotenkonglomerat von größer 200ml als Bulk (Yiallourous, 2016). Es bestehen also auch heutzutage noch sehr unterschiedliche Definitionen bezüglich dieses Parameters. Auf die sich daraus ergebenden Diskrepanzen hatten auch bereits *Willet et al.* aufmerksam gemacht. Sie untersuchten 53 Patienten mit frühen Stadien des Morbus Hodgkin bezüglich der Korrelation zwischen dem eindimensional bestimmten Verhältnis von größtem Mediastinaldiameter und Thoraxdiameter (entsprechend der Cotswolds-Kriterien) und des computertomographisch gemessenen Volumens des größten mediastinalen Lymphknotenkonglomerats. Dabei zeigte sich beispielweise, dass die Volumina von Patienten mit einem Mediastinum-Thorax-Verhältnis von 0,2 mit 30 bis 210 ml stark variierten. Weiterhin fiel auf, dass Patienten mit einem Mediastinum-Thorax-Verhältnis von 1:3 oder kleiner, welche also im Sinne der Cotswolds-Kriterien keine „bulky disease“ auswiesen, häufiger ein Tumorrezidiv entwickelten, wenn das initial bestimmte mediastinale Tumolvolumen 200ml übertraf (Willet et al., 1988). Ob die volumetrisch bestimmte (mediastinale) Tumorlast insbesondere im Hinblick auf eine entsprechende Therapieentscheidung zusätzlichen Nutzen bringen kann, ist also ebenfalls noch Gegenstand der aktuellen Diskussion und kann hier nicht abschließend geklärt werden.

Ein wesentlicher Nachteil der verwendeten manuellen Segmentierungsmethode ist der Zeitaufwand. Je nach Tumorgröße und rekonstruierter Schichtdicke betrug die Zeit für die OsiriX-Messungen zwischen 12 Sekunden und 40 Minuten (Mittelwert: 9 Minuten 53 Sekunden \pm 8 Minuten 12 Sekunden). Ursächlich ist zum einen die teils sehr komplexe Form der mediastinalen Raumforderungen und zum anderen ein recht hoher Korrekturbedarf der Tumorgrenzen, wenn man die Funktion wählte, nicht segmentierte Schichten automatisch vom Programm zu ergänzen. Demzufolge ist diese Methode für einen routinemäßigen Einsatz in einer multizentrischen Studie eher ungeeignet. Wesentliche Zeitersparnis können dahingehend semiautomatische oder automatische Segmentierungsmethoden bieten. Während diese Methoden für Lungentumore und -metastasen

(Marten et al., 2006; Zhao et al., 2009) sowie für Lebertumore (Keil et al., 2009c; Dubus et al., 2011) bereits seit längerem verwendet werden, ist der Einsatz solcher Verfahren für das Hodgkin-Lymphom noch begrenzt. Einige Autoren haben die (semi-)auto-matische Segmentierung bereits an Lymphknoten (-metastasen) untersucht. Obwohl die halbautomatischen Segmentierungsergebnisse insgesamt als gut bewertet wurden, müssen derzeit noch viele Segmentierungsergebnisse korrigiert werden, was insbesondere thorakale Lymphknoten betrifft (Feuerstein et al., 2009; Keil et al., 2009a; Keil et al., 2009b; Buerke et al., 2010a).

Trotz möglicher Zeitersparnis und besserer Reproduzierbarkeit (semi-)automatischer gegenüber manueller Segmentierungsmethoden ist eine Standardisierung der Volumetrie aufgrund von unterschiedlichster verwendeter Software und Algorithmen nur eingeschränkt möglich. Diese spielt jedoch wiederum eine entscheidende Rolle für die Vergleichbarkeit klinischer Studien (Wormanns, 2005).

Gegen den Einsatz der Tumorvolumetrie in einer multizentrischen Studie, wie der EuroNet-PHL-Studie, sprechen weiterhin die hohen Anforderungen an die Qualität der Bilddaten. Der Einfluss verschiedener Bildakquisitions- und Rekonstruktionsparameter auf das Ergebnis der Tumorvolumetrie wurde bereits ausführlich geprüft. Die dreidimensionalen Segmentierungsalgorithmen beispielsweise liefern umso bessere Ergebnisse, je größer das zu untersuchende Objekt und je dünner die Schichtdicke ist (Buerke et al., 2010b). In einer Studie mit Phantomen belegten *Prionas et al.*, dass bei einer Schichtdicke von 5 mm das wahre Volumen um bis zu 39 % unterschätzt wurde (Prionas et al., 2010). Auch der Rekonstruktionsalgorithmus oder der Rekonstruktionsfilter können je nach Tumortyp und -lokalisierung das Segmentierungsergebnis beeinflussen (Ko et al., 2003; Kuhnigk et al., 2006). *Das et al.* untersuchten den Einfluss des CT-Gerätetyps (vier verschiedene 16-Zeilen-Multidetektor-CT) auf das Ergebnis der Volumensegmentierung anhand von Phantomen, wobei statistisch signifikante Unterschiede mit einem prozentualen Fehler von bis zu 14,3 % resultierten (Das et al., 2007). Aufgrund der Tatsache, dass die Bilddaten innerhalb einer länderübergreifenden Studie wie der EuroNet-PHL-Studie hinsichtlich der genannten Parameter natürlich variieren, wurden die in Abschnitt 2.1.1 beschriebenen Qualitätskriterien festgelegt, um die Tumorvolumina mit Hilfe des Segmentierungsverfahrens möglichst exakt bestimmen zu können. Allerdings erfüllten nur etwa 40 % aller in diese Untersuchung eingeschlossenen Patienten diese Anforderungen, wobei die meisten (insgesamt 80) aufgrund einer fehlenden kontrastmittelgestützten Staging- oder Restaging-Untersuchung ausgeschlossen werden mussten. Werden gewisse Qualitätsstandards jedoch nicht erfüllt, muss mit abweichender Genauigkeit bei der Bestimmung des mediastinalen Tumorumfanges gerechnet werden.

5.5 Ausblick: Die Verlaufsbeurteilung des Hodgkin-Lymphoms im Kindesalter

Insgesamt liegen kaum Daten zur vergleichenden Beurteilung der verwendeten Tumorgrößenbestimmungsmethoden beim Morbus Hodgkin im Kindes- und Jugendalter vor. Während über viele Jahre die Cotswolds-Kriterien mit Hilfe der zweidimensionalen Abschätzung der Tumorlast die Grundlage für die Therapieverlaufsevaluierung beim Morbus Hodgkin darstellten, wird der Einsatz der RECIST-Kriterien bis heute kontrovers diskutiert, da spezielle Gegebenheiten pädiatrischer Tumore nicht berücksichtigt werden (McHugh und Kao, 2003; Barnacle und McHugh, 2006; McCarville, 2008). Insbesondere die Annahme einer kugelförmigen Tumormorphologie mit Empfehlung zur eindimensionalen Ausmessung der Lymphome in der axialen Ebene stellt eine potentielle Fehlerquelle dar, da hier die Komplexität und Heterogenität v. a. der mediastinalen Lymphome nicht berücksichtigt wird. Ob die Cheson-Kriterien oder die Tumorummetrie bezüglich der Therapieverlaufsevaluierung und letztlich für die Prognose entscheidende Vorteile beim Morbus Hodgkin liefern, ist noch immer Gegenstand der aktuellen Diskussion, wobei sich jedoch zumindest letztere bisher in multinationalen Studien nicht durchsetzen konnten.

Mit Hilfe der funktionellen Bildgebung stehen dem Onkologen heutzutage im Rahmen der Staging- und Follow-up-Untersuchungen jedoch auch Methoden zur Verfügung, die sich nicht allein auf die anatomische Größenänderung beschränken. Denn Blutfluss und Stoffwechselaktivität sind Parameter, die sich unter onkologischer Therapie sogar frühzeitig ändern können als die Tumorgröße. Gerade für das Hodgkin-Lymphom ist die FDG-PET ein in Studien fest etabliertes Verfahren zur Beurteilung des Therapieansprechens (Kobe et al., 2008; Cheson et al., 2014; Barrington und Kluge, 2017). Hier erhöht sie zum einen die Sensitivität der morphologischen Bildgebung bei der Detektion befallener Lymphknoten und liefert zum anderen auch wichtige prognostische Hinweise. Noch nicht vollständig geklärt ist die Bedeutung PET-negativer Befunde von in der Bildgebung jedoch weiterhin erkennbaren Tumorresiduen nach initialer Chemotherapie (Furth et al., 2009). Dies ist einer der Gründe, warum die Kombination aus anatomischer und funktioneller Bildgebung sicherlich auch in Zukunft die Basis für die Therapieverlaufsbeurteilung beim Morbus Hodgkin im Kindes- und Jugendalter darstellt, während gleichzeitig der Einsatz neuer bildgebender Verfahren und Biomarker untersucht wird.

6. ZUSAMMENFASSUNG

Die vorliegende Arbeit untersuchte verschiedene Messverfahren zur Bestimmung des mediastinalen Tumorumfangs beim Hodgkin-Lymphom hinsichtlich ihrer Übereinstimmung und des Therapieansprechens, in diesem Fall der Tumorumfangsreduktion. Für den Vergleich wurden gebräuchliche radiologische Methoden zur Evaluierung des Therapieansprechens in onkologischen Studien ausgewählt: eine eindimensionale Methode auf Grundlage der RECIST-Kriterien, eine zweidimensionale Methode auf Grundlage der Cheson-Kriterien sowie eine dreidimensionale Ellipsoid-Methode auf Basis der größten drei orthogonalen Tumordurchmesser aus der EuroNet-PHL-Studie. Als Referenzstandard diente hier die dreidimensionale Tumorumfangsmessung auf Grundlage eines manuellen Segmentierungsverfahrens. Um die Reproduzierbarkeit aller Methoden zu untersuchen, erfolgte zusätzlich die Berechnung der Interobserver-Variabilität.

Es konnte gezeigt werden, dass die Ergebnisse aller Methoden hinsichtlich der Bestimmung des exakten Tumorumfangs von denen der Referenzmethode abwichen, wobei die Ellipsoid-Methode der EuroNet-PHL-Studie als beste Alternativmethode (mittlere prozentuale Abweichung ca. -17 %) durchaus als Surrogat für die Abschätzung des mediastinalen Tumorumfangs in Frage kommt. Die auf Grundlage der Cheson-Richtlinien errechneten Volumina überschätzen das Referenzvolumen mäßig und sind daher weniger geeignet, das Tumorumfangs beim Morbus Hodgkin möglichst exakt wiederzugeben (mittlere prozentuale Abweichung ca. 24 %). Aufgrund der zum Teil deutlichen Überschätzung der Volumina, welche auf Basis der RECIST-Kriterien ermittelt wurden (mittlere prozentuale Abweichung ca. 1200 %), ist diese Methode für die Abschätzung des mediastinalen Tumorumfangs beim Hodgkin-Lymphom ungeeignet. Hauptursache für diese Abweichungen ist die vereinfachende Annahme dieser Methoden, die Form eines Tumors bzw. der mediastinalen Lymphome ähnele einer Kugel.

Hinsichtlich der Bewertung des Therapieansprechens, also in der vorliegenden Arbeit hinsichtlich der prozentualen Tumorumfangsreduktion, überschätzen alle drei metrischen Methoden die Referenzmethode nur leicht (Ellipsoid-Methode: 4,5 %, Cheson-Methode 2D: 5,7 %; RECIST-Methode: 3 %) und können daher alternativ eingesetzt werden.

Beim Vergleich der Interobserver-Variabilität wies die Tumorumfangsmessung auf Grundlage des Segmentierungsverfahrens die höchste Reproduzierbarkeit auf (VK: 9,4 %). Die Ellipsoid-Methode (VK: 16,7 %), die Cheson-Methode (VK: 17,8 %) und die RECIST-Methode (VK: 18,4 %) zeigten ähnliche, jedoch insgesamt etwas höhere Interobserver-Variabilitäten.

Der Vergleich aller Methoden zeigt also, dass die auf Grundlage der Segmentierungsmethode berechneten Volumina am genauesten und am besten reproduzierbar sind. Gegen den Einsatz dieser Methode in einer multinationalen Studie sprechen jedoch der enorme Zeitaufwand, der mitunter für den einzelnen Patienten benötigt wird, und die hohen Anforderungen an die Bildqualität, aufgrund derer sich weniger als die Hälfte der Patienten aus diesem Kollektiv tatsächlich dafür

eigneten, das Tumolvolumen möglichst exakt zu bestimmen und damit die Tumorgrößenreduktion genau zu beurteilen. Aufgrund des geringeren Zeitaufwands und der mäßigen Abweichungen bezüglich des exakten Tumolvolumens ist die Ellipsoid-Methode in großen Therapieoptimierungsstudien sicher praktikabler.

Für das Hodgkin-Lymphom ist die prognostische Relevanz einer großen, zusammenhängenden Tumormasse vielfach belegt worden. Welchen zusätzlichen Nutzen die exakte Bestimmung des Volumens im Gegensatz zum größten Durchmesser oder der größten Fläche des Bulks insbesondere im Hinblick auf die Therapieentscheidung und Prognose liefert, ist jedoch weiterhin Gegenstand aktueller Diskussion und kann hier nicht abschließend geklärt werden.

7. LITERATURVERZEICHNIS

- Barnacle AM, McHugh K (2006) Limitations with the response evaluation criteria in solid tumors (RECIST) guidance in disseminated pediatric malignancy. *Pediatr Blood Cancer* 46:127–134
- Barrington SF, Kluge R (2017) FDG PET for therapy monitoring in Hodgkin and non-Hodgkin lymphomas. *Eur. J. Nucl. Med. Mol. Imaging* 44:97–110
- Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Müller SP, Schwartz LH, Zucca E, Fisher RI, Trotman J, Hoekstra OS, Hicks RJ, O'Doherty MJ, Hustinx R, Biggi A, Cheson BD (2014) Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *Journal of Clinical Oncology* 32:3048–3058
- Biggar RJ, Frisch M, Goedert JJ (2000) Risk of cancer in children with AIDS. AIDS-Cancer Match Registry Study Group. *JAMA* 284:205–209
- Bland M (2006) How should I calculate a within-subject coefficient of variation? <http://www-users.york.ac.uk/~mb55/meas/cv.htm>. Zugegriffen: 22. August 2017
- Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310
- Bland JM, Altman DG (1996) Measurement error proportional to the mean. *BMJ* 313:106
- Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* 8:135–160
- Bradley AJ, Carrington BM, Lawrance JA, Ryder WD, Radford JA (1999) Assessment and significance of mediastinal bulk in Hodgkin's disease: comparison between computed tomography and chest radiography. *J. Clin. Oncol.* 17:2493–2498
- Buerke B, Puesken M, Mütter S, Weckesser M, Gerss J, Heindel W, Wessling J (2010a) Measurement accuracy and reproducibility of semiautomated metric and volumetric lymph node analysis in MDCT. *AJR Am J Roentgenol* 195:979–985
- Buerke B, Puesken M, Beyer F, Gerss J, Weckesser M, Seifarth H, Heindel W, Wessling J (2010b) Semiautomatic lymph node segmentation in multislice computed tomography: impact of slice thickness on segmentation quality, measurement precision, and interobserver variability. *Invest Radiol* 45:82–88
- Cheson BD, Horning SJ, Coiffier B, Shipp MA, Fisher RI, Connors JM, Lister TA, Vose J, Grillo-Lopez A, Hagenbeek A, Cabanillas F, Klippensten D, Hiddemann W, Castellino R, Harris NL, Armitage JO, Carter W, Hoppe R, Canellos GP (1999) Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI Sponsored International Working Group. *J Clin Oncol* 17:1244

- Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, Coiffier B, Fisher RI, Hagenbeek A, Zucca E, Rosen ST, Stroobants S, Lister TA, Hoppe RT, Dreyling M, Tobinai K, Vose JM, Connors JM, Federico M, Diehl V (2007) Revised response criteria for malignant lymphoma. *J. Clin. Oncol.* 25:579–586
- Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, Lister TA (2014) Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 32:3059–3068
- Chung E-J, Lee N-J, Baek S-K, Kwon S-Y, Woo J-S, Jung K-Y (2009) Clinical Efficacy of Primary Tumor Volume Measurements: Comparison of Different Primary Sites. *Clin Exp Otorhinolaryngol* 2:78
- Cuccaro A, Bartolomei F, Cupelli E, Galli E, Giachelia M, Hohaus S (2014) Prognostic factors in hodgkin lymphoma. *Mediterranean journal of hematology and infectious diseases* 6:e2014053
- Das M, Ley-Zaporozhan J, Gietema HA, Czech A, Mühlenbruch G, Mahnken AH, Katoh M, Bakai A, Salganicoff M, Diederich S, Prokop M, Kauczor H-U, Günther RW, Wildberger JE (2007) Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners. *Eur Radiol* 17:1979–1984
- Dubus L, Gayet M, Zappa M, Abaleo L, Cooman A de, Orioux G, Vilgrain V (2011) Comparison of semi-automated and manual methods to measure the volume of liver tumours on MDCT images. *Eur Radiol* 21:996–1003
- Eisenhauer E, Therasse P, Bogaerts J, Schwartz L, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J (2009) New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer* 45:228–247
- Erasmus JJ, Gladish GW, Broemeling L, Sabloff BS, Truong MT, Herbst RS, Munden RF (2003) Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. *J. Clin. Oncol.* 21:2574–2582
- Fabel M, Tengg-Kobligk H von, Giesel FL, Bornemann L, Dicken V, Kopp-Schneider A, Moser C, Delorme S, Kauczor H-U (2008) Semi-automated volumetric analysis of lymph node metastases in patients with malignant melanoma stage III/IV-a feasibility study. *Eur Radiol* 18:1114–1122
- Ferrari A, Miceli R, Meazza C, Casanova M, Favini F, Morosi C, Trecate G, Marchianò A, Luksch R, Cefalo G, Terenziani M, Spreafico F, Polastri D, Podda M, Catania S, Schiavello E, Giannatempo P, Gandola L, Massimino M, Mariani L (2010) Comparison of the prognostic value of assessing tumor diameter versus tumor volume at diagnosis or in response to initial chemotherapy in rhabdomyosarcoma. *J. Clin. Oncol.* 28:1322–1328

- Feuerstein M, Deguchi D, Kitasaka T, Iwano S (2009) Automatic Mediastinal Lymph Node Detection in Chest CT. Graduate School of Information Science, Nagoya University, Japan. https://www.researchgate.net/publication/242156208_Automatic_mediastinal_lymph_node_detection_in_chest_CT. Zugegriffen: 10. März 2018
- Filipovich AH, Mathur A, Kamat D, Shapiro RS (1992) Primary immunodeficiencies; Genetic risk factors for lymphoma. *Cancer research* 52:5465s-5467s
- Furth C, Steffen IG, Amthauer H, Ruf J, Misch D, Schönberger S, Kobe C, Denecke T, Stöver B, Hautzel H, Henze G, Hundsdoerfer P (2009) Early and late therapy response assessment with 18Ffluorodeoxyglucose positron emission tomography in pediatric Hodgkin's lymphoma: analysis of a prospective multicenter trial. *J. Clin. Oncol.* 27:4385–4391
- Galanis E, Buckner JC, Maurer MJ, Sykora R, Castillo R, Ballman KV, Erickson BJ (2006) Validation of neuroradiologic response assessment in gliomas: measurement by RECIST, two-dimensional, computer-assisted tumor area, and computer-assisted tumor volume methods. *Neuro-oncology* 8:156–165
- Gavrielides MA, Kinnard LM, Myers KJ, Petrick N (2009) Noncalcified lung nodules: volumetric assessment with thoracic CT. *Radiology* 251:26–37
- Gobbi PG, Ghirardelli ML, Solcia M, Di Giulio G, Merli F, Tavecchia L, Bertè R, Davini O, Levis A, Brogna C, Maffè GC, Ilariucci F, Dore R, Ascari E (2001) Image-aided estimate of tumor burden in Hodgkin's disease: evidence of its primary prognostic importance. *J. Clin. Oncol.* 19:1388–1394
- Gobbi PG, Brogna C, Di Giulio G, Mantelli M, Anselmo P, Merli F, Zinzani PL, Rossi G, Callea V, Iannitto E, Paulli M, Garioni L, Ascari E (2004) The clinical value of tumor burden at diagnosis in Hodgkin lymphoma. *Cancer* 101:1824–1834
- Handels H (2009) Segmentierung medizinischer Bilddaten. In: *Medizinische Bildverarbeitung*. Vieweg+Teubner
- Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, Lister TA, Bloomfield CD (2000) The World Health Organization classification of neoplastic diseases of the haematopoietic and lymphoid tissues: Report of the Clinical Advisory Committee Meeting, Airlie House, Virginia, November 1997. *Histopathology* 36:69–86
- Hasenclever D, Diehl V (1998) A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. *The New England journal of medicine* 339:1506–1514
- Hodgkin T (1832) On some Morbid Appearances of the Absorbent Glands and Spleen. *Medico-chirurgical transactions* 17:68–114
- Hopper KD, Diehl LF, Lynch JC, McCauslin MA (1991) Mediastinal bulk in Hodgkin disease. Method of measurement versus prognosis. *Invest Radiol* 26:1101–1110

- Hopper KD, Kasales CJ, van Slyke MA, Schwartz TA, TenHave TR, Jozefiak JA (1996a) Analysis of interobserver and intraobserver variability in CT tumor measurements. *AJR Am J Roentgenol* 167:851–854
- Hopper KD, Kasales CJ, Egli KD, TenHave TR, Belman NM, Potok PS, van Slyke MA, Olt GJ, Close P, Lipton A, Harvey HA, Hartzel JS (1996b) The impact of 2D versus 3D quantitation of tumor bulk determination on current methods of assessing response to treatment. *J Comput Assist Tomogr* 20:930–937
- Jagannath S, Velasquez WS, Tucker SL, Fuller LM, McLaughlin PW, Manning JT, North LB, Cabanillas FC (1986) Tumor burden assessment and its implication for a prognostic model in advanced diffuse large-cell lymphoma. *J Clin Oncol* 4:859–865
- James K, Eisenhauer E, Christian M, Terenziani M, Vena D, Muldal A, Therasse P (1999) Measuring response in solid tumors: unidimensional versus bidimensional measurement. *J. Natl. Cancer Inst.* 91:523–528
- Kaatsch P, Grabow D, Spix C. (2016) German Childhood Cancer Registry - Annual Report 2016 (1980-2015). Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI) at the University Medical. <http://www.kinderkrebsregister.de/dkk/ergebnisse/jahresberichte.html>. Zugegriffen: 12. Juni 2017
- Kaplan HS, Rosenberg SA (1966) The Treatment of Hodgkin's Disease. *Medical Clinics of North America*:1591–1610
- Keil S, Plumhans C, Behrendt FF, Stanzel S, Suehling M, Mühlenbruch G, Mahnken AH, Günther RW, Das M (2009a) Automated measurement of lymph nodes: a phantom study. *Eur Radiol* 19:1079–1086
- Keil S, Behrendt F, Stanzel S, Suehling M, Jost E, Mühlenbruch G, Mahnken A, Günther R, Das M (2009b) RECIST and WHO Criteria Evaluation of Cervical, Thoracic and Abdominal Lymph Nodes in Patients with Malignant Lymphoma: Manual versus Semi-Automated Measurement on Standard MDCT Slices. *Fortschr Röntgenstr* 181:888–895
- Keil S, Plumhans C, Behrendt FF, Stanzel S, Suehling M, Mühlenbruch G, Mahnken AH, Günther RW, Das M (2009c) Semi-Automated Quantification of Hepatic Lesions in a Phantom. *Investigative Radiology* 44:82–88
- Knowles DM (1999) Immunodeficiency-associated lymphoproliferative disorders. *Modern pathology* : an official journal of the United States and Canadian Academy of Pathology, Inc 12:200–217
- Ko JP, Rusinek H, Jacobs EL, Babb JS, Betke M, McGuinness G, Naidich DP (2003) Small pulmonary nodules: volume measurement at chest CT-phantom study. *Radiology* 228:864–870
- Kobe C, Dietlein M, Mauz-Körholz C, Engert A, Borchmann P, Sabri O, Schober O, Schicha H, Kluge R (2008) FDG-PET beim Hodgkin-Lymphom. *Nuklearmedizin* 6:235–238

- Körholz D, Wallace WH, Landman-Parker J (2006) EuroNet-Paediatric Hodgkin's Lymphoma Group: First international Inter-Group Study for classical Hodgkin's Lymphoma in Children and Adolescents [Studienprotokoll].
- Kraas JR, Underhill TE, D'Agostino RB, Williams DW, Cox JA, Greven KM (2001) Quantitative analysis from CT is prognostic for local control of supraglottic carcinoma. *Head Neck* 23:1031–1036
- Kriz J, Mueller R-P, Mueller H, Kuhnert G, Engert A, Kobe C, Haverkamp U, Eich HT (2012) Large mediastinal tumor mass as a prognostic factor in Hodgkin's lymphoma. Is the definition on the basis of a chest radiograph in the era of CT obsolete? *Strahlenther Onkol* 188:1020–1024
- Kuhnigk J-M, Dicken V, Bornemann L, Bakai A, Wormanns D, Krass S, Peitgen H-O (2006) Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. *IEEE Trans Med Imaging* 25:417–434
- Kuppers R, Klein U, Hansmann ML, Rajewsky K (1999) Cellular origin of human B-cell lymphomas. *The New England journal of medicine* 341:1520–1529
- Lee C-C, Chu S-T, Ho H-C, Lee C-C, Hung S-K (2008) Primary tumor volume calculation as a predictive factor of prognosis in nasopharyngeal carcinoma. *Acta Otolaryngol.* 128:93–97
- Lee CK, Bloomfield CD, Goldman AI, Levitt SH (1980) Prognostic significance of mediastinal involvement in Hodgkin's disease treated with curative radiotherapy. *Cancer* 46:2403–2409
- Lehmann T, Oberschelp W, Pelikan E, Repges R (1997) *Bildverarbeitung für die Medizin; Grundlagen, Modelle, Methoden, Anwendungen.* Springer-Verlag, Berlin-Heidelberg
- Lister TA, Crowther D, Sutcliffe SB, Glatstein E, Canellos GP, Young RC, Rosenberg SA, Coltman CA, Tubiana M (1989) Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. *J. Clin. Oncol.* 7:1630–1636
- Marten K, Auer F, Schmidt S, Kohl G, Rummeny EJ, Engelke C (2006) Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria. *Eur Radiol* 16:781–790
- Mauch P, Goodman R, Hellman S (1978) The significance of mediastinal involvement in early stage Hodgkin's disease. *Cancer* 42:1039–1045
- Mauz-Körholz C, Gorde-Grosjean S, Hasenclever D, Shankar A, Dorffel W, Wallace WH, Schellong G, Robert A, Korholz D, Oberlin O, Hall GW, Landman-Parker J (2007) Resection alone in 58 children with limited stage, lymphocyte-predominant Hodgkin lymphoma—experience from the European network group on pediatric Hodgkin lymphoma. *Cancer* 110:179–185

- Mauz-Körholz C, Lange T, Hasenclever D, Burkhardt B, Feller AC, Dorffel W, Kluge R, Vordermark D, Korholz D (2015) Pediatric Nodular Lymphocyte-predominant Hodgkin Lymphoma; Treatment Recommendations of the GPOH-HD Study Group. *Klinische Padiatrie* 227:314–321
- McCarville MB (2008) New frontiers in pediatric oncologic imaging. *Cancer Imaging* 8:87–92
- McHugh K, Kao S (2003) Can paediatric radiologists resist RECIST (response evaluation criteria in solid tumours)? *Pediatr Radiol* 33:739–743
- McNitt-Gray MF, Bidaut LM, Armato SG, Meyer CR, Gavrielides MA, Fenimore C, McLennan G, Petrick N, Zhao B, Reeves AP, Beichel R, Kim H-JG, Kinnard L (2009) Computed tomography assessment of response to therapy: tumor volume change measurement, truth data, and error. *Transl Oncol* 2:216–222
- Muenzel D, Engels H-P, Bruegel M, Kehl V, Rummeny EJ, Metz S (2012) Intra- and inter-observer variability in measurement of target lesions: implication on response evaluation according to RECIST 1.1. *Radiol Oncol* 46:8–18
- Mukherji SK, O'Brien SM, Gerstle RJ, Weissler M, Shockley W, Castillo M (1999) Tumor volume: an independent predictor of outcome for laryngeal cancer. *J Comput Assist Tomogr* 23:50–54
- Mukherji SK, Toledano AY, Beldon C, Schmalfuss IM, Cooper JS, Sicks JD, Amdur R, Sailer S, Loevner LA, Kousouboris P, Ang K (2005) Interobserver reliability of computed tomography-derived primary tumor volume measurement in patients with supraglottic carcinoma. *Cancer* 103:2616–2622
- Nawaratne S, Fabiny R, Brien JE, Zalberg J, Cosolo W, Whan A, Morgan DJ (1997) Accuracy of volume measurement using helical CT. *J Comput Assist Tomogr* 21:481–486
- Petrou M, Quint LE, Nan B, Baker LH (2007) Pulmonary nodule volumetric measurement variability as a function of CT slice thickness and nodule morphology. *AJR Am J Roentgenol* 188:306–312
- Prasad SR, Jhaveri KS, Saini S, Hahn PF, Halpern EF, Sumner JE (2002) CT tumor measurement for therapeutic response assessment: comparison of unidimensional, bidimensional, and volumetric techniques initial observations. *Radiology* 225:416–419
- Preda L, Lovati E, Chiesa F, Ansarin M, Cattaneo L, Fasani R, Gandini S, Flor N, Cornalba G, Bellomi M (2007) Measurement by multidetector CT scan of the volume of hypopharyngeal and laryngeal tumours: accuracy and reproducibility. *Eur Radiol* 17:2096–2102
- Prionas ND, Ray S, Boone JM (2010) Volume assessment accuracy in computed tomography: a phantom study. *J Appl Clin Med Phys* 11:3037
- Puesken M, Buerke B, Gerss J, Frisch B, Beyer F, Weckesser M, Seifarth H, Heindel W, Wessling J (2010) Prediction of lymph node manifestations in malignant lymphoma: significant role of volumetric compared with established metric lymph node analysis in multislice computed tomography. *J Comput Assist Tomogr* 34:564–569

- Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B, Kostis WJ (2006) On measuring the change in size of pulmonary nodules. *IEEE Trans Med Imaging* 25:435–450
- Robert Koch-Institut und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (2017) Krebs in Deutschland 2013/2014. https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/krebs_in_deutschland_inhalt.html. Zugegriffen: 10. März 2018
- Sagara Y, Hara AK, Pavlicek W, Silva AC, Paden RG, Wu Q (2010) Abdominal CT: comparison of low-dose CT with adaptive statistical iterative reconstruction and routine-dose CT with filtered back projection in 53 patients. *AJR Am J Roentgenol* 195:713–719
- Schwartz LH, Bogaerts J, Ford R, Shankar L, Therasse P, Gwyther S, Eisenhauer EA (2009) Evaluation of lymph nodes with RECIST 1.1. *European journal of cancer (Oxford, England : 1990)* 45:261–267
- Shah GD, Kesari S, Xu R, Batchelor TT, O'Neill AM, Hochberg FH, Levy B, Bradshaw J, Wen PY (2006) Comparison of linear and volumetric criteria in assessing tumor response in adult high-grade gliomas. *Neuro-oncology* 8:38–46
- Sohaib SA, Turner B, Hanson JA, Farquharson M, Oliver RT, Reznick RH (2000) CT assessment of tumour response to treatment: comparison of linear, cross-sectional and volumetric measures of tumour size. *Br J Radiol* 73:1178–1184
- Sorensen AG, Patel S, Harmath C, Bridges S, Synnott J, Sievers A, Yoon YH, Lee EJ, Yang MC, Lewis RF, Harris GJ, Lev M, Schaefer PW, Buchbinder BR, Barest G, Yamada K, Ponzio J, Kwon HY, Gemmete J, Farkas J, Tievsky AL, Ziegler RB, Salhus MR, Weisskoff R (2001) Comparison of diameter and perimeter methods for tumor volume calculation. *J. Clin. Oncol.* 19:551–557
- Spears CP (1984) Volume doubling measurement of spherical and ellipsoidal tumors. *Med. Pediatr. Oncol.* 12:212–217
- Specht L, Nissen NI (1988) Prognostic factors in Hodgkin's disease stage IV. *European Journal of Haematology* 41:359–367
- Specht L, Nordentoft AM, Cold S, Clausen NT, Nissen NI (1988) Tumor burden as the most important prognostic factor in early stage Hodgkin's disease. Relations to other prognostic factors and implications for choice of treatment. *Cancer* 61:1719–1727
- Suarez F, Mahlaoui N, Canioni D, Andriamanga C, Dubois d'Enghien C, Brousse N, Jais J-P, Fischer A, Hermine O, Stoppa-Lyonnet D (2015) Incidence, presentation, and prognosis of malignancies in ataxia-telangiectasia; A report from the French national registry of primary immune deficiencies. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 33:202–208
- Swerdlow AJ (2003) Epidemiology of Hodgkin's disease and non-Hodgkin's lymphoma. *Eur. J. Nucl. Med. Mol. Imaging* 30 Suppl 1:S3-12

- Tann M, Sopov V, Croitoru S, Nativ O, Moskovitz B, Bar-Meir E, Groshar D (2001) How accurate is helical CT volumetric assessment in renal tumors? *Eur Radiol* 11:1435–1438
- Therasse P, Arbuuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, van Glabbeke M, van Oosterom AT, Christian MC, Gwyther SG (2000) New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 92:205–216
- Tingelhoff K, Moral AI, Kunkel ME, Rilk M, Wagner I, Eichhorn KG, Wahl FM, Bootz F (2007) Comparison between manual and semi-automatic segmentation of nasal cavity and paranasal sinuses from CT images. *Conf Proc IEEE Eng Med Biol Soc* 2007:5505–5508
- Torricelli P, Grimaldi PL, Fiocchi F, Federico M, Romagnoli R (2004) Hodgkin's disease: a quantitative evaluation by computed tomography of tumor burden. *Clin Imaging* 28:239–244
- van der Vorst JR, van Dam RM, van Stiphout RSA, van den Broek MA, Hollander IH, Kessels AGH, Dejong CHC (2010) Virtual liver resection and volumetric analysis of the future liver remnant using open source image processing software. *World J Surg* 34:2426–2433
- Warren KE, Patronas N, Aikin AA, Albert PS, Balis FM (2001) Comparison of one-, two-, and three-dimensional measurements of childhood brain tumors. *J. Natl. Cancer Inst.* 93:1401–1405
- Way TW, Chan H-P, Goodsitt MM, Sahiner B, Hadjiiski LM, Zhou C, Chughtai A (2008) Effect of CT scanning parameters on volumetric measurements of pulmonary nodules by 3D active contour segmentation; A phantom study. *Physics in medicine and biology* 53:1295–1312
- Weissheimer A, Menezes LMd, Sameshima GT, Enciso R, Pham J, Grauer D (2012) Imaging software accuracy for 3-dimensional analysis of the upper airway. *Am J Orthod Dentofacial Orthop* 142:801–813
- Weßling J, Puesken M, Koch R, Kohlhase N, Persigehl T, Mesters R, Heindel W, Buerke B (2012) MSCT Follow-Up in Malignant Lymphoma: Comparison of Manual Linear Measurements with Semi-Automated Lymph Node Analysis for Therapy Response Classification. *Fortschr Röntgenstr* 184:795–804
- Weßling J, Schülke C, Koch R, Kohlhase N, Wassenaar L, Mesters R, Höink AJ, D'Anastasi M, Karpitschka M, Fabel M, Wulff AM, Santos DPd, Kiessling A, Graser A, Bornemann L, Dicken V, Heindel W, Buerke B (2014) Therapy Response Evaluation of Malignant Lymphoma in a Multicenter Study; Comparison of Manual and Semiautomatic Measurements in CT. *Fortschr Röntgenstr* 186:768–779
- Willett CG, Linggood RM, Leong JC, Miketic LM, Stracher MA, Skates SJ, Kushner DC (1988) Stage IA to IIB mediastinal Hodgkin's disease: three-dimensional volumetric assessment of response to treatment. *J. Clin. Oncol.* 6:819–824

- Winer-Muram HT, Jennings SG, Meyer CA, Liang Y, Aisen AM, Tarver RD, McGarry RC (2003) Effect of varying CT section width on volumetric measurement of lung tumors and application of compensatory equations. *Radiology* 229:184–194
- World Health Organization (1979) WHO handbook for reporting results of cancer treatment. World Health Organization.
<http://www.who.int/iris/handle/10665/37200>. Zugegriffen: 22. Februar 2018
- Wormanns D (2005) Radiologische Messverfahren zur Beurteilung des Therapieansprechens solider Tumoren 5:261–272
- Yiallouros M (2016) Hodgkin-Lymphom (Morbus Hodgkin); Krankheitsstadien und Therapielevel. https://www.kinderkrebsinfo.de/erkrankungen/lymphome/pohpatinfomh120060715/diagnostik_und_therapieplanung/therapieplanung/krankheitsstadien/index_ger.html. Zugegriffen: 16. März 2018
- Younes A, Hilden P, Coiffier B et al (2017) International Working Group consensus response evaluation criteria in lymphoma (RECIL 2017). *Annals of oncology: official journal of the European Society for Medical Oncology* 7:1436–1447
- Young LS, Murray PG (2003) Epstein-Barr virus and oncogenesis: from latent genes to tumours. *Oncogene* 22:5108–5121
- Zhao B, James LP, Moskowitz CS, Guo P, Ginsberg MS, Lefkowitz RA, Qin Y, Riely GJ, Kris MG, Schwartz LH (2009) Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non-small cell lung cancer. *Radiology* 252:263–272
- Zubrod CG, Schneiderman M, Frei E, Brindley C, Lennard Gold G, Shnider B, Oviedo R, Gorman J, Jones R, Jonsson U, Colsky J, Chalmers T, Ferguson B, Dederick M, Holland J, Selawry O, Regelson W, Lasagna L, Owens AH (1960) Appraisal of methods for the study of chemotherapy of cancer in man: Comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. *Journal of Chronic Diseases* 11:7–33

8. THESEN

1. Eine große zusammenhängende mediastinale Tumormasse (Bulk) ist ein international weithin anerkannter prognostischer Marker beim Hodgkin-Lymphom im Kindes- und Jugendalter und dient in vielen onkologischen Studienprotokollen zur Risikostratifizierung und Zuordnung der Patienten zu den entsprechenden Therapiegruppen.
2. Basierend auf der Annahme, dass mit Hilfe von dreidimensionalen Segmentierungsverfahren das wahre Tumolvolumen am exaktesten bestimmt werden kann, stellt die Ellipsoid-Methode der EuroNet-PHL-Studie im Vergleich dazu ein akzeptables Surrogat zur Abschätzung des mediastinalen Tumolvolumens dar.
3. Die Cheson-Methode ist im Vergleich zur Ellipsoid-Methode ungenauer im Hinblick auf die Bestimmung des wahren mediastinalen Tumolvolumens.
4. Die auf Grundlage der RECIST-Methode transformierten Volumina überschätzen das mediastinale Lymphomvolumen deutlich, sodass die RECIST-Methode ungeeignet ist, dieses zuverlässig zu bestimmen.
5. Die Annahme, die Tumorform beim Hodgkin-Lymphom mit Beteiligung des Mediastinums entspräche der einer Kugel, ist vereinfachend und bedingt hauptsächlich die erhaltenen Abweichungen der Alternativmethoden von der Referenzmethode.
6. Die prozentualen Abweichungen der Volumenmessergebnisse der Alternativmethoden gegenüber denen der Segmentierungsmethode nehmen mit abnehmender Tumorgröße zu, da schon geringe Messungenauigkeiten zu verhältnismäßig großen prozentualen Fehlern führen und der Anteil kleiner Tumolvolumina in dieser Untersuchung hoch war.
7. Bei der Übereinstimmung hinsichtlich der prozentualen Tumolvolumenreduktion zwischen Referenz- und Alternativmethoden weisen alle Methoden im Durchschnitt ähnliche Ergebnisse auf mit einer leichten Tendenz zur Überschätzung der Volumenreduktion.
8. Die manuelle Segmentierungsmethode weist die beste Reproduzierbarkeit auf, bei jedoch auch validen Ergebnissen der Ellipsoid-Methode, der Cheson-Methode und der RECIST-Methode.
9. Der je nach Komplexität der Lymphome enorme Zweitaufwand für die Volumetrie mittels Segmentierungssoftware und die hohen Anforderungen an die Bildqualität bedingen, dass sich manuelle Segmentierungsmethoden für den Einsatz in einer multizentrischen Studie nicht eignen.
10. Es besteht keine Einigkeit in der Diskussion über die zusätzliche prognostische Bedeutung der Abschätzung des mediastinalen Bulks über das exakte Volumen im Vergleich zur Bestimmung lediglich des größten bzw. der größten orthogonalen Durchmesser.

V. SELBSTSTÄNDIGKEITSERKLÄRUNG

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig und ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Ich versichere, dass ich für die inhaltliche Erstellung der vorliegenden Arbeit nicht die entgeltliche Hilfe von Vermittlungs- und Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen habe. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für die Arbeit erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Leinfelden-Echterdingen, den 20.11.2018

Susanne Taufer

VI. ERKLÄRUNG ÜBER FRÜHERE PROMOTIONSVERSUCHE

Hiermit erkläre ich, dass nur dieser Antrag auf Eröffnung eines Promotionsverfahrens von mir eingereicht wurde und auch keine anderen Promotionsversuche zu einem früheren Zeitpunkt an dieser oder einer anderen Universität erfolgt sind.

Leinfelden-Echterdingen, den 20.11.2018

Susanne Taufer

VII. DANKSAGUNG

Mein Dank gilt Herrn Prof. Dr. Dr. Wohlgemuth für die Möglichkeit in seiner Abteilung, der Universitätsklinik und Poliklinik für Radiologie an der Martin-Luther-Universität Halle-Wittenberg, zu promovieren.

Weiterhin danke ich Herrn Oberarzt Dr. med. Dietrich Stoevesandt für die Überlassung dieses Themas sowie für die fortwährende Unterstützung in der gesamten Promotionsphase. Danke für die vielen inhaltlichen und formalen Anregungen sowie kritischen Kommentare, die zum Gelingen dieser Arbeit wesentlich beigetragen haben.

Herrn Dr. Dirk Hasenclever vom Institut für Medizinische Informatik, Statistik und Epidemiologie der Universität Leipzig danke ich vielmals für die Unterstützung bei der statistischen Auswertung der vorliegenden Untersuchung.

Weiterhin möchte ich mich bei all meinen Freunden für ihre Unterstützung im gesamten Studium und die hilfreichen Anregungen bezüglich dieser Arbeit bedanken.

Mein besonderer Dank gilt meinem Mann Nico und meiner Familie für die unentwegte und tatkräftige Unterstützung und Motivation sowohl während des Studiums als auch während der Anfertigung dieser Dissertation.