

## 5. Diskussion

### 5.1 Volumenquantifizierung solider Tumoren in der Hochkontrastsituation

Ein Hochkontrastunterschied zwischen soliden Tumoren und ihrer Umgebung wird typischerweise bei intrapulmonalen Rundherden beobachtet. Neben Informationen wie der Anzahl, Lage, Morphologie und charakteristischen Veränderungen des umgebenden Parenchyms kann die Größe einen Hinweis zur Klärung der Dignität eines Rundherdes beitragen. In Patientenkollektiven, die hinsichtlich der Entwicklung eines Bronchialkarzinomes ein der Normalbevölkerung vergleichbares Risiko tragen, weisen rund 80% aller benignen intrapulmonalen Rundherde einen Durchmesser von unter 20mm (62) auf. In Screeningstudien konnte jedoch gezeigt werden, dass in vorbelasteten Patientengruppen rund 15% aller malignen Rundherde Durchmesser unter 10mm haben (63). Diese Zahlen demonstrieren, dass eine Größenbestimmung allein nicht einer zuverlässigen Bewertung der Dignität eines Rundherdes zugrunde gelegt werden kann. Die Quantifizierung von Größenveränderungen intrapulmonaler Rundherde kann vielmehr potentiell zur Beurteilung ihrer Dignität herangezogen werden.

Dieses Vorgehen wird in der klinischen Praxis vorrangig bei Rundherden zwischen 5 bis 10mm Durchmesser angewendet (64). Rundherde mit diesen Dimensionen entziehen sich aufgrund ihrer Größe häufig einer erfolgreichen Materialgewinnung zur histologischen Untersuchung, sie haben jedoch im Vergleich mit Rundherden unter 5mm statistisch häufiger eine maligne Genese (65). In Screeninguntersuchungen konnte gezeigt werden, dass bei Auftreten eines Rundherdes mit diesen Dimensionen in einem gefährdeten Patientenkollektiv ohne vorherigen Nachweis eines Malignoms bis zu 23% dieser Rundherde maligner Genese sind (66).

Der Parameter, der im Rahmen einer Verlaufskontrolle in erster Linie zur Einschätzung der Dignität herangezogen wird, ist die Zeit, in der sich das Volumen eines intrapulmonalen Rundherdes verdoppelt. Bei der Mehrzahl der malignen intrapulmonalen Neoplasien werden Tumorverdopplungszeiten zwischen 30 Tagen und 14 Monaten beobachtet. Liegt die Volumenverdopplungszeit unter 30 Tagen kann davon ausgegangen werden, dass ein Rundherd inflammatorischer Genese ist. Bei Rundherden mit Verdopplungszeiten von über 14 Monaten wird eine benigne Genese angenommen. Die typische Verdopplungszeit eines kleinzelligen Bronchialkarzinomes beträgt um 30 Tagen, die eines Adenokarzinoms zwischen 120 und 180 Tagen (65).

Große Studien konnten zeigen, dass die Computertomographie ein sensitives Screeningverfahren bei Patientenkollektiven, die aufgrund einer Häufung von pathogenetisch relevanten Faktoren ein

erhöhtes Risiko tragen ein Bronchialkarzinom zu entwickeln, darstellt (67). Sie trägt dazu bei insbesondere kleine, malignomsuspekte Rundherde in einem potentiell kurablen Stadium zu detektieren. Die Anzahl explorativer thoraxchirurgischer Eingriffe konnte durch den Einsatz eines flächendeckenden CT-Screenings von Risikokollektiven reduziert werden, so dass kein Studienteilnehmer mit einem benignen Rundherd thorakotomiert wurde. Dass ein CT-Niedrigdosis-Screening aufgrund einer die Mortalität pulmonaler Neoplasien zu senken vermag ist anzunehmen, kann jedoch bei derzeitiger Datenlage nicht bewiesen werden.

### **5.1.1. Tumorummetrie versus RECIST**

Die klinischen Konsequenzen, die aus einer Größenbestimmung von Tumoren resultieren sind in grundlegender Weise von der Reliabilität des Messverfahrens abhängig. Als Kriterien zur Beurteilung der Messverfahren wurden in der vorliegenden Untersuchung die Präzision und die Reproduzierbarkeit mehrfach wiederholter Messungen an einem Rundherdphantom bestimmt. Zunächst sollte ein anhand des maximalen Diameters berechnetes Volumen mit einem durch direkte Volumetrie bestimmten verglichen werden.

Die Auswertung idealisierter kugelförmiger Phantome bot den Vorteil, das Volumen aus dem Durchmesser direkt zu berechnen und so einen unmittelbaren Vergleich beider Vorgehensweisen anzustellen. Es konnte nachgewiesen werden, dass eine Volumetrie zu einer präziseren Bestimmung des Phantomvolumens führt als eine Berechnung anhand des maximalen Diameters. Der mittlere Fehler des anhand des Durchmessers bestimmten Volumens lag sowohl in der Hoch- als auch in der Niedrigkontraststudie über dem der Volumetrie.

Unter Annahme einer kugelförmigen Gestalt eines Rundherdes und der Vorstellung, dass dieser ein in alle Raumrichtungen gleichmäßiges Wachstum aufweist, müsste man bei einem 10mm durchmessenden Tumor und einer einfachen Volumenverdopplung lediglich eine Zunahme des Diameters um 2,6mm beobachten. Eine 50%ige Volumenzunahme führt bei Betrachtung derselben Dimensionen zu einer Zunahme des Durchmessers um lediglich 1,4mm.

Das Ergebnis der Phantomstudie und diese Überlegung legen nahe, dass Größenveränderungen in vivo ebenfalls eher durch eine Volumetrie als anhand der Veränderungen der Durchmesser erkannt werden. Unter Umständen führen beide Verfahren zu einer divergenten Bewertung des Wachstumsverhaltens eines intrapulmonalen Rundherdes.

Diese Hypothese wird durch eine Reihe von klinischen Studien gestützt. Einige Studien vergleichen an Patientenkollektiven mit intrapulmonalen Tumoren direkt eine Volumetrie mit dem Vorgehen nach RECIST (68, 69).

Die Gruppe von Marten et al. evaluierte an einem Patientenkollektiv pulmonale Metastasen verschiedener Primärtumoren anhand des maximalen Diameters sowie durch eine softwaregestützte, vollautomatische Segmentation des Volumens. Bei dieser Untersuchung konnte gezeigt werden, dass rund 28% der Patienten durch eine Rundherdvolumetrie anders hinsichtlich des Krankheitsverlaufes bewertet wurden. Es konnten signifikante Unterschiede zwischen der Volumetrie und dem manuell bestimmten Durchmesser bei der Bewertung eines konstanten Krankheitsverlaufes und eines Progresses gezeigt werden (69).

Zu vergleichbaren Ergebnissen kommen Jennings et al. bei einer Verlaufsbeobachtung an Patienten mit Bronchialkarzinomen. In dieser Studie wurden insgesamt 63 Patienten mit einem Bronchialkarzinom, das initial einen Durchmesser unter 3cm aufwies, im Verlauf beobachtet. Ein durch Volumetrie erfasstes Tumorwachstum wurde bei bis zu 37% der Patienten durch die Auswertung des Diameters nicht erkannt. Bei Tumoren mit einem Volumen unter 1.000m<sup>3</sup> waren es sogar bis zu 44% (70).

Die Ergebnisse der klinischen Studien und der Phantomuntersuchung zeigen, dass eine Volumetrie ermöglicht das Wachstumsverhältnis intrapulmonaler Tumoren präziser als der Durchmesser abzubilden. Bei der Beobachtung intrapulmonaler Metastasen ist zu vermuten, dass durch eine Metastasenvolumetrie ein potentielles Wachstum früher detektiert werden kann. Es ist denkbar, dass dies zu einer Änderung des therapeutischen Vorgehens führt.

Bei einer Verlaufsbeobachtung unklarer Rundherde kann aus der Anwendung einer Rundherdvolumetrie möglicherweise eine andere Einschätzung der Dignität des Rundherdes resultieren. Es ist anzunehmen, dass insbesondere kleine Veränderungen einer Bewertung anhand des Durchmessers entgehen und im schlimmsten Falle zu einer Fehlbewertung der Dignität führen.

### **5.1.2. Methoden der Volumenquantifizierung**

Wie eine Reihe aktueller Untersuchungen zeigen, kommt der Methode, die bei der Volumenbestimmung solider Tumoren in einer Hochkontrastumgebung angewendet wird, eine wichtige Rolle zu. Die Ergebnisse der Hochkontrastuntersuchung können im Zusammenhang mit einer Reihe von Studien betrachtet werden, die Volumetrieverfahren hinsichtlich ihrer Präzision an Phantomanordnungen unter Berücksichtigung einer Reihe von Einflussfaktoren evaluierten.

In der vorliegenden Studie ermöglichten die manuelle Segmentierung und die saapixelbasierte Methode die präzisesten Volumenbestimmungen. Bei Rekonstruktion in einem

Lungenfaltungskern betragen die mittleren absoluten Fehler beider Methoden 3,9 % bzw. 4,0%. Der mittlere Fehler der schwellenwertbasierten Methode lag bei vergleichbaren Bedingungen bei 5,6%. Die Wahl der Methode beeinflusste das Ergebnis der Volumenquantifizierung signifikant. Die Gruppe von Yankelevitz et al. bestimmte an einem kugelförmigen Phantom die Präzision einer vollautomatischen, algorithmengestützten Volumetrie-Software zur Bestimmung des Volumens intrapulmonaler Rundherde (68). Bei einer Schichtdicke von 1,0mm betragen in der Einzeilen-Spiral-CT die mittleren Fehler bei der Volumetrie zwischen rund 2,9% und 2,3% in Abhängigkeit von der Größe der Phantome und dem angewendeten Verfahren. Bei etwas kleineren Phantomen wurde dabei ein geringfügig höherer Fehler beobachtet. Die mittleren Fehler differierten bei einer Schichtdicke von 1mm vergleichsweise gering um rund 1,0% bzw. 1,6% von denen mit der saapixelbasierten Methode vorgenommenen Volumenbestimmungen bei einer Schichtdicke von 1,25mm.

Neben der Schichtdicke, die in der vorliegenden Untersuchung gering über der, der von Yankelevitz et al. vorgestellten Studie lag, unterschieden sich die Dimensionen der Phantome: in dieser Studie wiesen die Phantome Durchmesser zwischen 3 und 11mm, in der hier vorgestellten Hochkontrastuntersuchung zwischen 4mm und 30mm auf. Während in der vorliegenden Untersuchung insgesamt 8 verschiedene Größen von Phantomen untersucht wurden, waren dies in der Studie von Yankelevitz zwei. Darüber hinaus bestanden Unterschiede hinsichtlich der technischen Differenzierung, der bei der Volumetrie angewendeten Programme: Es kann angenommen werden, dass ein vollautomatisches Volumetrieprogramm, das Bestandteil einer kommerziell vertriebenen Auswertungssoftware ist, auf speziell adaptierten Algorithmen basiert. Aufgrund dessen ist eine Anwendung eines vergleichsweise optimierten Programmes unter Umständen einem manuell modifizierten, semiautomatischen Verfahren, das nicht explizit zur Tumorvolumetrie vorgesehen war, überlegen. Dies kann unter anderem ursächlich für den im Literaturvergleich geringfügig höheren Fehler der vorliegenden Untersuchung bei der Bestimmung des Phantomvolumens sein.

Neben der Untersuchung streng kugelförmiger Objekte wurde durch Yankelevitz et al. der Einfluss einer Veränderung der äußeren Objektform auf das Ergebnis einer Volumenbestimmung evaluiert. Es konnte gezeigt werden, dass bei einer isovolumetrischen Verformung der kugelförmigen Plastelinephantome eine Zunahme der mittleren Fehler um bis zu 3% resultiert (68).

Eine in diesem Zusammenhang zukunftsweisende Untersuchung stellte die Gruppe von Marten et al. vor (71). Mit einem Experimentalprototypen eines auf einem Flachbilddetektor basierenden Volumencomputertomographen wurde die Messgenauigkeit bei der Bestimmung des Volumens

synthetischer Rundherdphantome evaluiert. Der prozentuale Fehler vor und nach isovolumetrischer Deformierung wurde mit einer vollautomatischen, algorithmengestützten, schwellenwertbasierten Methode bestimmt. Für Phantome mit Volumina zwischen  $5,51\text{mm}^3$  und  $10,10\text{mm}^3$  (Äquivalenzdurchmesser zwischen  $2,41\text{mm}$  und  $2,95\text{mm}$ ) betrug der mittlere prozentuale Fehler rund  $1,5\%$  vor und  $3,5\%$  nach Deformierung. Bei Phantomvolumina zwischen  $14,2\text{cm}^3$  und  $20,1\text{cm}^3$  (Äquivalenzdurchmesser zwischen  $3,31\text{mm}$  und  $3,71\text{mm}$ ) konnten mittlere prozentuale Fehler vor Deformierung von rund  $1,3\%$ , danach von rund  $0,61\%$  beobachtet werden.

Die in dieser Studie angegebenen prozentualen Fehler bei Verwendung eines Flachbilddetektorsystems lagen deutlich unter denen vergleichbarer Untersuchungen mit der Mehrzeilen– Spiral –CT. Bei einer Weiterentwicklung dieser Systeme und einer anschließenden klinischen Anwendung ist zu erwarten, dass eine Bestimmung der Volumenzunahme kleinerer Rundherde innerhalb eines kürzeren Zeitintervalls zwischen den Untersuchungen ermöglicht wird. Die Autoren nennen einen erforderlichen zeitlichen Mindestabstand von lediglich 20 bzw. 43 Tagen zwischen zwei repetitiven Untersuchungen, um einen prozentualen Messfehler auszugleichen. Diese Angaben beziehen sich auf eine Volumetrie von Herden unter  $2\text{mm}$  unter Annahme einer Tumorverdopplungszeit von 177 bzw. 396 Tagen, wie diese bei Bronchialkarziomen in der Literatur angenommen werden (65). Durch Anwendung von Flachbilddetektoren, die aktuell nur im Rahmen von Experimentalstudien zugelassen sind, können möglicherweise früher kleine maligne Rundherde in potentiell kurablen Stadien detektiert werden.

Als weiteren Faktor, der eine Volumenbestimmung intrapulmonaler Tumoren potentiell beeinflussen kann untersuchten Goo et al. den Einfluss des Dichteunterschiedes zwischen einem Rundherd und dem Lungenparenchym an einem anthropomorphen Thoraxphantom (72). Bei dieser Untersuchung konnte nachgewiesen werden, dass bei höheren Dichteunterschieden zwischen dem simulierten Parenchym und dem kugelförmigen Rundherdphantom geringere prozentuale Fehler bei der Bestimmung des Volumens auftraten. Dieses Ergebnis scheint aufgrund einer besseren Abgrenzbarkeit der äußeren Kontur zwischen einem Objekt und seiner Umgebung bei Vorliegen eines höheren Kontrastunterschiedes plausibel. Die mittleren prozentualen Fehler betragen bei einer Schichtdicke von  $1\text{mm}$  und Phantomdurchmessern zwischen rund  $3\text{mm}$  und  $13\text{mm}$  in Abhängigkeit von der Kontrastdifferenz zur Umgebung zwischen  $5,4\%$  und  $16,7\%$ . Sie lagen somit über den mittleren Fehlern der in der vorgestellten Hochkontraststudie angewendeten Volumetrieverfahren.

Das Volumen wurde in der Studie von Goo et al. durch ein vollautomatisches

Segmentationsverfahren quantifiziert, das in seinem methodischen Ansatz mit der saapixelbasierten Methode vergleichbar ist. Nach dem manuellen Setzen einer ROI (region of interest) in das Zentrum einer Läsion und einer weiteren ROI in das umgebende Parenchym wurde das Phantomvolumen durch vollautomatische Segmentation bestimmt. Die äußere Kontur wurde algorithmengestützt dort festgelegt, wo der größte Anstieg innerhalb des Dichtespektrums zwischen den zwei ROIs detektiert wurde.

Die Gruppe um Ko et al. evaluierte ein Volumetrieverfahren mit einer vergleichbaren Methodik. Ziel dieser Untersuchung war unter anderem ein Vergleich mit Verfahren, die anhand von frei wählbaren oder vorgegebenen Schwellenwerten ein Volumen quantifizierten (73). Es konnte gezeigt werden, dass eine Methode, die aufgrund eines Dichtevergleiches zwischen einer Läsion und dem Parenchym das Volumen eines Rundherdes segmentierte, eine präzisere Volumetrie als eine schwellenwertbasierte Methode ermöglichte. Nach Meinung der Autoren ist dieser Unterschied darauf zurückzuführen, dass ausschließlich schwellenwertbasierte Algorithmen eine höhere Empfindlichkeit gegenüber dem Partialvolumeneffekt aufweisen. Dieser tritt insbesondere an der Grenzfläche von Hochkontrastobjekten auf, führt zu einer Fehlbewertung der äußeren Kontur und somit zu einem höheren Fehler bei der Volumenbestimmung.

Diese Aussagen können durch die vorliegende Untersuchung bestätigt werden. Mit der saapixelbasierten Methode, die sich bei der kontinuierlichen Festlegung der Objektkontur an dem höchstmöglichen Dichteunterschied orientierte, konnten präzisere Volumenbestimmungen als mit der schwellenwertbasierten Methode vorgenommen werden. Im Unterschied zu den durch Ko und Goo vorgestellten vollautomatischen Volumetrieverfahren bot die saapixelbasierte Methode die Möglichkeit, die festgelegte Kontur durch den visuellen Eindruck des Untersuchers modifizieren. Die prozentualen Unterschiede der mittleren Fehler zwischen beiden Methoden lagen in der vorliegenden Studie bei Evaluation im Lungenfaltungskern bei 1,6%. In der Gruppe von Ko betragen die mittleren prozentualen Fehler bei soliden Rundphantomen zwischen den Methoden rund 3,3% (73). Die algorithmengestützte Methode, die anhand des Vergleiches zwischen der Läsion und ihrer Umgebung eine Volumenbestimmung durchführte, wies einen mittleren prozentualen Fehler von rund 2,4%, im Vergleich zur schwellenwertbasierten Methode mit 5,7% auf.

### 5.1.3. Reproduzierbarkeit

Neben der Präzision, die nur in einer Phantomuntersuchung quantifiziert werden kann, stellt die Reproduzierbarkeit ein weiteres wichtiges Kriterium bei der Bewertung eines diagnostischen Messverfahrens dar. Ihr Stellenwert bei der Analyse intrapulmonaler Rundherde in vivo ist in einer Reihe von Untersuchungen unterstrichen worden (64, 65, 68, 75). In Patientenstudien wurde gezeigt, dass eine durch Algorithmen gestützte, automatisierte Volumetrie eine bessere Reproduzierbarkeit bei der Bestimmung des Volumens intrapulmonaler Metastasen aufweist als eine Messung des maximalen Durchmessers (64, 68). Die Übereinstimmungen zwischen den Ergebnissen eines Untersuchers und zwischen denen verschiedener Untersucher sind bei mehrfachen Messungen mitunter signifikant höher als bei einer Anwendung der RECIST – Kriterien (75). Auch zweidimensionale Messungen vermögen nach Meinung einiger Autoren die durch eine Volumetrie gewonnenen Aussagen über Größenveränderungen pulmonaler Rundherde nicht adäquat wiederzugeben und werden aus diesem Grunde bei einer Verlaufsbetrachtung von ihnen nicht favorisiert (64).

Bei Durchsicht der Literatur konnte keine weitere Studie gefunden werden, die die Reproduzierbarkeit verschiedener Verfahren zur Rundherdvolumetrie an kugelförmigen Phantomen untersucht und einen direkten Vergleich zwischen einer Bestimmung anhand des maximalen Durchmessers und einer Volumetrie anstellt.

In der vorliegenden Studie konnte gezeigt werden, dass selbst bei der Auswertung eines kugelförmigen Phantomes eine Volumetrie zu einer besser reproduzierbaren Bestimmung des Volumens führt. Das aus dem Durchmesser berechnete Volumen wich in Abhängigkeit von dem zur Rekonstruktion verwendeten Faltungskern bei wiederholten Messungen durchschnittlich zwischen 6,6% und 17,2% von den Mittelwerten der Einzelmessungen ab. Bei vergleichbaren Bedingungen differierten die mit der saapixelbasierten Methode bestimmten Volumina zwischen 4,0% und 10,6%, die durch manuelle Segmentation bestimmten zwischen 3,9% und 10,8%.

Die Gruppe um Bolte bestimmte die Reproduzierbarkeit einer softwaregestützten Volumetrie bei der Bestimmung des Volumens von Rundherdphantomen mit einer irregulären Oberfläche (76). Die aus einem Wachsgemisch bestehenden Phantome wiesen in dieser Studie Volumina zwischen  $20\text{mm}^3$  und  $391\text{mm}^3$  auf. Im Vergleich hierzu lagen die Volumina der Phantome in der vorgestellten Hochkontraststudie zwischen  $34\text{mm}^3$  und  $14.132\text{mm}^3$ . Das umgebende Lungenparenchym wurde durch eine luftdistendierte und in einem Unterdruckbehälter evakuierte Schweinelunge simuliert. Für die Reproduzierbarkeit der Volumenbestimmungen mit einer

vollautomatisierten Auswertesoftware geben die Autoren mittlere Abweichungen von 6,2% bei Auswertung desselben Datensatzes an. Die mittleren Abweichungen bei der Auswertung verschiedener Datensätze desselben Phantomaufbaues lagen bei 9,2%.

Wormanns et al. untersuchten in einer Patientenstudie, inwieweit das Ergebnis einer Volumenbestimmung zweier im Abstand weniger Minuten durchgeführter Niedrigdosis-CT Untersuchungen reproduzierbar ist (74). Hierzu wurden insgesamt 151 pulmonale Metastasen verschiedener Entitäten im Abstand von 10 Minuten untersucht und das Volumen durch eine vollautomatische Auswertungssoftware bestimmt. Die Volumina der evaluierten Metastasen lagen zwischen  $7,3\text{mm}^3$  und  $4833\text{mm}^3$ . Die mittlere prozentuale Abweichung zwischen den Einzeluntersuchungen betrug rund 10,8%. Die Autoren schlussfolgerten aus diesen Ergebnissen, dass eine Anwendung der untersuchten vollautomatischen Volumetrie-Software zu ausreichend reproduzierbaren Ergebnissen führt, um die klinisch relevante Größenveränderungen intrapulmonaler Rundherde zu detektieren.

Unter Berücksichtigung der Unterschiede zwischen den vorgestellten Studien hinsichtlich des Versuchsaufbaues, der Volumetrieverfahren und der Fragestellungen scheinen die Dimensionen der Ergebnisse der vorgestellten Untersuchung für eine Rundherdvolumetrie mit diesen vergleichbar.

#### **5.1.4. Einfluss der Rekonstruktionsalgorithmen**

In der Hochkontrastuntersuchung wurde der Einfluss verschiedener Faltungskerne exemplarisch für den Stellenwert von Bildrekonstruktionsparametern auf eine Tumolvolumetrie untersucht. Die zur Rekonstruktion verwendeten Faltungskerne unterschieden sich dabei im Hinblick auf die Kantenbetonung und bezüglich des resultierenden Signal-Rausch-Verhältnisses.

Bei der Verwendung kantenbetontender Faltungskerne zur Bildrekonstruktion wurden vergleichsweise geringe Differenzen bei der Präzision und Reproduzierbarkeit zwischen den Verfahren beobachtet. Bei Rekonstruktion mit weniger kantenbetonender Faltungskerne, traten deutliche Unterschiede zwischen den verschiedenen Segmentationsverfahren auf: Verfahren, die das Volumen durch Festlegung der Objektkontur unter Berücksichtigung des visuellen Untersuchereindrucks evaluieren - wie die Segmentierung und die saapixelbasierte Methode - zeigen bei weniger kantenbetonender Abbildung eine vergleichsweise höhere Präzision und bessere Wiederholbarkeit bei der Volumenbestimmung.

Die schwellenwertbasierte Methode nahm hinsichtlich der Präzision und Wiederholbarkeit eine



Mittelstellung zwischen dem Vorgehen nach RECIST und den anderen Volumetrieverfahren ein. Dies ist aus der Vorgehensweise bei der Bestimmung des Phantomvolumens mit dieser Methode abzuleiten. Der Untersucher hatte abgesehen von der Festlegung der Schwellenwerte keinen Einfluss auf die Determinierung der Objektkontur. Kam es zu einer unschärferen Abbildung der Grenzfläche zwischen dem Phantom und seiner Umgebung konnte keine weitere Modifikation der bestimmten Kontur aufgrund des visuellen Bildeindrucks vorgenommen werden. Diese im Vergleich zu einer semiautomatischen Volumetrie oder einer Segmentierung geringere Flexibilität bei der Festlegung der Objektkonturen spiegelt sich in einer unpräziseren und weniger reproduzierbaren Bestimmung des Volumens wieder.

Es konnten nachgewiesen werden, dass ein aus dem Durchmesser berechnetes Volumen deutlicher als ein durch Volumetrie bestimmtes von Bildrekonstruktionsparametern abhängig ist. Bei einer Reduktion der Kantenbetonung zeigte ein berechnetes Volumen eine deutlicher ausgeprägte Abnahme der Präzision und Wiederholbarkeit verglichen mit einem durch Volumetrie bestimmten.

Honda et al. untersuchten den Einfluss des Rekonstruktionsalgorithmus, des field of view (FOV) und der Schichtdicke auf die Volumenbestimmung intrapulmonaler Rundherde entzündlicher und neoplastischer Genese (76). Sie werteten die Datensätze von insgesamt 50 Rundherden unter Verwendung eines hoch- und niedrigauflösenden Faltungskernes, verschiedener FOVs zwischen 10 und 30mm Durchmesser und von Schichtdicken zwischen 1,25mm und 5mm aus. Als einziger der betrachteten Parameter zeigte der Rekonstruktionsalgorithmus einen signifikanten Einfluss auf die Bestimmung des Rundherdevolumens. Den Autoren zufolge ergaben sich bei vergleichsweise großen Rundherdevolumina bei Verwendung eines hochauflösenden Faltungskernes und eines größeren FOV die geringsten Abweichungen zwischen den einzelnen Volumina.

Ko et al. bestätigen die Bedeutung des Rekonstruktionsalgorithmus auf das Ergebnis einer Volumenbestimmung in einer Untersuchung an kugelförmigen Phantomen (73). Neben der Phantomgröße, der Phantomdichte und der Methode der Volumetrie beeinflusste auch in dieser Untersuchung die Wahl des Rekonstruktionsalgorithmus das Ergebnis einer Volumenbestimmung signifikant. Angaben über prozentuale Fehler werden in beiden Studien nicht gemacht.

Bildverarbeitungsparameter können unter Berücksichtigung der Literatur und der Ergebnisse der Hochkontraststudie das Ergebnis einer Tumorvolumetrie potentiell beeinflussen. Verglichen mit dem Vorgehen nach RECIST erwies sich eine Volumetrie bei einer Verschlechterung der Abbildungsbedingungen als stabiler.

## 5.2 Tumorvolumetrie in der Niedrigkontrastsituation

### 5.2.1. Volumetrieverfahren

Die vorliegenden Ergebnisse können mit einer Reihe von Literaturdaten verglichen werden, die Volumetrieverfahren an Phantomen in einer Niedrigkontrastumgebung untersuchen. Es existieren derzeit entweder Patientenstudien, die die Beurteilung des Krankheitsverlaufes nach RECIST mit einer Volumetrie vergleichen oder Phantomstudien, die an zumeist kugelförmigen Phantomen verschiedene Volumetrieverfahren evaluieren. Bislang gibt es jedoch keine Untersuchung, die beide Vorgehensweisen direkt in einer Niedrigkontrastsituation miteinander vergleicht.

Mahr et al. evaluierten sechs potentiell zur Bestimmung von Tumorumfängen einsetzbare Algorithmen in einer experimentellen Studie im Hinblick auf ihre klinische Anwendbarkeit bei der Volumetrie von Lebertumoren (25). Als Versuchsaufbau diente eine kontrastmittelperfundierte Rinderleber, in die als Phantome organische Objekte wie eine Pflaume, eine Kartoffel und eine Traube implementiert waren. Das reale Volumen der biologischen Phantome wurde zuvor durch Wasserverdrängung ermittelt. Neben dem absoluten und relativen Fehler bestimmten sie die Zeit, die eine vollständige Bestimmung des Volumens beanspruchte. Die Autoren favorisierten zwei der insgesamt fünf untersuchten Segmentationsverfahren im Hinblick auf ihre Präzision und Segmentationszeit. Der „region-growing“-Algorithmus wies in seiner Vorgehensweise Ähnlichkeiten mit der saapixelbasierten Methode auf. Der Algorithmus determiniert, ausgehend von einer durch den Untersucher gesetzten ROI, das Volumen der Objekte anhand der Dichteunterschiede zu umliegenden Voxeln. Die Grenze einer Läsion wurde dabei dort festgelegt, wo der größte Unterschied im Dichtekontinuum festgestellt wurde. Als zweiten befürworteten die Autoren einen semiautomatischen Delineationsalgorithmus. Die mittleren relativen Fehler dieser Verfahren bei einer Schichtdicke von 2mm lagen dabei um 4,1% bzw. 6,1%. Die saapixelbasierte Methode und die Schichtsegmentierung ermöglichten bei einer Schichtdicke von 2,5mm im Vergleich dazu eine Volumenquantifizierung mit einem mittleren Fehler von 3,3% bzw. 2,5%. Der mittlere relative Fehler der schwellenwertbasierten Methode lag bei 5,3%, der des Vorgehens nach RECIST bei 8,3%.

Die Arbeitsgruppe von Disler et al. untersuchte die Präzision einer dreidimensionalen Volumetrie mit der CT und MRT. Sie arbeiteten dabei mit einer Versuchsanordnung, die der Niedrigkontraststudie vergleichbar war (23). In einer wassergefüllten Wanne wurden

kontrastmittelgefüllte Phantome unterschiedlicher Dichte untersucht. Die Auswirkungen des FOV und der Schichtdicke auf die Volumenbestimmung wurden bestimmt. Zur Volumetrie wurden die Schichtsegmentierung und eine schwellenwertbasierte Methode verwendet. Wie in der Niedrigkontrastuntersuchung bestanden zwischen den beiden Methoden keine signifikanten Unterschiede bei der Volumenbestimmung. Die Wahl der Größe des FOV beeinflusste das Volumetrieergebnis nicht signifikant. Die mittleren Fehler betragen in Abhängigkeit von der Methode und der Gestalt der Phantome zwischen 5,6% und 24,8%. Durch mehrfache Messung eines Phantoms bei Schichtdicken zwischen 1,5 und 10mm konnte gezeigt werden, dass bei einer Abnahme der Schichtdicke ein geringerer absoluter Fehler bei der Volumenbestimmung resultierte. Diese Ergebnisse können für die saapixelbasierte Methode und die Schichtsegmentierung in der vorliegenden Untersuchung bestätigt werden.

Yim et al. analysierten neben der Präzision und Reproduzierbarkeit den Zeitaufwand einer semiautomatischen Segmentation im Vergleich zu einer manuellen Segmentierung an einem Phantommodell und an Patienten mit Lebermetastasen kolorektaler Karzinome (27). Der Vergleich einer algorithmengestützten Segmentation mit einer manuellen Segmentierung zeigten hinsichtlich der Präzision signifikante Unterschiede. Bei Phantomen, die ein Volumen von rund 63cm<sup>3</sup> aufwiesen betragen die mittleren Fehler bei der Volumenbestimmung rund 9,2% bei einer algorithmengestützten und 11,1% bei einer manuellen Segmentation. Die algorithmengestützte Volumetrie führte nicht nur zu präziseren Ergebnissen bei der Volumenbestimmung, sondern sie zeigte darüber hinaus eine geringere Variabilität zwischen verschiedenen Untersuchern. In der vorliegenden Niedrigkontraststudie waren die semiautomatische Methode und die Schichtsegmentierung hinsichtlich der Präzision und Reproduzierbarkeit vergleichbar.

Die mittlere Zeit, die zu einer algorithmengestützten Volumenbestimmung benötigt wurde, lag mit durchschnittlich 105s unter der einer Segmentierung. Beide Verfahren unterschieden sich nicht signifikant im Hinblick auf die Segmentationszeit ( $p=0,08$ ). Mahr et al. bestimmten ebenfalls die Segmentationszeit von verschiedenen Volumetriealgorithmen (24). Die zwei hier näher vorgestellten Algorithmen ermöglichten mit mittleren Segmentationszeiten zwischen 32 und 33 Sekunden eine Volumenbestimmung in einem zeitlich akzeptablen Rahmen.

In der vorliegenden Studie wurde die Zeit, die bei einer Volumenbestimmung benötigt wurde, nicht quantifiziert, dennoch bestanden bei der Anwendung deutliche Unterschiede zwischen den Verfahren. Das Vorgehen nach den RECIST-Kriterien erforderte, gefolgt von der schwellenwertbasierten Methode den geringsten Zeitaufwand. Die manuelle Segmentierung erwies sich von den Verfahren mit Abstand als am zeitaufwändigsten. Die saapixelbasierte Methode lag hinsichtlich des Zeitaufwandes zumeist deutlich unter dem einer Segmentierung,

jedoch über dem einer Anwendung der RECIST-Kriterien. Mit der saapixelbasierten Methode ist es möglich mit einem deutlichen geringeren Zeitaufwand und einer vergleichbaren Präzision und Reproduzierbarkeit wie mit der manuellen Segmentierung das Tumolvolumen zu bestimmen. Eine zukünftige Schlüsselrolle wird der Optimierung des Segmentationsprozesses von Tumoren zukommen. Die Entwicklung leistungsfähiger und ebenso schneller Verfahren zur Bestimmung des Tumolvolumens wird dazu beitragen, die klinische Akzeptanz einer Volumetrie solider Tumoren zu erhöhen.

### **5.2.2. Einfluss der Schichtdicke**

In der Phantomuntersuchung wurde nachgewiesen, dass eine Bestimmung des Tumolvolumens von der Auswertung eines dünnere Schichtigen Datensatzes profitiert. Während bei der Durchsicht der Literatur keine weiteren Daten gefunden wurden, die diese Erkenntnis für solide Tumoren in einer Niedrigkontrastumgebung bestätigen, existieren eine Reihe von Studien, die dies für Tumoren in Hochkontrastumgebungen nachweisen. Die Studien stützen sich dabei auf Phantom- und Patientenuntersuchungen.

Winer–Muran et al. zeigten, dass Volumina von Rundherdephantomen zwischen 11% und 27% bei Auswertung derselben Datensätze überschätzt wurden, wenn dieser in einer größeren Schichtdicke rekonstruiert wurden (78). Der prozentuale Fehler bei der Bestimmung des Phantomvolumens war direkt mit der Schichtdicke korreliert und verhielt sich invers zur Größe der Phantome. Die Gruppe von Goo wies ebenfalls nach, dass bei einer Zunahme der Schichtdicke der mittlere prozentuale Fehler bei der Bestimmung des Rundherdvolumens zunahm (72). Für ein kugelförmiges Phantom mit einem Durchmesser von rund 13mm gaben die Autoren eine Zunahme des mittleren Fehlers von 7,0% auf rund 13,6% bei einer Erhöhung der Schichtdicke von 0,75mm auf 5,0mm an. Der mittlere prozentuale Fehler der saapixelbasierten Methode bei der Bestimmung des Phantomvolumens stieg von 1,4% auf 6,5%, wenn die Schichtdicke von 0,625mm auf 3,6mm zunahm.

In einer Querschnittsstudie an intrapulmonalen Metastasen wiesen Zhao et al. nach, dass bei einer Schichtdicke von 3,75mm im Vergleich zu 7,5mm durch eine softwaregestützte Volumetrie im Mittel ein höheres Volumen der Metastasen bestimmt wurde (79). Die Ergebnisse der Volumenbestimmung bei Verwendung derselben Methode unterschieden sich dabei signifikant. Die Autoren folgern daraus, dass die Schichtdicke ein entscheidender Einflussfaktor bei der Verlaufsbewertung intrapulmonaler Filae ist.

In der vorliegenden Arbeit führte eine Reduktion der Schichtdicke in der Niedrigkontrastuntersuchung nur bei der semiautomatischen Methode und der Schichtsegmentierung zu einer präziseren und besser reproduzierbaren Volumenbestimmung der Phantome. Die schwellenwertbasierte Methode und das Vorgehen nach den RECIST-Kriterien zeigen bei einer geringeren Schichtdicken vergleichsweise unpräzisere und schlechter reproduzierbare Volumetrieergebnisse. Diese Unterschiede sind auf ein Zusammenwirken des Segmentationsprozesses und der Abbildungsbedingungen zurückzuführen. Die Vorteile der Untersuchung eines dünn-schichtigen Datensatzes liegen in einer höheren Detailauflösung und einer Reduktion des Partialvolumeneffektes. Gleichzeitig kommt es bei einer Reduktion der Schichtdicke zu einer Zunahme des Bildrauschens (80). Theoretisch resultiert aus einer Halbierung der Schichtdicke eine Zunahme des Signal-Rausch-Verhältnisses um den Faktor  $\sqrt{2}$  und eine Erhöhung der Bilddetailauflösung um den Faktor 2. Bei einer geringeren Schichtdicken überwiegt eine Erhöhung der Ortsauflösung rechnerisch die Zunahme des Bildrauschens. Verfahren, die das Volumen durch eine Konturbestimmung der gesamten Läsion quantifizieren, profitieren von der höheren Detailauflösung und der Reduktion des Partialvolumeneffektes. Die schwellenwertbasierte Methode und das Vorgehen nach RECIST sind bei der Volumenbestimmung vielmehr vom Bildeindruck einer singulären Schicht abhängig. Nach der Festlegung der Schwellenwerte bzw. der Bestimmung des Durchmessers ließen sie darüber hinaus keine manuelle Modifikation der Objektkontur zu. Trotz eines relativ homogenen Objektes scheinen bei beiden Verfahren die Nachteile eines von einem höheren Bildrauschen geprägten Bildeindruckes die Vorteile einer höheren Detailauflösung bei einer geringeren Schichtdicke zu überwiegen, so dass ein höherer Fehler und eine geringe Reproduzierbarkeit der Einzelmessungen zu resultieren.

Bei einer Reihe onkologischer Fragestellungen ist in einer Niedrigkontrastsituation eine dünn-schichtige Rekonstruktion mit einem Gewinn an diagnostischer Information verbunden. Bei der Charakterisierung fokaler Leberläsionen führt die Evaluation eines dünn-schichtigeren Datensatzes aufgrund einer detailgenaueren Darstellung zu einer besseren Detektierbarkeit und Abgrenzbarkeit der Kontur der Läsionen (80). Die Überlegenheit einer dünn-schichtigen Rekonstruktion konnte darüber hinaus bei der Beurteilung von Lebermetastasen (81, 82), Pankreaskarzinomen (82), Peritonealmetastasen (83) und Kolonkarzinomen (84, 85) nachgewiesen werden.

Catalano et al. verglichen bei der Ausbreitungsdiagnostik des Pankreaskarzinomes die Auswertung von Untersuchungen bei Schichtdicken von 1mm und 5mm (82). Die 1-mm-Schichten ermöglichen dabei neben einem zuverlässigeren Staging vor allem eine sicherere

Detektion von Lebermetastasen (83%% im Vergleich zu 91%). Pannu et al. berichten von einer Reduktion falsch negativer Ergebnisse bei der Detektion von Peritonealmetastasen gynäkologischer Entitäten unter Verwendung eines vergleichsweise dünnere Datensatzes sowie der Anfertigung multiplanarer Rekonstruktionen (83).

Wie bei Durchsicht der Literatur und der Auswertung der vorliegenden Studie gezeigt werden konnte, bietet eine Untersuchung von Niedrigkontrastobjekten bei einer geringeren Schichtdicke eine Reihe verschiedener Vorteile. Das Arbeiten mit dünnere Datensätzen stellt jedoch höhere Anforderungen an Datenverarbeitungssysteme und erfordert einen größeren Zeitaufwand von Seiten des Untersuchers. Ziel zukünftiger Anwendungen muss es sein, den Gewinn an diagnostischer Information in einem zeitlich praktikablen Rahmen nutzbar zu machen.

### **5.3. Limitationen der Phantomuntersuchung**

Einige Limitationen der vorgestellten Phantomstudie müssen bei dem Vergleich mit anderen Untersuchungen und bei der Übertragung der Ergebnisse auf die Situation solider Tumoren in vivo berücksichtigt werden.

Wir verwendeten in dieser Studie ein idealisiertes Phantom, das weder die heterogene histologische Zusammensetzung von Lungenrundherde und Lebermetastasen, noch atem- und pulsationsbedingte Artefakte oder das durch den Weichteilmantel bedingte Bildrauschen als potentielle Einflussparameter berücksichtigte.

Die Gestalt der Phantome entsprach einer dem Untersucher bekannten geometrischen Figur, die es im Verlauf der Auswertung repetitiv zu segmentieren galt. Es ist anzunehmen, dass dieser Vorgang aufgrund eines Lerneffektes und der vergleichsweise regulären Kontur leichter durchzuführen war als bei Tumoren, die häufig eine irreguläre Kontur aufweisen.

Unter Berücksichtigung dieser Umstände ist zu erwarten, dass die Unterschiede zwischen den Verfahren hinsichtlich der Präzision bei der Volumetrie solider Tumoren in vivo deutlicher ausgeprägt sind. Es ist anzunehmen, dass diese Faktoren insbesondere die Volumetrie kleinster Läsionen unter physiologischen Bedingungen erschweren.

Die Größe der Phantome lag mit Durchmessern zwischen 4mm und 30mm nur partiell innerhalb der Dimensionen, bei denen eine Verlaufskontrolle dignitätsunklarer intrapulmonaler Rundherde mit einer Volumetrie durchgeführt wird. Eine Verlaufskontrolle kommt in der klinischen Anwendung meist bei Rundherden zwischen 5 und 10mm zur Anwendung (64). Größere Rundherde werden nach einer Punktion zumeist einer histologischen Untersuchung zugeführt.

Die simplifizierte Darstellung der Tumoren durch kugelförmige Phantome bot den Vorteil, anhand der mathematischen Beziehung zwischen Diameter und Volumen, einen direkten Vergleich eines nach RECIST errechneten Volumens mit einem durch Volumetrie bestimmten anzustellen. Selbstverständlich weisen in vivo die wenigsten Tumoren eine kugelförmige Gestalt auf oder zeigen ein in alle gleich Richtungen ausgeprägtes Wachstumsverhalten. Eine Reihe von Studien untersuchte den Einfluss einer isovolumetrischen Deformierung eines Phantomes auf das Ergebnis der Volumenbestimmung (68, 71). Sie kamen zu dem Ergebnis, dass die Volumenbestimmung eines Phantoms mit einer irregulären Gestalt meist mit einem höheren Fehler behaftet war. Es ist anzunehmen, dass dies auch für die Anwendung der hier untersuchten Volumetrieverfahren bei soliden Tumoren in vivo zutrifft.

Eine weitere Einschränkung der Studie betrifft die schwellenwertbasierte Methode. Da eine Volumenbestimmung ausschließlich durch manuell eingestellte Schwellenwert erfolgte und abgesehen davon keine weitere Modifikation bei der Bestimmung der Objektkonturen vorgenommen wurde, ist diese Methode nicht ohne Einschränkungen auf die Situation solider Tumoren übertragbar. Es ist wahrscheinlich, dass bei Tumoren, die weniger hohe Dichteunterschiede zu den umgebenden Strukturen aufweisen, ein höherer Fehler bei der Volumenbestimmung im Vergleich zur Phantomuntersuchung auftritt.

In der Mehrzahl der publizierten Studien wurden bei der Auswertung vollautomatische Algorithmen zur Volumenbestimmung von Lungenrundherden eingesetzt (73–75). Diese waren Bestandteil einer speziell zu diesem Zweck entwickelten Auswertesoftware und zeichneten sich zumeist durch einen geringeren Untersuchereinfluss im Vergleich zu den in der Phantomstudie angewendeten Verfahren aus. Ein Vergleich mit den Literaturdaten sollte daher in erster Linie im Hinblick auf die grundsätzlichen Vorgehensweisen bei einer Volumenbestimmung angestellt werden. Komplexere Segmentationsalgorithmen zur Volumetrie von Tumoren befinden sich derzeit vielfach im Entwicklungsstadium. Eine zukünftige Aufgabe wird darin bestehen, diese zu optimieren und in weiteren experimentellen und klinischen Untersuchungen zu validieren.

#### **5.4. Vergleichende Bewertung der Patientenstudie**

Der Stellenwert einer Volumetrie bei der Verlaufsbeobachtung einer systemischen Therapie wird in der Literatur unterschiedlich bewertet (17, 18, 28-32). Einige Autoren sehen keinen klinischen Vorteil in der Anwendung einer Tumorummetrie im Vergleich zu einer ein- oder zweidimensionalen Messung des Tumors (18, 28). Andere befürworten eine Evaluation des

Tumorvolumens bei der Verlaufsbeobachtung einer systemischen Therapie (17, 29, 30).

Hopper et al. zeigten in einer Studien an rund 130 Patienten mit verschiedenen Entitäten, dass bei der Evaluation von Tumormanifestationen durch Bestimmungen der maximalen Durchmesser in den Raumebenen eine bemerkenswerte Varianz zwischen verschiedenen Untersuchern besteht (29). Zunächst stratifizierte die Gruppe die Tumoren bei der Auswertung hinsichtlich ihrer Form und Begrenzung. Konnten die Untersucher die Referenzläsionen frei wählen, so wurden runde, gut abgrenzbare Läsionen vor irregulären und unscharf begrenzten bevorzugt. Diese Präferenz spiegelt sich auch bei der Bestimmung der Durchmesser wieder. Für irregulär konturierte Tumoren wurden Variabilitäten bei der Bestimmung des Durchmessers von bis zu 16% zwischen verschiedenen Untersuchern bei der gleichen Läsion beobachtet. Bei gut abgrenzbaren Läsionen betrug diese Variabilitäten lediglich bis zu 11 %.

Zu einem vergleichbaren Resultat kommt die Gruppe von van Hoe et al. Sie wiesen nach, dass bei Mehrfachmessungen von Lebermetastasen eine Bestimmung des Tumorvolumens zu geringeren Abweichungen als eine Auswertung des maximalen Durchmessers führt (86). Die Reproduzierbarkeiten der Messungen wurden durch die Größe der Läsion signifikant beeinflusst. Größere Läsionen zeigten in dieser Studie besser reproduzierbare Ergebnisse bei der Volumenbestimmung.

Die Ergebnisse beider Studien weisen auf klinisch potentiell relevante Unterschiede bei der Bestimmung des Durchmessers in Abhängigkeit von der geometrischen Konfiguration und der Größe von Tumoren hin. Ein Vorteil der Volumetrie besteht in der vollständigen Erfassung der Ausbreitung von unregelmäßig begrenzten Tumoren. Die Bestimmung des Tumorvolumens wird dabei weniger als der Durchmesser durch die Morphologie einer Läsion beeinflusst, was in einer besser reproduzierbaren Volumenbestimmung resultiert. Zu diesem Ergebnis kommt die Analyse der Reproduzierbarkeiten der Volumina von Lebermetastasen in der vorgestellten Patientenstudie. Es konnte gezeigt werden, dass ein durch Volumetrie bestimmtes Metastasenvolumens besser reproduzierbar ist als ein aus dem maximalen Durchmesser errechnetes. Die Reproduzierbarkeiten einer Volumetrie und einer Volumenbestimmungen anhand des Durchmessers wiesen signifikante Unterschiede auf. Die Reproduzierbarkeiten der saapixelbasierten Methode und der Schichtsegmentierung zeigten, wie auch in der Phantomuntersuchung, geringere Standardabweichungen als die der schwellenwertbasierten Methode.

Inwieweit die Unterschiede zwischen einem Vorgehen nach RECIST und einer Tumorvolumetrie zu klinisch relevanten Unterschieden bei der Beurteilung einer systemischen Therapie führen, wird in der Literatur unterschiedlich bewertet.



Die Arbeitsgruppe von Dachmann et al. befürwortet eine Bewertung des Wachstumsverhaltens von Lebermetastasen kolorektaler Karzinome anhand des Durchmessers (28). Sie weisen eine hohe Korrelation ( $r: 0,86$ ) zwischen einem durch Segmentierung bestimmten Volumen und dem Durchmesser nach. Die Autoren schlussfolgern daraus, dass der Durchmesser das Wachstumsverhalten der Metastasen zuverlässig repräsentiert. Nach ihrer Meinung stellt eine Metastasenvolumetrie eine zeitaufwendige Prozedur dar, die zu keinem anderen Ergebnis bei der Bewertung des Krankheitsverlaufes führt.

Neben der geringen Anzahl von insgesamt 9 Patienten mit 69 Metastasen, stellt das methodische Vorgehen bei der Auswertung eine Limitation dieser Studie dar. Der Nachweis einer guten Korrelation zwischen dem Volumen und dem Durchmesser eines Tumors erlaubt nur eingeschränkt Rückschlüsse darüber, inwieweit beide Verfahren komplementär bei der Bewertung des Krankheitsverlaufes sind. Eine Beurteilung der Verfahren hinsichtlich ihrer Eignung zur Therapieevaluation kann aus diesen Daten nicht abgeleitet werden. Bei der Bewertung von Verfahren zur Quantifizierung von Tumorumfängen ist in einer Reihe von Publikationen darauf hingewiesen worden, dass vielmehr eine gute Reproduzierbarkeit als eine hohe Korrelation ein Kriterium zur Bewertung der Validität eines Verfahrens darstellt (27, 75, 86). Trotz gut korrelierender Werte zeigt eine Analyse der Reproduzierbarkeiten mitunter deutliche Unterschiede zwischen verschiedenen Verfahren bei klinischen Messungen in vivo (87, 88).

Eine weitere Studie, die die Anwendung einer Tumorumfängenmetrie in der klinische Praxis eher skeptisch bewertet, ist die von Sohaib et al. an insgesamt 16 Patienten mit Lymphomen und Keimzelltumoren durchgeführte Untersuchung. Die Anwendung der RECIST- und WHO-Kriterien wurde hier mit einer Volumetrie bei der Verlaufsbeurteilung unter systemischer Therapie verglichen (18). Es konnte bei diesem Patientenkollektiv eine 90%ige Übereinstimmung zwischen einer Klassifikation nach den RECIST - und den WHO - Kriterien nachgewiesen werden. Bei allen Patienten stimmte die Klassifikation des Krankheitsverlaufes zwischen den WHO – Kriterien und dem durch manuelle Segmentierung ermittelten Volumen überein. Die evaluierten Patienten wiesen mehrheitlich einen konstanten Krankheitsverlauf auf, keiner von ihnen zeigte einen Krankheitsprogress.

Die Autoren räumen ein, dass die hohe Übereinstimmung durch Besonderheiten der untersuchten Entitäten bedingt sein kann. Sohaib et al. stellen die These auf, dass Lymphome und Keimzelltumoren eher zu umschriebenen, nodulären Tumormanifestationen neigen. Unter der Voraussetzung, dass diese ein gleichmäßiges Wachstumsverhalten entlang der Raumrichtungen aufweisen, leiten sie ab, dass Unterschiede zwischen einer Volumetrie und einer

Verlaufskontrolle aufgrund des Durchmessers bei diesen Entitäten im Vergleich zu anderen möglicherweise geringer ausgeprägt sind. Um potentielle Unterschiede aufzuzeigen, müsste nach Meinung der Autoren eine größere Anzahl von Patienten in Studie untersucht werden.

In einer Querschnittsstudie an insgesamt 140 Patienten mit verschiedenen Tumorentitäten evaluieren Hopper et al. das durch Segmentierung ermittelte Tumolvolumen und die maximale Querschnittsfläche der Tumoren (29). Anhand beider Parameter bestimmten sie das Ansprechen auf die Therapie. Bei der Hälfte der Patienten wurden im Verlauf neue Tumorfoci detektiert, so dass beiden Methoden zufolge ein Krankheitsprogress vorlag. Bei den übrigen Patienten unterschied sich die Bewertung des Krankheitsverlaufes in rund 37% der Patienten zwischen beiden Methoden. Bei rund 25% der reklassifizierten Patienten wurde durch Evaluation des Volumens ein Krankheitsprogress anstatt eines konstanten Krankheitsverlaufes detektiert. In dem Patientenkollektiv wurden signifikante Unterschiede bei der Therapiebewertung anhand einer Volumetrie und der maximalen Querschnittsfläche nachgewiesen (29).

Die Arbeitsgruppe von Prasad et al. wies bei der Analyse der Lebermetastasen eines Kollektives von Mammakarzinompatientinnen, die in einer Studie zum Wirksamkeitsvergleich einer Monotherapie mit einem Taxel und einer Kombination von Capecitabine und einem Taxel teilnahmen ebenfalls Unterschiede bei der Klassifikation des Therapieverlaufes nach (17). Ein Vorgehen nach den RECIST - und den WHO - Kriterien führte im Vergleich zu einer Volumetrie in 32% bzw. 34% der Patientinnen zu einer anderen Bewertung des Krankheitsstatus. In der Hälfte der reklassifizierten Patienten wurde der Therapieerfolg durch eine Volumetrie besser beurteilt, in der anderen Hälfte schlechter als bei einer Anwendung der RECIST - Kriterien.

Beide Studien kommen in zwischen 37% und 23% der Patienten durch eine Tumolvolumetrie zu einer unterschiedlichen Bewertung des Krankheitsverlaufes als durch ein- oder zweidimensionalen Messungen. In der vorgestellten Verlaufsuntersuchung lag der Anteil der divergent zu RECIST klassifizierten Patienten in Abhängigkeit vom jeweiligen Segmentationsverfahren zwischen 23% und 17%. Verglichen mit anderen Studien unterschieden sich die Tendenzen bei der Bewertung des Krankheitsverlaufes. In der vorgestellten Untersuchung führte eine Volumetrie der Metastasen bei allen unterschiedlich klassifizierten Patienten zur Detektion eines Krankheitsprogresses anstatt eines konstanten Krankheitsverlaufes. Hopper et al. wiesen ebenfalls bei der Mehrzahl der reklassifizierten Patienten einen Progress nach (29). Bei der Untersuchung von Prasad führt eine Volumenbestimmung zwar zu einer anderen Klassifikation, zeigte jedoch keine eindeutige Tendenz (31).

Eine frühe Detektion eines Tumorprogresses hat möglicherweise Einfluss auf das klinische Vorgehen bei einer onkologischen Therapie. Bei einem Ausbleiben des Ansprechens kann ohne

Zeitverlust eine andere Therapie eingeleitet werden, die unter Umständen zu einem Erfolg führt. Es ist denkbar, dass aus der Kombination eines Therapieversagens und der Option eines frühzeitigen Therapiewechsels eine Verbesserung der Prognose onkologischer Patienten resultiert. Inwieweit eine Verlaufsbeobachtung des Tumorumfanges dabei einen Einfluss auf klinische Endpunkte wie das rezidivfreie Zeitintervall oder das mittlere Überleben des Patienten hat, wird in prospektiven, klinischen Studien an großen Patientenkollektiven evaluiert werden müssen.

Der grundlegende Vorteil einer Auswertung des Tumorumfanges liegt in der Möglichkeit einen soliden Tumor vollständig zu erfassen und seine Beurteilung unabhängig von der Ausbreitung in einer singulären Schicht durchzuführen. Häufig angeführte Nachteile einer Volumetrie sind der hohe Zeit- und Untersucheraufwand verglichen mit einer eindimensionalen Messung. Der Entwicklung speziell adaptierter und optimierter Verfahren wird in Zukunft eine entscheidende Rolle bei der klinischen Verbreitung einer Tumorumfangmetrie zukommen. Für die Volumetrie intrapulmonaler Rundherde existieren bereits eine Reihe von Verfahren, die derzeit eine zuverlässige Volumenbestimmung in einem akzeptablen Zeitrahmen ermöglichen. Für die Volumenbestimmung von Tumoren in einer Niedrigkontrastumgebung wird derzeit in Experimentalstudien über einzelne Verfahren zur Volumenbestimmung von Lebermetastasen berichtet (27).

Ein Aspekt, bei dem Unterschiede zwischen einer eindimensionalen Messung und einer Bestimmung des Volumens von Tumoren möglicherweise zum Tragen kommen, ist die Bewertung neuer onkologischer Therapiekonzepte im Rahmen der klinischen Testung (10, 89). Über 80% der in klinischer Erprobung befindlichen Medikamente erhalten derzeit keine Zulassung am Markt (10). Die meisten Neuentwicklungen scheitern dabei an den Anforderungen der Phase-III, bei der im Vergleich mit der Standardtherapie die Überlegenheit eines neuen Therapeutikums im Hinblick auf die Wirksamkeit nachgewiesen werden muss (89). Die Grundlage der quantitativen Bewertung der Wirksamkeitsanalyse bilden zumeist die metrischen Analysen nach den WHO- oder den RECIST-Kriterien.

Angesichts einer fortschreitenden Differenzierung der Ansatzpunkte systemischer Therapien sind die resultierenden Unterschiede zwischen einzelnen Therapiergebnissen unter Umständen geringer. Unter der Voraussetzung, dass eine Beobachtung des Tumorumfanges zu einer valideren Bewertung im Vergleich zu etablierten Konzepten der onkologischen Therapieevaluation führt, könnte dies zu einer differenzierteren Betrachtung neuer Therapien beitragen. Möglicherweise können klinisch relevante Unterschiede, die einer Bewertung durch eindimensionale Verfahren verborgen geblieben wären, durch eine Beurteilung des

Tumorvolumens aufgezeigt werden. Die sich daraus ableitenden Konsequenzen für die klinische Medizin, die pharmakologische Forschung und den Arzneimittelmarkt sind weiterreichend und werden in Zukunft näher betrachtet werden müssen.