

Medizinische Fakultät  
der  
Universität Duisburg–Essen

Aus dem Fachbereich angewandte Naturwissenschaften und Elektrotechnik  
Westfälische Hochschule Gelsenkirchen Bocholt Recklinghausen

Eignung der Niedrigdosis–Computertomographie  
mit modellbasierter iterativer Rekonstruktion für das Lungenkrebs–Screening.  
– Eine quantitative Phantomstudie –

I n a u g u r a l – D i s s e r t a t i o n  
zur  
Erlangung des Doktorgrades der Naturwissenschaften in der Medizin  
durch die Medizinische Fakultät  
der Universität Duisburg–Essen

Vorgelegt von  
Britta König  
aus Recklinghausen  
2021

# DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT  
DUISBURG  
ESSEN

*Offen im Denken*

ub | universitäts  
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

**DOI:** 10.17185/duepublico/76529

**URN:** urn:nbn:de:hbz:465-20220914-123359-7

Alle Rechte vorbehalten.

Dekan: Herr Univ.–Prof. Dr. med. J. Buer  
1. Gutachter: Herr Priv.–Doz. Dr. med. H. Kühl  
2. Gutachter: Herr Prof. Dr. med. Ch. Pöttgen  
3. Gutachter: Herr Prof. Dr. med. K. Darwiche

Tag der mündlichen Prüfung: 28. Juni 2022

## Publikationen

Stand: 31.10.2021

---

- 2021** König, B., Kühl, H., Zylka, W. (2021): CT lung–cancer screening: does iterative reconstruction matter? Shanghai Chest.  
<https://dx.doi.org/10.21037/shc-21-9>  
Journal–Artikel.
- 2020** König, B., Guberina, N., Kühl, H., Zylka, W. (2020): Validation of iterative CT reconstruction by inter and intra observer performance assessment of artificial lung foci. Curr. Dir. Biomed. Eng. 6, 534–537.  
<https://doi.org/10.1515/cdbme-2020-3137>  
Journal–Artikel.
- 2020** König, B., Guberina, N., Kühl, H., Zylka, W. (2020): Validation of iterative CT reconstruction by inter and intra observer performance assessment of artificial lung foci. Biomed. Eng.–Biomed. Tech. 65(s1), 249.  
<https://doi.org/10.1515/bmt-2020-6043>  
Abstract, Poster, 54th DGBMT Annual Conference on Biomedical Engineering, Leipzig.
- 2020** König, B., Gernhardt, C., Guberina, N., Kühl, H., Forsting, M., Zylka, W. (2020): Suitability of iterative reconstruction for lung–cancer screening with low–dose computer tomography: First results of a phantom study.  
<https://dx.doi.org/10.26044/ecr2020/C-02150>  
Abstract, Poster, European Society of Radiology, Wien.
- 2020** König, B., Guberina, N., Kühl, H., Zylka, W. (2020): Eignet sich die statistische modellbasierte iterative Rekonstruktion mit low–dose Protokollen der Computertomographie für Lungenkrebs–Screening? Erste Ergebnisse einer Phantomstudie. Rofo. 192(s01), 110.  
<https://doi.org/10.1055/s-0040-1703452>  
Abstract, e–Poster, Deutscher Röntgenkongress, Leipzig.

- 2019** König, B., Guberina, N., Kühl, H., Zylka, W. (2019): Design and first results of a phantom study on the suitability of iterative reconstruction for lung-cancer screening with low-dose computer tomography. *Curr. Dir. Biomed. Eng.* 5, 593–596.  
<https://doi.org/10.1515/cdbme-2019-0149>  
Journal-Artikel.
- 2019** König, B., Guberina, N., Kühl, H., Zylka, W. (2019): Design and first results of a phantom study on the suitability of iterative reconstruction for lung-cancer screening with low-dose computer tomography. *Biomed. Eng.–Biomed. Tech.* 64(s2), 261.  
<https://doi.org/10.1515/bmt-2019-6023>  
Abstract, Poster, 53rd DGBMT Annual Conference on Biomedical Engineering, Frankfurt am Main.



# Inhalt

<b>1</b>	<b>Einleitung</b>	<b>7</b>
1.1	Motivation . . . . .	7
1.2	Stand der Forschung . . . . .	9
1.3	Zielsetzung der vorliegenden Dissertation . . . . .	13
<b>2</b>	<b>Material und Methoden</b>	<b>14</b>
2.1	Grundlagen der Computertomographie . . . . .	14
2.1.1	Aufbau klinischer Standard-CT-Systeme . . . . .	14
2.1.2	Bildrekonstruktion . . . . .	16
2.1.3	Darstellung und Bildqualitätsmaße . . . . .	26
2.1.4	CT-geeignete Dosisbegriffe . . . . .	27
2.1.5	Ansätze zur Dosisreduzierung . . . . .	30
2.2	Methodische Grundlagen der Statistik . . . . .	31
2.2.1	Maßzahlen einer Stichprobe . . . . .	31
2.2.2	$t$ -Test für zwei gepaarte Stichproben . . . . .	32
2.2.3	Friedman-Test . . . . .	33
2.2.4	Überschreitungswahrscheinlichkeit . . . . .	34
2.2.5	Kappa-Koeffizient nach Cohen . . . . .	35
2.2.6	Kappa-Koeffizient nach Fleiss . . . . .	38
2.3	Messphantome . . . . .	39
2.3.1	Allgemeine Spezifikationen . . . . .	39
2.3.2	Lungenrundherdmodelle . . . . .	40
2.3.3	Lungman-N1-Phantom . . . . .	41
2.3.4	QRM-Thorax-Phantom mit Extensionsringen . . . . .	42
2.4	Akquisitions- und Rekonstruktionsmethodik . . . . .	44
2.4.1	CT-Dosisprotokolle und Scanparameter . . . . .	44
2.4.2	CT-Experimente . . . . .	45
2.5	Methoden der Bildqualitätsanalyse . . . . .	47
2.5.1	Soft- und Hardware . . . . .	47
2.5.2	Messqualitätsprüfungen . . . . .	48
2.5.3	Bildqualitätsanalyse mit Lungman-N1 . . . . .	49
2.5.4	Bildqualitätsanalyse mit QRM-Thorax . . . . .	50
2.5.5	Analyse der Strahlenexposition . . . . .	51

---

<b>3</b>	<b>Ergebnisse</b>	<b>52</b>
3.1	Qualitätsmessungen . . . . .	52
3.2	Lungman-N1 . . . . .	54
3.2.1	Likert-Skala . . . . .	54
3.2.2	Signal-Rausch-Verhältnis . . . . .	57
3.2.3	Kontrast-Rausch-Verhältnis . . . . .	60
3.2.4	Fallbetrachtung . . . . .	63
3.3	QRM-Thorax . . . . .	64
3.3.1	Rangordnungstest mit Probe 1 . . . . .	65
3.3.2	Rangordnungstest mit Probe 2 . . . . .	69
3.4	Strahlendosisbewertung . . . . .	74
<b>4</b>	<b>Diskussion</b>	<b>76</b>
<b>5</b>	<b>Zusammenfassung</b>	<b>88</b>
<b>6</b>	<b>Literaturverzeichnis</b>	<b>89</b>
<b>A</b>	<b>Anhang</b>	<b>98</b>
A.1	Projektoren . . . . .	98
A.2	Datenblätter und Abnahmeprotokoll . . . . .	99
A.3	QRM-Thorax-Phantom mit Granulatfüllung . . . . .	103
A.4	Technische Spezifikationen des CT-Scanners . . . . .	104
A.5	Prüfbild 212 mit ROI-Messregionen für die Konstanzprüfung . . . . .	104
A.6	HU-Profile der 100 HU- und -50 HU-Kugel . . . . .	105
A.7	Rangordnungstestergebnisse für variable Phantomdurchmesser . . . . .	107
A.8	Tabellenverzeichnis . . . . .	109
A.9	Abbildungsverzeichnis . . . . .	111
A.10	Abkürzungsverzeichnis . . . . .	114
<b>B</b>	<b>Danksagung</b>	<b>117</b>
<b>C</b>	<b>Lebenslauf</b>	<b>118</b>

# 1 Einleitung

## 1.1 Motivation

Die Einführung eines Lungenkrebs-Screenings mit Niedrigdosis-Computertomographie (CT) für Hochrisikopatienten, (ehemals) starke Raucher<sup>1</sup>, wird in Deutschland erwartet. Erst kürzlich, seit März 2021, empfiehlt die U.S. Preventive Services Task Force (Rockville, Vereinigte Staaten von Amerika) eine jährliche dosissparende CT-Untersuchung zur Lungenkrebsfrüherkennung bei Hochrisikopatienten (United States Preventive Services Task Force, 2021). Gegenwärtig bewertet das Bundesamt für Strahlenschutz (BfS, Salzgitter, Deutschland), ob der zu erwartende Nutzen der Prozedur das potentielle Strahlenrisiko überwiegt (Siegmund-Schultze, 2019; Bundesamt für Strahlenschutz, 2019). Die Bewertung wurde zuletzt im Jahr 2018 durch die Ergebnisse der belgisch-niederländischen Nederlands-Leuven Longkanker Screenings Onderzoek-Studie (NELSON) motiviert (Koning et al., 2018). Für die Hochrisikopatienten zeigte sich darin, dass die Screeningprozedur die Lungenkrebs-Mortalität für Männer innerhalb einer Periode von zehn Jahren um 26 % herabsetzen konnte. Abhängig von der Nachbeobachtungszeit, acht und zehn Jahre, ließ sich für Frauen eine variable Verringerung von 39 % bis 61 % beobachten (International Association for the Study of Lung Cancer, 2018). Zuvor demonstrierte der US-amerikanische National Lung Screening Trial (NLST) im Jahr 2011, dass ein Lungenkrebs-Screening mit Niedrigdosis-CT das relative Risiko, an einem Lungenkrebs zu sterben, für Risikopatienten um 20 % senken könnte, was einer absoluten Risikoreduktion von 0,3 % entspricht (Center for Statistical Sciences, 2011). In beiden Studien wurde festgestellt, dass es sich bei der Mehrzahl der entdeckten Lungenkrebsfälle um Karzinome im Stadium I handelte. Die neu geschaffene Evidenzlage belegte die Fähigkeit der Niedrigdosis-CT-Prozedur, Lungenkarzinome frühzeitig zu erkennen (Center for Statistical Sciences, 2011; Yousaf-Khan et al., 2017). Der von der deutschen Bundesregierung eingesetzte Gemeinsame Bundesausschuss der großen Selbstverwaltungsorganisationen im Gesundheitswesen (G-BA, Berlin, Deutschland) beauftragte das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG, Berlin, Deutschland) mit der Bewertung des Lungenkrebs-Screenings mit Niedrigdosis-CT. Im Oktober 2020 publizierte das IQWiG einen Abschlussbericht, der den Nutzen der Prozedur für Hochrisikopatienten feststellte (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2020). Offen bleibt jedoch die Frage nach einer einheitlichen Screeningprozedur, Diagnostik und rechtlichen Regelungen (Bundesamt für Strahlenschutz, 2019).

Insbesondere der Niedrigdosis-CT-Ansatz unterliegt keinem einheitlichen Konzept. Vielmehr existiert eine Vielzahl von dosissparenden Hochkontrast-CT-Protokollen für die Detektion

---

<sup>1</sup>In der vorliegenden Dissertation wird bei Personenbezeichnungen das generische Maskulinum verwendet. Damit werden – falls nicht explizit anders angegeben – alle biologischen und soziologischen Geschlechter inkludiert.

von Lungenrundherden (de Lorenzo, 2011, S. 41). Um die hohen Qualitätsansprüche bei der Dosisreduktion an die Bildqualität zu erfüllen, bieten Hersteller unterschiedliche Lösungen an. Neben Hardwareansätzen existieren moderne Software-Implementationen, wie die Modulation des Röhrenstroms und der Röhrenspannung, die iterative Rekonstruktion (IR) sowie die Deep-Learning (DL)-Bildrekonstruktion (Siemens Healthineers, 2020; Kim et al., 2021; König et al., 2021). Wird die gefilterte Rückprojektion (FBP; engl.: Filtered Back Projection) als analytisches Standard-Rekonstruktionsverfahren in Verbindung mit der Niedrigdosis-CT angewendet, entsteht ein verrauschter Bildeindruck. Dieser Bildqualitätsverlust kann einen negativen Einfluss auf die diagnostische Leistung von Radiologen nehmen (König et al., 2019). Ungeachtet dessen ist die Dosisreduktion für die Untersuchung des Lungenorgans mit einer hohen relativen Strahlensensibilität (Nagel, 2002, S. 14) für das Lungenkrebs-Screening von Priorität. Moderne modellbasierte IR-Algorithmen erzeugen einen rauschreduzierten, veränderten Bildeindruck und bieten eine vielversprechende Alternative zur FBP-Rekonstruktion bei der Niedrigdosis-CT.

Ein Hauptvorteil der IR-Rekonstruktionsalgorithmen besteht in ihrer inhärenten diskreten mathematischen Formulierung. Diese ermöglicht eine vollständige Modellierung der Akquisitionspipeline. Hierzu zählen die Modellierungen: der CT-Geometrie und -Komponenten (z. B. Detektordesign und -empfindlichkeit, Fokausdehnung und -typ, Strahlform), der Physik (z. B. Strahlenspektrum, Wechselwirkungsprozesse), von Vorwissen (z. B. elektronische Rauscheigenschaften, Objekteigenschaften) und der Zählstatistik der Röntgenquanten am Detektor (z. B. Poisson- oder Gaußverteilung). Die zuletzt genannte Zählstatistik kann im Algorithmus der Rekonstruktion, in den Rohdaten oder in den Bilddaten implementiert werden. Die Methoden werden zudem als statistische IR-Rekonstruktionsalgorithmen bezeichnet und zählen zu den stark rauschreduzierenden Implementierungen (Beister et al., 2012; Willeminck und Noël, 2019). Eine moderne und kommerziell verfügbare modellbasierte IR-Implementierung ist die Advanced Modeled Iterative Reconstruction (ADMIRE; Siemens Healthineers, Forchheim, Deutschland). Dieser Algorithmus basiert auf einem erweiterten statistischen IR-Ansatz mit einer optimierten Modellierung der CT-Geometrie und -Komponenten (Siemens Healthineers, 2020; Ramirez-Giraldo et al., 2017). Der veränderte Bildeindruck der Niedrigdosis-CT-Bilder mit modellbasierter IR steht Vorbehalten von Radiologen und einem möglichen Einfluss auf die diagnostische Leistung bei einem Lungenkrebs-Screening gegenüber.

Eine umfangreiche Bildqualitätsanalyse zur Bewertung der Eignung der Niedrigdosis-CT mit modellbasierter IR-Rekonstruktion für das Lungenkrebs-Screening bedarf eines bivalenten Ansatzes. Neben der objektiven Bildqualitätsbeurteilung anhand geeigneter Bildqualitätsmaße ist die subjektiv-visuelle Wahrnehmung der CT-Bilder von Radiologen zu untersuchen. Um diagnostische Lesegewohnheiten festzustellen, ist der Vergleich mit FBP-rekonstruierten CT-Bildern indiziert. Das weitreichende Angebot dosissparender Hochkontrast-CT-Protokolle inspiriert die Analyse mit mehreren Varianten. Zur Vollständigkeit empfiehlt sich ein klinisch eingesetztes Standard-Hochkontrast-CT-Protokoll sowie zwei weitere Niedrigdosis-CT-Protokolle (Nagel, 2002, S. 49 f.; Prokop, 2001): Low-Dose (LD; 120 kV, 40 mAs) und Ultra-Low-Dose (ULD; 120 kV, 20 mAs). Das breite Spektrum von untersuchten Rekonstruktionskernen eines aktuellen

Forschungsansatzes zur Lungenrundherddetektion mit einem DL–Computer–Aided Detection (CAD; deutsch: Computergestützte Detektion)–System motiviert dazu, den Rekonstruktionskern als variablen Analyseparameter zu definieren (Blazis et al., 2021). Die erforderliche Anzahl an CT–Experimenten geht mit einer hohen applizierten Strahlendosis einher. Aus Gründen des Strahlenschutzes von möglichen Probanden ergibt sich daher die Überlegung einer ex–vivo–Studie unter Verwendung von anthropomorphen Phantomen. Neben der beständigen Form sind diese mit artifiziellen Lungenrundherden erweiterbar und bilden einen definierten experimentellen Aufbau. Gleichwohl ist die Dokumentation und Bewertung der Strahlendosis ein unerlässliches Kriterium für die Akzeptanz eines Lungenkrebs–Screenings, da das CT–Experiment primär an gesunden Patienten erfolgt. Überdies können Rückschlüsse auf die Verhältnismäßigkeit zwischen Bildqualität und Dosisreduktion gezogen werden.

## 1.2 Stand der Forschung

Das substanzielle Ziel europäischer und US–amerikanischer randomisierter kontrollierter Studien zum Lungenkrebs–Screening mit Niedrigdosis–CT war die Untersuchung des potentiellen Einflusses auf die Lungenkrebs–Mortalität. Eine Studienübersicht, geordnet nach Startjahr und unter Angabe der Teilnehmer, wird in Tab. 1.1 vorgestellt.

Tab. 1.1: Europäische und US–amerikanische randomisierte kontrollierte Studien mit dem primären Endpunkt, den Effekt des Lungenkrebs–Screenings mittels Niedrigdosis–CT auf die Lungenkrebsmortalität zu untersuchen (Prosch und Ebner, 2019). Die Studien sind nach Startjahr und unter Angabe der Teilnehmer sortiert worden.

Studie	Startjahr, Land	Teilnehmer
DANTE	2001, Italien	2472
NLST	2002, Vereinigte Staaten von Amerika	53456
NELSON	2003, Niederlande und Belgien	15822
ITALUNG	2004, Italien	3206
DLCST	2004, Dänemark	4104
MILD	2005, Italien	4479
LUSI	2007, Deutschland	4000

Anhand des italienischen Detection and screening of early lung cancer by novel imaging technology and molecular assays Trial (DANTE) wurde in der CT–Screeninggruppe eine erhöhte Lungenrundherddetektionsrate gegenüber der Kontrollgruppe identifiziert (Infante et al., 2008; Infante et al., 2009). Eine signifikante Lungenkrebs–Mortalitätsreduktion konnte nicht nachgewiesen werden. Hierzu vergleichbare Ergebnisse hatte der dänische Danish Lung Cancer Screening Trial (DLCST; Saghiri et al., 2012). Der italienische Multicentric Italian Lung Detection Trial (MILD) demonstrierte einen nicht signifikanten Anstieg der Mortalität in der CT–Screeninggruppe (Pastorino et al., 2012). Die gepoolte Analyse mit den Ergebnissen der DANTE– und MILD–Studie ergab eine nicht signifikante Gesamtmortalitätsreduktion von 11 % durch das Lungenkrebs–Screening (Infante et al., 2017). Vergleichbare Erkenntnisse

erbrachte der Italian Lung Cancer Screening Trial (ITALUNG; Paci et al., 2017). Erstmals gelang der Nachweis einer signifikant reduzierten Mortalität durch Lungenkrebs-Screening mit dem US-amerikanischen NLST (Center for Statistical Sciences, 2011). Dies ergaben auch die Ergebnisse der belgisch-niederländischen NELSON-Studie (Koning et al., 2018). Mit der deutschen Lung Cancer Screening Intervention-Studie (LUSI) konnte eine signifikante Mortalitätsreduktion ausschließlich für Frauen festgestellt werden (Becker et al., 2020). Eine jüngst publizierte systematische Übersicht und Meta-Analyse randomisierter kontrollierter Studien zum Lungenkrebs-Screening mit Niedrigdosis-CT für Hochrisikopatienten ergab eine um 12 % signifikant verringerte Lungenkrebssterblichkeit (Hunger et al., 2021).

Sekundäre Untersuchungsziele der Studien zum Lungenkrebs-Screening waren: Kosteneffizienz, Überdiagnosen, Überlebensraten, Lebensqualität, Stadium der Lungenrundherde bei Diagnose und geeignete Screening-Intervalle (Hunger et al., 2021; Prosch und Ebner, 2019; O'Dowd und Baldwin, 2018; Yousaf-Khan et al., 2017; Center for Statistical Sciences, 2011). Offene Risikoaspekte der Screeningprozedur, insbesondere die negativen Auswirkungen auf die Raucherentwöhnung sowie unerwünschte Vorkommnisse bei der diagnostischen Aufarbeitung, diskutierten Nanavaty et al. (Nanavaty et al., 2014). Kauczor et al. berichten von nicht ausreichend betrachteten inhärenten Risikoaspekten (Kauczor et al., 2020). Im besonderen Fokus stand die strahleninduzierte Morbidität, welche sich aus wiederholten CT-Untersuchungen und der diagnostischen Aufarbeitung von Zufallsbefunden ergab. Mit den Ergebnissen der italienischen Continuous Observation of Smoking Subjects-Studie (COSMOS) wurde eine Risikoberechnung für einen strahleninduzierten Lungentumor durchgeführt. Die Resultate legten offen, dass über eine zehnjährige Lungenkrebs-Screening-Periode (Screeningintervall: 1 Jahr) mit 259 detektierten Lungenkarzinomen in jedem 108-ten Fall ein strahleninduzierter Lungentumor zu erwarten sei (Mascalchi und Sali, 2017). Mit dem United Kingdom Lung Cancer Screening Trial (UKLS) wurden praktische und emotionale Barrieren der Hochrisikopatienten bei der Teilnahme an einem Lungenkrebs-Screening identifiziert (Ali et al., 2015).

Entscheidend für die Akzeptanz des Lungenkrebs-Screenings mit Niedrigdosis-CT in der Allgemeinbevölkerung wird die Reduktion der Strahlendosis unter Erhalt einer ausreichenden diagnostischen Leistung sein. Miller et al. wiesen eine hohe Sensitivität bei der Lungenrundherddetektion für die Ultra-Niedrigdosis-CT mit modellbasierter IR-Rekonstruktion nach (Miller et al., 2019). Die Strahlenexposition war vergleichbar mit der einer normalen Radiographie (effektive Dosis: 0,13 mSv). Derartige Strahlendosisreduktionen könnten zu einer relevanten Reduktion der kumulativen Strahlenexposition bei einem Screening führen.

Moderne Dosisüberwachungssysteme (DMS; engl.: Dose Monitoring System) erfassen, bearbeiten und verwalten Daten über die Strahlendosis von Patienten bei diagnostischen oder therapeutischen Verfahren mit ionisierender Strahlung. Eine sinnvolle Dosisüberwachung schließt die Schwellenwertanpassung der implementierten Dosisalarmfunktionen für Lungenkrebs-Screening geeignete Niedrigdosis-CT-Protokolle gegenüber Standard-CT-Protokollen ein. Somit wären potentielle Statistik-Ausreißer bei der Dokumentation vermeidbar. Künftig ist zu erwarten, dass die Funktion von DMS über Dosismetriken hinausgeht und mitunter für Bildqualitätsanalysen eingesetzt werden könnte. Die automatische Erfassung von objektiven Bildqua-

litätsmaßen könnte ermöglichen, Rückschlüsse aus den Bildqualitäts- und Strahlendosisparametern für die Optimierung von CT-Dosisprotokollen zu nutzen (Tsalafoutas et al., 2020).

Gegenwärtig existiert eine Vielzahl dosissparender, kommerzieller IR-Implementationen. Die Algorithmen führender CT-Hersteller sind geordnet nach Rekonstruktionsverfahren in Tab. 1.2 gelistet.

Tab. 1.2: Kommerzielle IR-Implementationen der führenden Hersteller. Die geordnete Darstellung umfasst statistische und modellbasierte Algorithmen. Die Zuordnungen und kumulierten Akronyme orientieren sich an (Willemink und Noël, 2019).

Kommerzielle IR-Implementationen		
Hersteller	statistisch	modellbasiert
GE Healthcare	ASiR, ASiR-V	Veo
Siemens Healthineers	SAFIRE	ADMIRE
Philips Healthcare	iDose	IMR
Canon Healthcare	AIDR 3D	FIRST

Darüber hinaus gibt es IR-Algorithmen zur Metallartefaktkorrektur, z. B. die Implementierungen: Iterative Metal Artifact Reduction (iMAR; Siemens Healthineers, 2021) sowie Metal Artifact Reduction for Orthopedic Implants (O-MAR; Philips Healthcare, 2021).

Fortschrittliche DL-Implementationen, wie TrueFidelity (GE Healthcare, Chicago, Vereinigte Staaten von Amerika), finden in der klinischen Routine zunehmend Anwendung (Hsieh et al., 2019). Der DL-Ansatz basiert auf einem tiefen neuronalen Netzwerk (DNN; engl.: Deep Neuronal Network). Die Implementation bietet mit einem komplexen DL-Modell sowie einer intelligenten Rauschunterdrückung für die Niedrigdosis-CT das Potential analytische und iterative Rekonstruktionsverfahren in Bildqualität und Rechengeschwindigkeit zu übertreffen (Hsieh et al., 2019). Ersten Erkenntnissen nach führt die DL-Bildrekonstruktion bei der Niedrigdosis-CT des Thorax unter Erhalt einer exzellenten Bildqualität zu einem höheren Signal-Rausch-Verhältnis gegenüber der modellbasierten IR- und der FBP-Rekonstruktion (Kim et al., 2021; Zeng et al., 2021).

Macri et al. führten eine erste bivalente Bildqualitätsanalyse an Lungen-CT-Bildern durch (Macri et al., 2016). Die Untersuchung befasste sich mit dem Einfluss der statistischen IR-Implementation SAFIRE in Verbindung mit LD- und ULD-Protokollen auf die anatomische Bildqualität. Die in variablen IR-Stärken rekonstruierten CT-Bilder wurden einer objektiven sowie subjektiv-visuellen Bildqualitätsanalyse unterzogen und anatomische Strukturen quantitativ vermessen. Die statistische IR-Rekonstruktion mit dem LD-Protokoll wurde als effiziente Lungen-CT-Rekonstruktion vorgeschlagen. Hohe statistische IR-Stärken mit dem ULD-Protokoll eigneten sich als guter Kompromiss hinsichtlich der anatomischen Bildqualität und der Dosisreduktion. Weitere Studien mit modellbasierten IR-Implementationen anderer Hersteller befassten sich gleichermaßen mit dem Lungenkrebs-Screening (Hassani et al., 2018) oder anderen Untersuchungen, z. B. der Pulmonalisangiographie (Kurfürst, 2018).

Objektive Bildqualitätsmaße sind eine weit verbreitete Methode zur Bildqualitätsanalyse von IR-rekonstruierten Lungen-CT-Bildern. Aktuelle Phantomstudien führen diese Analysen unter Verwendung von LD- und ULD-Protokollen durch (Afadzi et al., 2020; Lee und Goo, 2018). Andersen et al. untersuchten für verschiedene Hersteller die Bildqualität in IR-rekonstruierten Lungen-CT-Bildern (Andersen et al., 2018). Zur Anwendung kamen herstellereigene: Algorithmen (FBP, statistische und modellbasierte IR), CT-Scanner, Lungen-CT-Dosisprotokolle und IR-Rekonstruktionsempfehlungen. Es wurden CT-Scans in drei Dosisstufen bei variablem Phantomdurchmesser akquiriert. Die erfassten Bildqualitätsmaße in den CT-Bildern waren: das Kontrast-Rausch-Verhältnis, die Rauschleistungsfunktion sowie die Modulationstransferfunktion. Alle modellbasierten IR-Methoden verbesserten das Kontrast-Rausch-Verhältnis mehr als andere IR-Techniken bei gleichbleibender räumlicher Auflösung. Verglichen mit der FBP-Rekonstruktion konnten die IR-Methoden bis auf eine Ausnahme die Rauschtextur in den CT-Bildern erhalten oder verbessern. Eine ähnliche Studie von Afadzi et al. bestätigte in Verbindung mit ULD-CT Protokollen die Verbesserung des Kontrast-Rausch-Verhältnisses mit modellbasierter IR-Rekonstruktion (Afadzi et al., 2020).

Verdun et al. diskutieren mathematische Modellbeobachter (MO; engl.: Model Observer) als eine vielversprechende Methode bei der Beurteilung der medizinischen Bildqualität (Verdun et al., 2015). Die MO erfüllen klinisch relevante Aufgaben, hierzu zählen u. a. diagnostische Detektionsaufgaben in IR-rekonstruierten Lungen-CT-Bildern. Die MO-Beobachterleistung bzw. der Vergleich mit menschlicher Beobachterleistung lässt sich mit Grenzwertoptimierungskurven (ROC; engl.: Receiver Operating Characteristic-Curve) und ihrer Fläche unterhalb der Kurve (AUC; engl.: Area Under the Curve) bewerten. In einer Phantomstudie untersuchten Rampado et al. die Detektionsleistung in IR-rekonstruierten Niedrigkontrast-CT-Bildern mit variabler IR-Stärke für zwei Phantomdurchmesser sowie die Strahlendosis (Rampado et al., 2021). Simulierte Leberläsionen wurden von menschlichen Beobachtern und einem MO untersucht und die Detektionsleistung über AUC verglichen. Beide Beobachter lieferten konsistente Ergebnisse und zeigten eine unbeeinflusste Detektionsleistung in den IR-Rekonstruktionen für beide Durchmesser bis IR-Stärke ASIR-V 50 % (bei 70 % Strahlendosisreduktion). Bei höherem IR-Stärkeniveau und höherer Strahlendosisreduktion nahm die Detektionsleistung signifikant ab. Der Forschungsansatz von Han und Baek stellt einen Leistungsvergleich eines konventionellen MO mit einem anthropomorphen MO auf Basis eines neuronalen Faltungsnetzwerks (CNN; engl.: Convolutional Neural Network) in simulierten CT-Rekonstruktionen der Brust vor (Han und Baek, 2020). Der anthropomorphe CNN-Beobachter sagte die Leistung eines menschlichen Beobachters besser voraus als der konventionelle Modellbeobachter.

Aktuelle Forschungsarbeiten bewerten maschinelles Lernen (ML; engl.: Machine Learning) und CAD-Software-Systeme anhand von ROC in Lungenphantomen (Fu et al., 2020; Wielpütz et al., 2015). Erstmals erforschten Blazis et al. den Einfluss von CT-Rekonstruktionsparametern auf die Lungenrundherd-Detektionsleistung eines kommerziellen DL-CAD-Systems (Blazis et al., 2021). Mit zwei CT-Scannern wurden klinische Lungen-CT-Scans akquiriert und mit scannerspezifischen Parametern rekonstruiert: FBP und statistische bzw. modellbasierte IR-Algorithmen, drei Rekonstruktionskerne (standard, weich, hart) und zwei Schichtdicken (1 mm



und 3 mm). In den CT-Bildern erfolgte die Lungenrundherddetektion durch die Software und durch Radiologen. Die Ergebnisse beider Detektionsmethoden wurden verglichen. Ein Einfluss auf die Leistungsparameter des DL-CAD-Systems mit zunehmendem IR-Stärkeniveau gegenüber der FBP-Rekonstruktionen konnte nachgewiesen werden. Dies wurde an einem Verlust der System-Sensitivität bei gleichzeitiger Präzisionszunahme beobachtet. Der Effekt fiel für dickere Schichten (3 mm) mit modellbasierter IR-Rekonstruktion deutlich geringer aus. Die zur Diagnostik und Therapie von Lungenrundherden eingesetzten DL-CAD-Systeme und ML-Algorithmen bieten durch automatische Detektion, Segmentierung, Quantifizierung, Klassifikation sowie Verlaufsprädikation auch für weitere Lungenerkrankungen, wie z. B. parenchymatische Veränderungen, Verbesserungspotential (Prayer et al., 2020).

### 1.3 Zielsetzung der vorliegenden Dissertation

Die vorliegende Dissertation widmete sich der Untersuchung der Eignung der Niedrigdosis-CT mit modellbasierter IR-Rekonstruktion für das Lungenkrebs-Screening. Zur Erfüllung dieser Aufgabe wurde eine bivalente Bildqualitätsanalyse im Rahmen einer quantitativen Phantomstudie umgesetzt, wobei die Überlegungen zum Studiendesign aus Kap. 1.1 berücksichtigt worden sind.

Die spezialisierten Analysen verwendeten die Rohdaten aus zwei Niedrigdosis-CT-Experimenten. Zu jedem Experiment wurde ein definierter experimenteller Aufbau eines kommerziellen anthropomorphen Lungenphantoms konzipiert. Die mit modellbasierter IR-Rekonstruktion erstellten CT-Bilder der Experimente wurden verwendet, um die Interdependenz zwischen modellbasierten IR-Stärken, Rekonstruktionskernen und Niedrigdosis-CT-Protokollen zu untersuchen und einen Vergleich mit konventionellen FBP-Rekonstruktionen durchzuführen. Hierfür erörterte die vorliegende Promotionsarbeit die folgenden Fragestellungen:

- Frage 1:** Ist die modellbasierte IR-Rekonstruktion in Verbindung mit Niedrigdosis-CT als Alternative zur FBP-Rekonstruktion für das Lungenkrebs-Screening geeignet?
- Frage 2:** Welche modellbasierte IR-Stärke eignet sich für die Detektionsaufgabe bei einem Lungenkrebs-Screening mit Niedrigdosis-CT am besten?
- Frage 3:** Korreliert die von Radiologen subjektive wahrgenommene Eignung der IR-rekonstruierten CT-Bilder mit deren radiologischer Erfahrung?

## 2 Material und Methoden

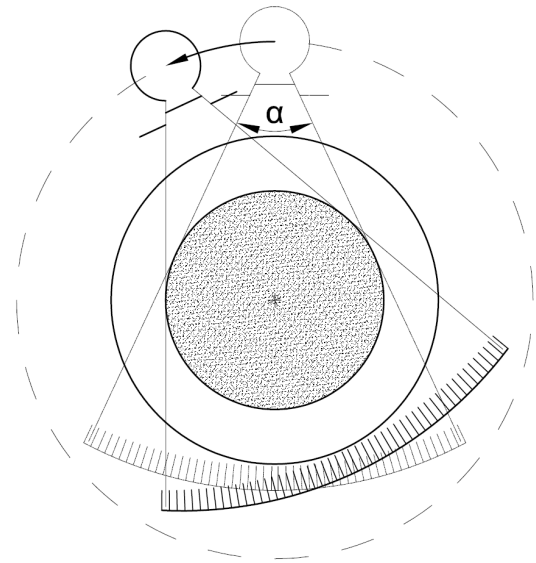
### 2.1 Grundlagen der Computertomographie

#### 2.1.1 Aufbau klinischer Standard-CT-Systeme

Die Systemhardware eines CT gliedert sich in die Scannerapparatur und die Anzeige- bzw. Bedienhardware (Dössel und Buzug, 2014, S. 70). Das Scannerdesign von klinisch eingesetzten Standard-CT-Systemen wird dem Entwicklungsstand der dritten CT-Generation zugeordnet, dargestellt in Abb. 2.1 (a), S. 15. Die technische Realisierung dieser Entwicklungsstufe beläuft sich auf drei Hauptkomponenten: die Röntgenröhre (Quelle) zur Strahlungserzeugung mit der gekoppelten Messeinheit (Detektor) sowie der Positionierungsvorrichtung (Tisch). Dabei wird die Quellen-Detektor-Kombination von der CT-Gantry getragen. Die Aufnahmegeometrie rotiert während des CT-Experiments auf einer definierten Kreisbahn um das Isozentrum im Messfeld. Das Prinzip der dritten CT-Generation sowie ein rechtshändiges Koordinatensystem  $(x, y, z)$  für die Orientierung im Raum bei der Vorstellung technischer Instrumente und Funktionen werden in Abb. 2.1 (b) und (c), S. 15, vorgestellt. Ein spezifisches Konstruktionsmerkmal der dritten CT-Generation repräsentiert eine Quelle mit vorgeschaltetem Kollimator, durch dessen Öffnung in der  $xy$ -Ebene ein Fächerstrahl (engl.: fan beam) mit einem Öffnungswinkel  $\alpha_{fb}$  (engl.: fan beam angle) zwischen  $40^\circ$  und  $60^\circ$  erzeugt wird. Moderne Quellen besitzen einen sich periodisch bewegendenden Fokus (engl.: flying focal spot) in  $z$ -Richtung (Flohr et al., 2004). Zudem besteht bei der Fächerstrahl- bzw. Einzeilen-CT das Detektorarray aus einer geraden oder gebogenen Detektorreihung, sog. Zeile, von mehreren Einzeldetektorelementen, sog. Dixel (Buzug, 2014, S. 49 f.). CT-Systeme nach aktuellem Stand der Technik beinhalten eine weiterentwickelte Mehrzeilen-CT-Aufnahmegeometrie. Hierbei wird das quellenseitige Kollimatorsystem zusätzlich in  $z$ -Richtung geöffnet und eine Kegelstrahlgeometrie (engl.: cone beam) erzeugt. Analog zur Kollimatoröffnung wird das Detektorarray in  $z$ -Richtung um parallel angeordnete Detektorzeilen erweitert (Dössel und Buzug, 2014, S. 74 ff.). Konstruktive Varianten der Detektoren für die Mehrzeilen-CT stellen der Hybriddetektor und der Matrix-Array-Detektor dar. Der Hybriddetektor besitzt Detektorzeilen variabler Dixelgrößen in  $z$ -Richtung. Hingegen besteht das Matrix-Array aus einer strukturierten Matrixanordnung gleichgroßer Dixel (Stiller, 2011). Hinsichtlich der Aufnahmegeometrie existieren für die dritte CT-Generation Dual-Source-CT-Systemvarianten mit zwei unter  $90^\circ$  implementierten Quellen-Detektor-Kombinationen (Dössel und Buzug, 2014, S. 106).



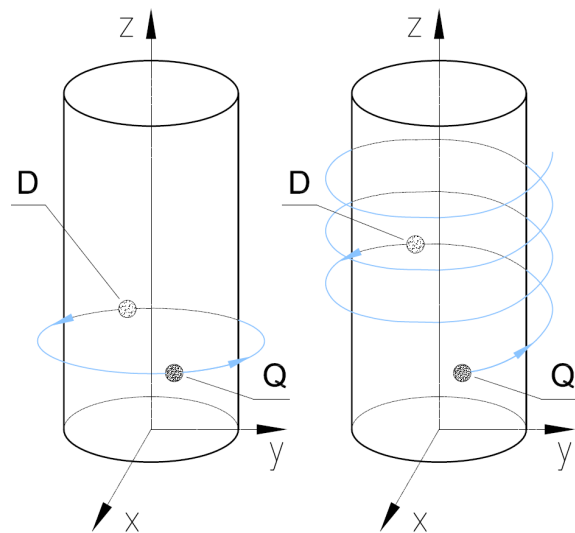
(a) State of the Art CT-Scanner



(b) Prinzip der dritten CT-Generation



(c) CT-Gantry mit Koordinatensystem



(d) Trajektorie, links Axial-, rechts Spiral-CT

Abb. 2.1: Die dritte CT-Generation mit einer frontalen Ansicht eines (a) Dual-Source-CT-Scanners mit Tisch und CT-Gantry sowie dem (b) generationsspezifischen, schematischen Prinzip. Anhand eines (c) festgelegten, rechtshändigen Koordinatensystems, hier in der CT-Gantry, lassen sich die möglichen (d) Trajektorien der Quellen-Detektor-Kombination sowie der technische Aufbau raumorientiert einheitlich beschreiben.

Während der Scanprozedur wird das Untersuchungsobjekt aus einer Vielzahl von Winkeln durchstrahlt. Die technische Umsetzung bedingt eine durch die Aufnahmegeometrie festgelegte Trajektorie der Quelle um das Untersuchungsobjekt. Bei der sequenziellen, axialen CT-Technik bewegt sich der Lagerungstisch nur in den Messpausen um einen definierten Schichtabstand in  $z$ -Richtung. Demzufolge resultiert zu jedem Zeitpunkt der Messung eine kreisförmige Trajektorie der Quelle. Eine weitere Aufnahmetechnik ist die Spiral-CT. Dabei wird der Tisch während der Messung kontinuierlich vorgeschoben und es bildet sich eine Helix-Trajektorie aus (Dössel und Buzug, 2014, S. 62 ff.). Schematische Darstellungen der genannten Trajektorien werden in Abb. 2.1 (d) geboten.

**Tuy–Smith–Kriterium.** Die Vollständigkeit der gemessenen Daten für zwei- und dreidimensionale Aufnahmebereiche kann mit dem Tuy–Smith–Kriterium bewertet werden (Tuy, 1983). Nach dem Kriterium gilt, dass eine exakte Rekonstruktion einer einzelnen Schicht durch ein Objekt dann möglich ist, wenn alle Strahlen, die die zu rekonstruierende Fläche schneiden, die Bahn der Quellen–Detektor–Bewegung mindestens einmal schneiden. Für Rekonstruktionsvolumina gilt, dass alle Flächen durch das Rekonstruktionsvolumen mindestens einmal die Bahn der Quellen–Detektor–Bewegung schneiden müssen. Das Vollständigkeitskriterium gewährleistet damit, dass für jeden Punkt innerhalb eines Objektes aus allen Winkelbereichen Informationen vorliegen (Buzug, 2014, S. 310 ff.). Die technische Verwirklichung basiert für die dritte CT–Generation mit einer kreisförmigen Trajektorie auf einem Aufnahme- bzw. Projektionswinkel  $\gamma$  von  $180^\circ$  plus dem Fächerwinkel  $\alpha_{fb}$ . Die Helix–Trajektorie bei der Spiral–CT erfasst alle Punkte in einem Objekt entlang der  $z$ –Richtung mindestens einmal, jedoch liegen für eine beliebige Schichtlage  $z_r$  nicht aus allen Winkelbereichen, gemäß dem Tuy–Smith–Kriterium, Informationen vor. Für die exakte Rekonstruktion werden die Teilinformationen mathematisch anhand der  $z$ –Interpolation für den festgelegten Projektionswinkel ergänzt. Daten aus Winkeln, die für  $z_r$  nicht gemessen wurden, werden aus Informationen umliegender, tatsächlich gemessener Winkel der Helixbahn interpoliert (Dremel, 2017, S. 7; Buzug, 2014, S. 255 ff.).

## 2.1.2 Bildrekonstruktion

### Grundlegendes Prinzip

Die Aufgabe computertomographischer Rekonstruktionsmethoden besteht in der Lösung eines inversen Problems. Dies beruht auf der Umrechnung der aus verschiedenen Winkeln durch den Schnitt eines Objektes gemessenen Intensitätsprofile von Röntgenstrahlen, in auf Volumenelemente (Voxel) verteilte Schwächungskoeffizienten. Die mathematische Beschreibung des exponentiellen Abfalls der Intensität beim Durchlaufen homogener Materie einer Dicke  $d$  mit dem Schwächungskoeffizienten  $\mu$  lässt sich mit dem Gesetz von Lambert–Beer (Buzug, 2014, S. 22 u. S. 36 ff.) formulieren,

$$I(d) = I(0)e^{-\mu d}, \quad \text{mit } \mu = \mu_s + \alpha, \quad (2.1)$$

wobei  $I(0)$  die Intensität vor der Materie ist und  $\mu$  materialspezifisch und energieabhängig. Zudem setzt  $\mu$  sich additiv aus dem Streukoeffizienten  $\mu_s$  und dem Absorptionskoeffizienten  $\alpha$  zusammen. Er wird in der Einheit  $\text{m}^{-1}$  angegeben. In den Betrachtungen wird auf die Energieauflösung der räumlich verteilten  $\mu$  verzichtet und ein monochromatisches Spektrum angenommen.

Für den Schwächungsprozess in einem inhomogenen Objekt folgt aus Gl. (2.1):

$$I(s) = I(0) \exp\left(-\int_0^s \mu(\eta) d\eta\right). \quad (2.2)$$

Durch Logarithmieren von Gl. (2.2) ergibt sich die Projektion

$$p(s) = -\ln\left(\frac{I(s)}{I(0)}\right) = \int_0^s \mu(\eta) d\eta. \quad (2.3)$$

Im Integral auf der rechten Seite sind alle  $\mu(\eta)$  entlang des Strahls zwischen 0 und  $s$ .

Im Folgenden wird die CT-Rekonstruktion für die Parallelstrahlgeometrie betrachtet. Die Rekonstruktion für die Kegelstrahlgeometrie unterscheidet sich hiervon um geometrische Faktoren. Alle parallelen Strahlen durch das Objekt definieren eine Schnittebene in einem ruhenden Laborsystem  $(x, y)$ . Das Projektionsintegral wird als Funktion der Detektorcoordinate  $\xi$  und des Projektionswinkels  $\gamma$  angegeben. Dabei ist  $\gamma$  jeder Winkel aus dem Intervall  $[0^\circ, 180^\circ]$  (Buzug, 2014, S. 109 ff.). Die radiale Objektabtastung erfordert die Einführung eines auf dem Detektor liegenden, um das Isozentrum rotierenden Koordinatensystems  $(\xi, \eta)$ , siehe Abb. 2.2.

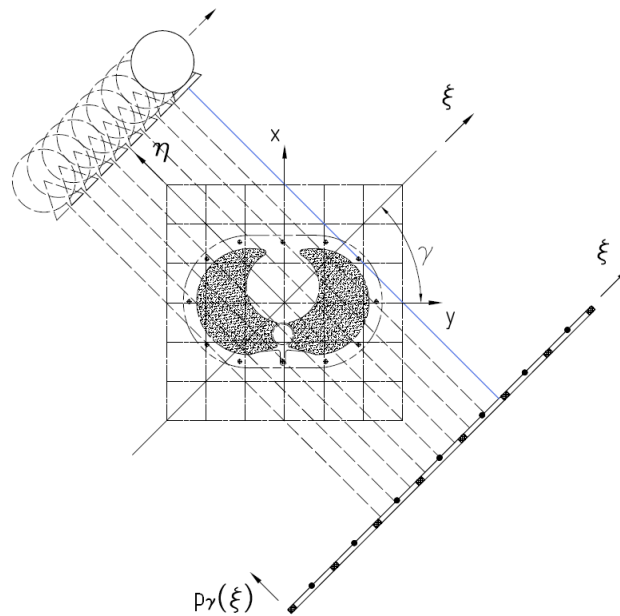


Abb. 2.2: Schema einer zweidimensionalen räumlichen Abtastung mit einer Parallelstrahlgeometrie in der CT. Die parallelen Strahlen einer linear verschobenen Quelle unter dem Projektionswinkel  $\gamma$  definieren eine Schnittebene in dem ruhenden Laborsystem  $(x, y)$ . Auf dem Detektor liegt das rotierende Koordinatensystem  $(\xi, \eta)$  mit der Detektorcoordinate  $\xi$  und der Coordinate des Strahls  $\eta$ . Die Projektion  $p_\gamma(\xi)$  stellt die logarithmierten gemessenen Intensitätswerte auf dem Detektor bzw. die entlang von  $\eta$  integrierten Werte von  $\mu$  dar.

Das Projektionsintegral ergibt für das  $(\xi, \eta)$ -System mit  $\gamma$ ,

$$p_\gamma(\xi) = \int_{-\infty}^{\infty} \mu(\xi, \eta) d\eta. \quad (2.4)$$

Die Integrationsgrenzen auf der rechten Seite gelten unter der Annahme, dass  $\mu$  außerhalb des Untersuchungsobjekts verschwindet. Für ein festes  $\gamma$  stehen das feststehende Labor- und

das rotierende Koordinatensystem in Beziehung (Buzug, 2014, S. 112 f.):

$$x = \xi \cos \gamma - \eta \sin \gamma, \quad (2.5)$$

$$y = \eta \cos \gamma + \xi \sin \gamma. \quad (2.6)$$

Die Menge aller Projektionen  $p_\gamma(\xi)$  heißt Sinogramm. In der Mathematik wird das Sinogramm als Radon- bzw. Rohdatenraum bezeichnet. Mit allen Strahlen durch die Schnittebene unter jedem  $\gamma$  ist die Gesamtheit aller  $p_\gamma(\xi)$  gleich einer zweidimensionalen Radontransformierten der Funktion  $\mu(x, y)$  (Radon, 1917).

**Fourier–Scheiben–Theorem.** Die Rücktransformation der gemessenen Rohdaten, das heißt die Rekonstruktion des Schwächungskoeffizienten  $\mu(x, y)$ , wird durch das Fourier–Scheiben–Theorem motiviert. Dabei wird die eindimensionale Fouriertransformierte  $P_\gamma(k)$  der gemessenen Projektionen  $p_\gamma(\xi)$  mit  $k$  der zu  $\xi$  gehörigen Ortsfrequenz in der zweidimensionalen Fouriertransformierten  $M(k_x, k_y)$  von  $\mu(x, y)$  unter dem Projektionswinkel  $\gamma$  gefunden. Für die kartesischen Ortsfrequenzkoordinaten gilt:  $k_x = k \cos \gamma$  bzw.  $k_y = k \sin \gamma$ . Das Fourier–Scheiben–Theorem sichert die folgende Gleichheit zu (Dössel und Buzug, 2014, S. 79):

$$P_\gamma(k) = M(k \cos \gamma, k \sin \gamma) = M(k_x, k_y). \quad (2.7)$$

Die Rekonstruktion der Schwächungskoeffizienten  $\mu(x, y)$  kann direkt über die inverse zweidimensionale Fouriertransformation von  $M(k_x, k_y)$  erfolgen. Realisiert wird die inverse zweidimensionale Fouriertransformation über einen schnellen Fouriertransformationsalgorithmus (FFT; engl.: Fast–Fourier–Transformation). Der Algorithmus bedingt, dass die radial abgetasteten  $P_\gamma(k)$  im Fourierraum in ein kartesisches Koordinatensystem überführt werden. Die Interpolation von  $P_\gamma(k)$  in das kartesische  $(k_x, k_y)$ –Koordinatensystem wird als Regridting bezeichnet. Durch den Koordinatensystemwechsel werden die hohen Ortsfrequenzen im Allgemeinen unterdrückt, dies ist mit Bildqualitätseinbußen verbunden (Buzug, 2014, S. 125 f.).

**Shannon–Nyquist–Kriterium.** Durch den Abtastvorgang bei einem CT–Experiment wird das kontinuierliche Signal digitalisiert. Anstatt der kontinuierlich formulierten Projektionsintegrale der vorhergehenden Betrachtungen liegen bei einem realen CT–Experiment Projektionssummen vor. Die vollständige Abbildung eines ordinalen Signals wird durch das Shannon–Nyquist–Kriterium bei der diskreten Abtastung gewährleistet. Dafür muss die Abtastfrequenz  $f_a$  größer als das Zweifache der größten Signalfrequenz  $k_{max}$  sein,

$$f_a > 2k_{max}. \quad (2.8)$$

### Gefilterte Rückprojektion

Die FBP ist ein analytisches Rekonstruktionsverfahren, welches die Bildqualitätseinbußen durch das Regridding bei der direkten Fourierrekonstruktion eliminiert. Die nachfolgenden formellen Beziehungen zeigen, wie mit der FBP die Funktion  $\mu(x, y)$  aus den Projektionsdaten erhalten wird. Aus dem Fourier–Scheiben–Theorem ist bekannt, dass die inverse zweidimensionale Fouriertransformation von  $M(k_x, k_y)$  die Funktion  $\mu(x, y)$  ergibt (Dössel und Buzug, 2014, S. 80),

$$\mu(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(k_x, k_y) \exp(2\pi i(k_x x + k_y y)) dk_x dk_y. \quad (2.9)$$

Dabei ist  $i$  die imaginäre Einheit mit der Eigenschaft  $i^2 = -1$ . Durch Substitution von  $k_x = k \cos \gamma$  bzw.  $k_y = k \sin \gamma$  erfolgt der Wechsel von kartesischen in polare Koordinaten

$$\mu(x, y) = \int_0^{2\pi} \int_{-\infty}^{\infty} M(k \cos \gamma, k \sin \gamma) \exp(2\pi i k(x \cos \gamma + y \sin \gamma)) k dk d\gamma. \quad (2.10)$$

Durch Aufteilen des äußeren Integrals in die Winkelbereiche  $\gamma = [0, \pi)$  und  $\gamma = [\pi, 2\pi)$  und mit den Symmetrieeigenschaften der Fouriertransformation für reelle Ortsdaten  $M(k, \gamma + \pi) = M(-k, \gamma) = M(k, \gamma)$  kann der Winkelbereich auf  $\gamma = [0, \pi]$  beschränkt werden (Dössel und Buzug, 2014, S. 80 f.). Unter Verwendung des Fourier–Scheiben–Theorems, siehe Gl. (2.7), ergibt sich

$$\mu(x, y) = \int_0^{\pi} \int_{-\infty}^{\infty} P_{\gamma}(k) |k| \exp(2\pi i k \xi) dk d\gamma = \int_0^{\pi} h_{\gamma}(\xi) d\gamma. \quad (2.11)$$

Der erhaltene Ausdruck zeigt das grundlegende Vorgehen der FBP. Die fouriertransformierten Projektionen  $P_{\gamma}(k)$  werden mit der Gewichtsfunktion  $|k|$  im Fourierraum multipliziert. Diese mathematische Operation entspricht einer Filterung. Der Filter  $|k|$  stellt den mathematisch–„natürlichen“ Rampen–Filter dar. Für  $k < 1$  werden die Amplituden der Frequenzen von  $P_{\gamma}(k)$  gedämpft, für  $k > 1$  angehoben. Das innere Integral im mittleren Teil entspricht einer inversen eindimensionalen Fouriertransformation und lässt sich als gefilterte Projektion  $h_{\gamma}(\xi)$  im Radonraum auffassen. Die rechte Seite stellt die Rückprojektion von  $h_{\gamma}(\xi)$  für den Winkelbereich  $\gamma = [0, \pi]$  dar.

**Filter.** Die reale Anwendung der FBP–Rekonstruktion kann mit unterschiedlichen Filterkernfunktionen  $G(k)$  (Filter) erfolgen.

$$G(k) = |k|W(k). \quad (2.12)$$

Dabei ist  $W(k)$  eine wählbare Filterfensterfunktion (Buzug, 2014, S. 192). Das Einsetzen von  $G(k)$  anstatt  $|k|$  in Gl. (2.11) liefert

$$h_{\gamma}(\xi) = \int_{-\infty}^{\infty} P_{\gamma}(k) G(k) \exp(2\pi i k \xi) dk = \int_{-\infty}^{\infty} p_{\gamma}(z) g(\xi - z) dz. \quad (2.13)$$

Zudem lässt sich  $h_\gamma(\xi)$ , äquivalent zu der Multiplikation mit  $G(k)$  im Fourierraum, durch eine Faltungsoperation im Radonraum erhalten. Die Faltung erfolgt mit der Faltungskernfunktion  $g(\xi)$  (Kern), der inversen Fouriertransformierten von  $G(k)$ . (Buzug, 2014, S. 194).

Reale Betrachtungen gehen mit einer Regularisierung von  $G(k)$  einher. Die folgenden Darstellungen orientieren sich an (Buzug, 2014, S. 189 ff.). Die Signale werden auf ein normiertes Intervall  $k \in [-k_\epsilon, k_\epsilon]$  mit  $k_\epsilon = 0,5$  definiert. Nach dem Abtasttheorem von Shannon–Nyquist, Kap. 2.1.2, ergibt sich mit  $k_{max} = 1/(2\Delta\xi) = 0,5$  eine Abtastrate von  $\Delta\xi = 1$ . In Abb. 2.3 werden drei bandbegrenzte Filter mit variierendem  $W(k)$  im Fourierraum (a) und die zugehörigen Kerne im Radonraum (b) vorgestellt. Der Filter nach Ramachandran und Lakshminarayanan (Ram–Lak) entspricht der Bandbegrenzung von  $|k|$ . Das zu starke Anheben hoher Frequenzen wird mit dem Ram–Lak–Filter vermieden, jedoch entstehen aufgrund der scharfen Spitzen bei  $\pm k_\epsilon$  durch die Begrenzung des Frequenzbandes Gibbs–Artefakte (engl.: side lobe artifact) im Bild. Eine kantenbetonende Filteranwendung wird als hart bezeichnet, analog gilt dies für zugehörige Kerne.

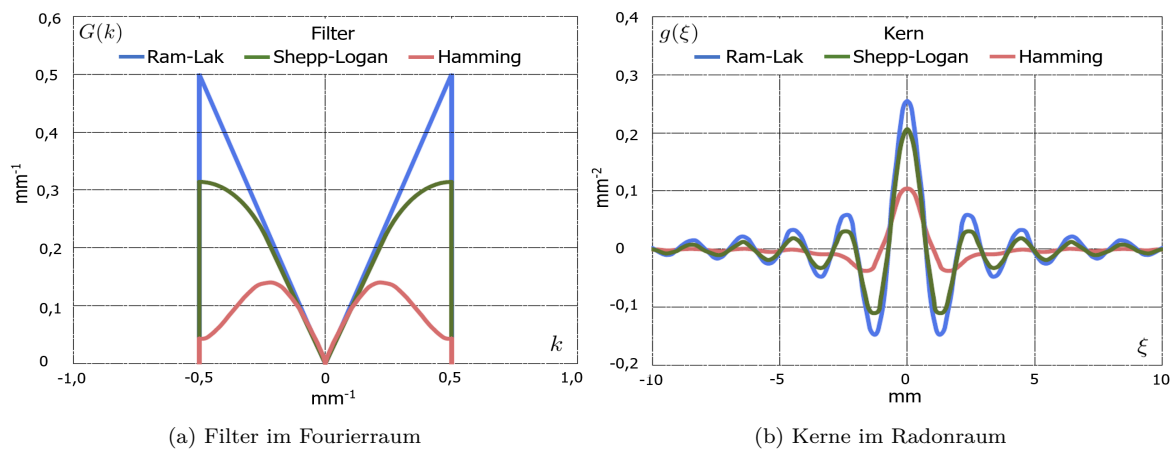


Abb. 2.3: Bandbegrenzter Ram–Lak–, Shepp–Logan– und Hamming–Filter im Fourierraum (links) mit dem jeweils zugehörigen Kern im Radonraum (rechts) für das Abtastintervall  $\Delta\xi = 1$ .

Der Shepp–Logan–Filter glättet die scharfen Spitzen des Ram–Lak–Filters und reduziert Gibbs–Artefakte. Zudem ist der Shepp–Logan–Filter bzw. –Kern in modernen CT–Systemen eine Standardimplementierung. Der stark glättende Hamming–Filter generiert eine hohe Rauschunterdrückung. Filter und Kerne, welche starke Glättungseigenschaften besitzen, werden als weich bezeichnet. Die Darstellung der Kerne im Radonraum zeigt mit zunehmender Glättung ein kleineres und breiteres Maximum mit weniger Oszillationen in den Randbereichen. Formelle Ausdrücke für den Ram–Lak– und Shepp–Logan–Filter bzw. –Kern sind einsehbar in (Buzug, 2014, S. 191 f. u. S. 195 f.) und für den Hamming–Filter in (Huesman et al., 1977, S. 53 f.).



### Iterative Rekonstruktionsmethoden

Das Rekonstruktionsproblem wird für die IR-Rekonstruktion diskret formuliert. Die Lösung des inversen Problems kann z. B. durch die modellbasierte IR-Implementation ADMIRE, Tab. 1.2, S. 11, genähert werden. Die Methode ist ein erweiterter statistischer IR-Ansatz, welcher über ein optimiertes CT-Systemmodell (CT-Geometrie und -Komponenten) verfügt. Zur einfachen Betrachtung der Konstruktion und Ausführungsschritte der modellbasierten IR-Implementation folgen die verfahrensspezifischen Grundlagen. Anschließend folgt eine praxisorientierte Zusammenfassung für den Algorithmus.

**Projektionen.** Aus den Summen der Schwächungskoeffizienten  $\mu_j$  entlang eines Strahls  $i$  ergeben sich die Projektionen  $p_i$  und werden in dem Projektionsvektor  $\mathbf{p}$  angeordnet. Dabei ist  $i = 1, \dots, M$  mit der Anzahl aller Strahlen  $M = N_p D$  aus allen Dixeln  $D$  des Detektors und allen Projektionen  $N_p$  aus allen Winkeln. Die Anordnung von  $\mu_j$  aller zu rekonstruierenden Voxel  $N$  einer Schicht mit  $j = 1, \dots, N$  erfolgt in dem Vektor  $\boldsymbol{\mu}$ . In einer diskreten Rekonstruktionsebene (CT-Schicht) ergibt sich das lineare Gleichungssystem:

$$\mathbf{p} = \mathbf{A}\boldsymbol{\mu}. \quad (2.14)$$

Ein Vergleich mit Gl. (2.4), S. 17, zeigt, dass die Systemmatrix  $\mathbf{A}$  mit den Elementen  $a_{ij}$  auf  $\boldsymbol{\mu}$  wie eine Radontransformation wirkt. Für die Gleichheit von  $N$  mit  $M$  ist das Gleichungssystem durch den Operator  $\mathbf{A}^{-1}$  lösbar (Buzug, 2014, S. 56 f.). Die Lösung ist für idealisierte physikalische Bedingungen exakt bestimmbar. Die Messwerte bei einer CT enthalten Rauschen, das auf der statistischen Natur der Röntgenquanten und der Elektronik beruht. Dies motiviert Näherungsverfahren zur Berechnung von  $\boldsymbol{\mu}$  (Dremel, 2017, S. 25).

**Statistischer iterativer Rekonstruktionsansatz.** Für die Näherungslösung von  $\boldsymbol{\mu}$  aus Gl. (2.14) inkorporieren statistische IR-Methoden die Zählstatistik der Röntgenquanten am Detektor in die Rekonstruktionsprozedur. Die Anzahl der detektierten Quanten  $n$  wird als poissonverteilte Zufallsvariable  $\mathcal{N}$  angenommen. Ausgehend von statistisch voneinander unabhängigen Detektionsprozessen lässt sich die Wahrscheinlichkeit  $P$  der Detektion von  $n$  durch die Poissonverteilung

$$P(\mathcal{N} = n) = \frac{(p\langle n \rangle)^n}{n!} e^{-p\langle n \rangle}, \quad \text{mit } 0 \leq p \leq 1, \quad (2.15)$$

ausdrücken, mit  $p$  als Grundwahrscheinlichkeit und  $\langle n \rangle$  als Erwartungswert von  $n$  (Buzug, 2014, S. 35 f.).

Die Maximum-Likelihood-Methode (ML-M) orientiert an (Buzug, 2014, S. 178 ff.) ist ein statistisches Schätzverfahren, welches den Erwartungswert  $\langle \boldsymbol{\mu} \rangle$  von  $\boldsymbol{\mu}$  aus der Stichprobe der Projektionswerte  $\mathbf{p}$  schätzt. Die Likelihood-Funktion (Likelihood) ist das Produkt der in Gl. (2.15) beschriebenen Verteilung über die Stichprobe  $\mathbf{p}$ . Die Likelihood als Funktion der

variablen Verteilung der Erwartungswerte  $\langle \mu_j \rangle$  zu formulieren, wird durch den Ausdruck

$$I_i \propto \langle n_i \rangle = n_0 \exp\left(-\sum_{j=1}^N a_{ij} \langle \mu_j \rangle\right) \quad (2.16)$$

motiviert. Die Anzahl der Quanten  $n_i$  im Detektor  $i$  ist proportional zu der Röntgenintensität  $I_i$ , damit kann Gl. (2.16) ausgehend von Gl. (2.1) für  $\langle n_i \rangle$  formuliert werden. Dabei ist  $n_0$  die Anzahl der von der Quelle emittierten Photonen. Durch Einsetzen der rechten Seite in die Likelihood entsteht die an die CT angepasste Funktion. Die logarithmierte Likelihood-Funktion (Log-Likelihood) stellt sich dar als,

$$\ln(L(\langle \boldsymbol{\mu} \rangle)) = \sum_{i=1}^M \left( n_i \ln(n_0) - n_i \sum_{j=1}^N a_{ij} \langle \mu_j \rangle - \ln(n_i!) - n_0 \exp\left(-\sum_{j=1}^N a_{ij} \langle \mu_j \rangle\right) \right). \quad (2.17)$$

Das additive Verfahren sichert die Positivität von  $\langle \boldsymbol{\mu} \rangle$ . Die Verteilung mit  $\langle \boldsymbol{\mu} \rangle$ , für welche die Log-Likelihood maximal wird, ergibt die Lösung der ML-M und entspricht am wahrscheinlichsten den gemessenen Werten. Eine Implementation zur Lösung der ML-M in der CT ist der Expectation-Maximization (EM)-Algorithmus. Jede Iteration  $m$  berechnet den Erwartungswert  $\langle \mu_r \rangle^m$  am Bildpunkt  $r$ . Mit  $\langle \mu_r \rangle^m$  wird  $\langle n_i \rangle$  nach Gl. (2.16) berechnet und mit  $n_i$  in der Update-Gleichung  $\langle \mu_r \rangle^{m+1}$  des EM-Algorithmus als gewichteter Quotient verglichen. Die Log-Likelihood wird durch den Quotienten in jeder Iteration maximiert (Buzug, 2014, S. 175 u. S. 181 f.).

Die Kleinste-Quadrate-Schätzung (LS; engl.: Least Squares) nach Darstellung in (Buzug, 2014, S. 185 ff.) ist ein varianter Näherungsansatz für  $\boldsymbol{\mu}$ , welcher das Maximierungsproblem der Log-Likelihood approximiert. Hierfür minimiert die LS-Methode eine Summe von quadrierten Residuen aus allen Projektionen. Der gesuchte Ausdruck, welcher in  $\langle \mu_r \rangle^{m+1}$  eingeht, kann aus Gl. (2.17) durch eine Taylor-Entwicklung mit den erwarteten Projektionssummen

$$\langle p_i \rangle = \sum_{j=1}^N a_{ij} \langle \mu_j \rangle, \quad (2.18)$$

an der Stelle  $p_i = \ln\left(\frac{n_0}{n_i}\right)$  bis hin zur zweiten Ordnung konstruiert werden zu

$$\ln(L(\langle \boldsymbol{\mu} \rangle)) \approx -\frac{1}{2} \sum_{i=1}^M \left( n_i \left( \sum_{j=1}^N a_{ij} \langle \mu_j \rangle - p_i \right)^2 \right) + \sum_{i=1}^M c(n_i). \quad (2.19)$$

Der linke Summand ist das zu minimierende Funktional. Der rechte Summand ist unabhängig von  $\boldsymbol{\mu}$  und kann für die Iterationsprozedur ignoriert werden.

Statistische IR-Methoden, die im Rohdatenraum arbeiten, gewichten die detektierten Röntgenphotonen. Kantenerhaltende zweidimensionale Rauschalgorithmen unterdrücken rauschintensive und verstärken rauscharme Signalbereiche. Das gewichtete Sinogramm wird in die Rekonstruktionsprozedur integriert. Überdies findet die modellierte Zählstatistik der Röntgenquanten im Bildraum zur iterativen Rauschreduktion Anwendung. Die Kombination

des Modells mit einem Regularisierungsfunktional wird in den nachfolgenden Abschnitten beschrieben (Beister et al., 2012).

**Modellbasierter iterativer Rekonstruktionsansatz.** Die modellbasierte IR-Methode ermöglicht die realitätsnahe Modellierung der Akquisitionsprozedur. Hardwarebedingte Limitationen, z. B. Rechenzeiten und Speicherkapazitäten, beschränken die Modellkonstruktion (Dremel, 2017, S. 26). Iterative Verfahren, welche im Algorithmus implementierte Modelle enthalten werden der modellbasierten IR-Methode zugeordnet. Hierzu zählt auch der statistische IR-Ansatz (Beister et al., 2012).

Die Realisierung der in Gl. (2.14), S. 21, vorgestellten Systemmatrix  $\mathbf{A}$  wird als Projektor bezeichnet. Die Elemente  $a_{ij}$  von  $\mathbf{A}$  können als das vom Strahl  $i$  überstrichene Volumenelement des Voxels  $j$  modelliert werden, dies ist für Pixel in Abb. 2.4 veranschaulicht (Buzug, 2014, S. 155).

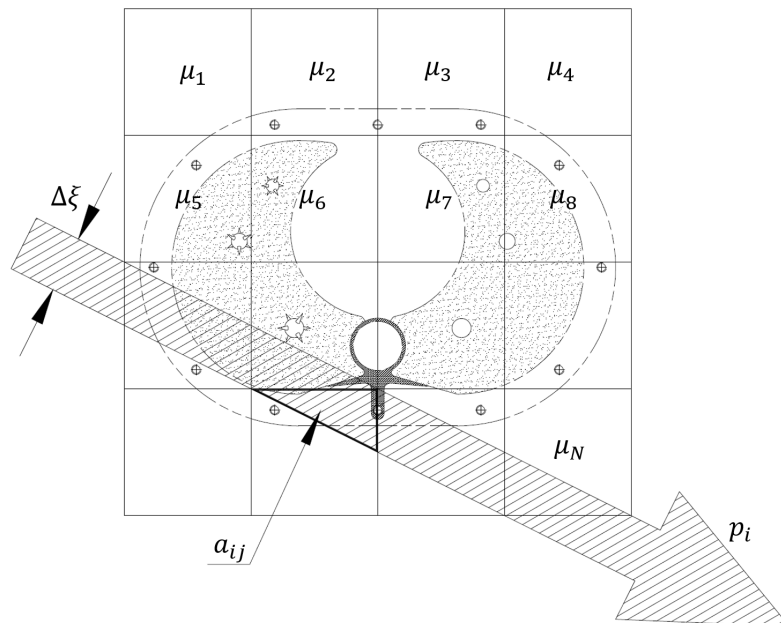


Abb. 2.4: Die zweidimensionale Veranschaulichung zeigt die von einem Röntgenstrahl der Breite  $\Delta\xi$  überstrichene Teilfläche  $a_{ij}$  eines Pixels  $j$  beim Durchlauf durch das Gewebe. Dabei geht  $a_{ij}$  als Gewichtung der Schwächungskoeffizienten  $\mu_j$  in die Berechnung einer Projektion  $p_i$  aus der Strahlsumme entlang des Strahls  $i$  ein.

In der Rekonstruktionsprozedur gewichtet der Vorwärtsprojektor  $\mathbf{A}$  mit den Elementen  $a_{ij}$  die Werte der Schwächungskoeffizienten  $\mu_j$  und es werden Strahlsummen entlang des Strahls  $i$  in dem Projektionsvektor  $\mathbf{p}$  gebildet. Entsprechend ist  $\mathbf{A}^{-1}$  der Rückprojektor, der die Werte des Pseudosinogramms auf das Voxelrastrer von  $\mu_j$  verteilt (Dremel, 2017, S. 25).

Auch andere Eigenschaften des CT-Systems, z. B. die Geometrie, können über die Gewichte in  $\mathbf{A}$  in das Modell der IR eingebracht werden. Je präziser die Modellierung, desto optimaler ist die räumliche Auflösung des zu modellierenden Systems. Zudem wird die numerische Instabilität verringert. Ein Ansatz ist die Modellierung der Fokus- und Dexelflächen mit geeigneten Voxeln. Die Implementation der Modellierungen in die Vorwärtsprojektoren geht

mit der Forderung spezieller Varianten von  $\mathbf{A}$  einher. Dabei wird grundlegend in strahlbasierte und voxelbasierte Projektoren unterschieden (Beister et al., 2012). Die Prinzipdarstellung der Projektoren und Literaturempfehlungen sind im Anhang A.1 aufgeführt.

Modelliertes Vorwissen über das Objekt wird zur Rauschreduktion in die Rekonstruktionsprozedur implementiert. Verfügbare Algorithmen werden als Regularisierungen bezeichnet. Zudem kann eine geeignete Regularisierung die numerische Instabilität der ML–M kompensieren. Es existieren glättende und kantenbetonende Regularisierungen. Letztere sind nichtlineare Verfahren, welche selbst Artefakte generieren können. Stark glättende Regularisierungsalgorithmen führen zu einem überglätteten, artifiziellen Bildeindruck. Die Regularisierungsparameter nehmen zudem Einfluss auf die Konvergenzgeschwindigkeit der IR–Prozedur (Beister et al., 2012).

Im Folgenden wird ein statistischer Ansatz zur Rauschunterdrückung aus (Buzug, 2014, S. 182 ff.) vorgestellt. Dabei sei die Bedingung, dass die Grauwerte von benachbarten Voxeln im Mittel nicht stark abweichen. Ein A–priori–Modell nutzt das Bayes–Schätzverfahren, weiter bezeichnet als Maximum–a–Posteriori (MAP)–Verfahren. Die Methode verwendet zwei statistische Modellierungen. Das erste Modell entspricht der Poissonstatistik der Röntgenquanten. Die Konstruktion des zweiten A–priori–Modells ergibt sich aus mathematisch formuliertem Vorwissen in dem Prior  $R(\langle \boldsymbol{\mu} \rangle)$ ,

$$R(\langle \boldsymbol{\mu} \rangle) = \frac{1}{Z} e^{-\lambda^q \sum_{c \in C} V_c(\langle \boldsymbol{\mu} \rangle)}. \quad (2.20)$$

Das Modell entspricht der Wahrscheinlichkeitsverteilung des Originalbildes. Zudem ist das Bild als Markoffsches Zufallsfeld (MRF) modelliert. Dabei entspricht Gl. (2.20) einer Gibbs–Verteilung mit den lokalen Voxelgruppen (Cliques)  $c$  aus dem Gesamtnachbarschaftsvoxelsystem  $C$  (Deserno et al., 1997, S. 183 f.). Der Regularisierungsparameter  $\lambda^q$  steuert den Regularisierungseinfluss mit  $1 \leq q \leq 2$ . Mit  $Z$  der Normierungskonstante und  $V_c(\langle \boldsymbol{\mu} \rangle)$  der Potentialfunktion von  $c$ . Die Ansätze für Potentialfunktionen zur Rauschunterdrückung sind weitreichend. Für direkt benachbarte Voxel (Nachbarn) mit der Differenz  $\mu_i - \mu_j$  der Cliques innerhalb des Potentials  $V_c = w_{jk} |\mu_j - \mu_k|^q$  ergibt sich mit einheitlicher Gewichtung  $w_{jk} = 1$  für alle Nachbarn und mit  $q = 2$  das Gauß–MRF–Potential. Für den Fall, dass  $w_{jk} = 1$  für orthogonale Nachbarn bzw.  $w_{jk} = 2^{-\frac{1}{2}}$  für diagonale Nachbarn und mit kleinen Werten für  $q$  wird das generalisierte Gauß–MRF–Potential erhalten. Letzteres erzeugt rauschunterdrückte Bilder mit schärferen Kanten.

**Zusammenfassung.** Die Realisierung der ADMIRE-Implementation beruht auf zwei Schleifenroutinen. Die Projektionsschleife (Schleife 1) arbeitet über den Rohdaten- und Bildraum und entfernt hauptsächlich Spiralartefakte (Buzug, 2014, S. 387 f.). Sekundär korrigiert die Routine Rauschen. Letzteres wird hauptsächlich durch die Rauschkorrekturschleife (Schleife 2) im Bildraum korrigiert, welche zudem eine überglättete Bildtextur beschränkt. Die Spezifikationen der modellbasierten IR-Methode sind (Ramirez-Giraldo et al., 2017):

- Der Rauschalgorithmus zur Gewichtung der poissonverteilten Photonen am Detektor.
- Das in die Vorwärtsprojektion implementierte CT-Systemmodell, Schleife 1.
- Das Poisson-Rauschmodell und das Regularisierungsfunktional mit dreidimensionalen Voxelcliquen, Schleife 2.

In Abb. 2.5 sind die einzelnen Routinen des ADMIRE-Algorithmus schematisch dargestellt. Die Schleife 1 berechnet ausgehend von einem Initialvolumen mit der Vorwärtsprojektion ein Pseudosinogramm. Es folgt der Vergleich mit dem gewichteten Sinogramm aus der CT-Messung. Das Abbruchkriterium besteht aus einer vordefinierten Updateanzahl. Ist das Kriterium nicht erfüllt, wird das Update bzw. der nächste Korrekturterm bestimmt und auf die aktuelle Schätzung rückprojiziert. Die Schleife 1 startet erneut mit der Vorwärtsprojektion der aktuellen Schätzung, bis das Abbruchkriterium den Start der Schleife 2 einleitet. Die Optimierung der aktuellen Bilddaten erfolgt mit dem Regularisierungsfunktional und dem Poisson-Rauschmodell bis ein A-priori-Abbruchkriterium erfüllt ist und die Rekonstruktion mit der Ausgabe des Finalvolumens endet.

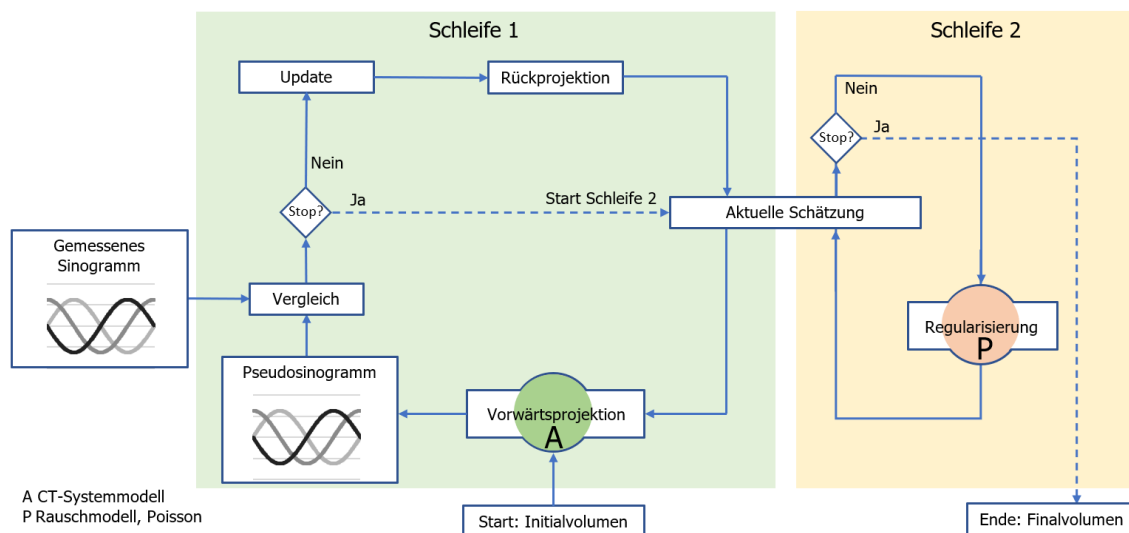


Abb. 2.5: Schema der ADMIRE-Implementation aus der Projektionsschleife (Schleife 1, links) und der Rauschkorrekturschleife (Schleife 2, rechts). Die IR startet in Schleife 1 mit einem Initialvolumen und der Vorwärtsprojektion mit dem CT-Systemmodell. Das berechnete Pseudosinogramm wird iterativ mit dem gemessenen Sinogramm verglichen. Das erzeugte Update wird auf die aktuelle Bildschätzung rückprojiziert. Die nächste Vorwärtsprojektion erzeugt das neue Pseudosinogramm. Dies erfolgt, bis ein Abbruchkriterium Schleife 2 startet. Die Korrektur der geschätzten Bilddaten erfolgt mit einem Regularisierungsfunktional und dem Poisson-Rauschmodell bis das endgültige Abbruchkriterium die Ausgabe des Finalvolumens einleitet.

Die modellbasierte IR-Methode kontrolliert in fünf IR-Stärken mit den Niveaus  $S_1$  bis  $S_5$  das Bildrauschen. Dabei werden die Anzahl der Iterationen im Rohdatenbereich sowie die Rechengeschwindigkeit nicht beeinflusst. Die durch den Hersteller empfohlene Standardeinstellung zur Rauschkorrektur ist das IR-Stärkenniveau  $S_3$ . Für die Rekonstruktion kann ein Kern gewählt werden.

### 2.1.3 Darstellung und Bildqualitätsmaße

#### Hounsfield-Skala und Fenster

Ein Ansatz für die Darstellung der Schwächungskoeffizienten in Graustufen ist die Umrechnung in CT-Zahlen,

$$\text{CT-Zahl} = \frac{\mu - \mu_w}{\mu_w} 1000 \text{ HU}. \quad (2.21)$$

Dabei werden die Schwächungskoeffizienten  $\mu$  mit dem Schwächungskoeffizienten  $\mu_w$  von Wasser in Beziehung gesetzt. Die CT-Zahl wird in Hounsfield-Einheiten (HU; engl.: Hounsfield Units) angegeben. Die 12 Bit-Hounsfield-Skala erstreckt sich über einen Wertebereich von 4069 HU. Für ausgewählte CT-Scanner bieten Hersteller eine 16 Bit-Skala an. Die Zuordnung der CT-Zahlen beginnt bei  $-1000$  HU für Luft, ist gleich  $0$  HU für das Referenzmaterial Wasser und im positiven Wertebereich theoretisch unbegrenzt. Technisch ergibt sich für 12 Bit der limitierte Wertebereich von  $-1024$  HU bis  $3071$  HU (Buzug, 2014, S. 403 f.).

Ausgehend von der Vigilanz des Betrachters und der Helligkeit am Betrachtungsort vermag das menschliche Auge 20 bis 50 Graustufen zu unterscheiden (Hofer, 2000, S. 12). Die Grauwertdiskretisierung an den Endgeräten ist auf 256 Stufen begrenzt. Für die Bildbetrachtung werden Ausschnitte, sog. Fenster, der Hounsfield-Skala verwendet, welche auf dem Grauwertebereich abgebildet werden. Die Abbildung folgt der stückweise definierten Fensterfunktion, wobei das Fenster durch die Fenstermitte (engl.: level) und die Fensterweite (engl.: width) beschrieben wird (Buzug, 2014, S. 407 f.).

**Aufhärungsartefakte.** Das Spektrum der Quelle ist polychromatisch. Ein energieabhängiger Röntgenstrahl wird auf seinem Weg entlang der Orte  $\eta$  durch  $\mu(\xi, \eta, E)$  geschwächt. Durch die Absorption via Photoeffekt wird die mittlere Strahlenenergie erhöht. Misst der Detektor die Photonen ohne Energieauflösung, führt dies zu inkonsistenten Werten von  $p_\gamma(\xi)$ . Der resultierende Bildfehler wird als Aufhärungsartefakt bezeichnet. Eine adäquate HU-Zuordnung erfolgt nach rechnerischer A-priori-Korrektur für Gewebe mit HU nahe dem Referenzmaterial. Die Korrekturinformation kann zum Beispiel einer Wasserphantom-Kalibrationsmessung entnommen werden (Buzug, 2014, S. 404).

#### Bildqualitätsmaße

Die statistische Natur der Röntgenquanten als die physikalische Ursache von Quantenrauschen sowie das thermische Rauschen am Detektor wurden in Kap. 2.1.2 vorgestellt. Das

Signal–Rausch–Verhältnis (SNR; engl.: Signal–to–Noise Ratio) wird in der CT als objektiver Bildqualitätsparameter verwendet und lässt sich als Betrag aus dem Quotienten des Mittelwerts  $\bar{x}_{\text{ROI}}$  mit der zugehörigen Standardabweichung  $\sigma_{\text{ROI}}$  formulieren,

$$\text{SNR} = \left| \frac{\bar{x}_{\text{ROI}}}{\sigma_{\text{ROI}}} \right|. \quad (2.22)$$

Das Suffix ROI (engl.: Region of Interest) bezeichnet eine definierte Region (Buzug, 2014, S. 363). Bei der realen Anwendung unterliegt das SNR strahlenschutzbedingten Limitationen. Für die aus Gl. (2.15) verteilte Zufallsvariable  $n$  gilt für das SNR der Zusammenhang

$$\text{SNR} = \frac{\langle n \rangle}{\sqrt{\langle n \rangle}} = \sqrt{\langle n \rangle}. \quad (2.23)$$

Der Quotient erhöht sich mit der Wurzel aus dem Erwartungswert von  $\langle n \rangle$  (Buzug, 2014, S. 364). Zudem gilt ein proportionaler Zusammenhang des SNR mit der Dosis  $D$ ,  $\text{SNR} \propto \sqrt{D}$ .

Ein weiterer objektiver Bildqualitätsparameter in der CT ist das Kontrast–Rausch–Verhältnis (CNR; engl.: Contrast–to–Noise Ratio). Der Parameter wird zwischen einer gewählten ROI und einer als Referenzregion definierten  $\text{ROI}_{\text{ref}}$  bestimmt. Das CNR ergibt sich mit

$$\text{CNR} = \frac{|\bar{x}_{\text{ROI}} - \bar{x}_{\text{ROI}_{\text{ref}}}|}{\sqrt{\frac{\sigma_{\text{ROI}}^2 + \sigma_{\text{ROI}_{\text{ref}}}^2}{2}}}, \quad (2.24)$$

aus dem Verhältnis der Differenz der Mittelwerte  $\bar{x}_{\text{ROI}}$  und  $\bar{x}_{\text{ROI}_{\text{ref}}}$  zur Wurzel aus der halben Summe der Varianzen  $\sigma_{\text{ROI}}^2$  und  $\sigma_{\text{ROI}_{\text{ref}}}^2$  (Macri et al., 2016).

## 2.1.4 CT–geeignete Dosisbegriffe

### Volumen–CT–Dosisindex

Der CT–Dosisindex (CTDI) zählt zu den CT–spezifischen lokalen Dosisgrößen. Er beschreibt die lokale Dosisexposition in einer nominellen Schichtdicke mitsamt den außerhalb der Schicht liegenden Dosisbeiträgen. Für Mehrschicht–CT–Systeme ist der

$$\text{CTDI} = \frac{1}{N h} \int_{-z_0}^{z_0} D(z) dz \quad (2.25)$$

definiert, mit der nominellen Schichtdicke  $h$  und mit der Anzahl der simultan gescannten Schichten  $N$ . Das Energiedosisprofil  $D(z)$  wird parallel zur Rotationsachse ( $z$ –Achse) des CT–Scanners erfasst (Nagel, 2002, S. 42). Die Energiedosis  $D = d\bar{\epsilon}/dm$  ist der Quotient aus der mittleren Energiedifferenz  $d\bar{\epsilon}$  pro Massenelement  $dm$ . Die SI–Einheit des CTDI bzw. von  $D$  ist Gray,  $\text{Gy} = \text{J}/\text{kg} = \text{m}^2/\text{s}^2$  (Bundesamt für Strahlenschutz, 2007).

Die US–amerikanische Verordnung 21 CFR 1020.30:2011–04 definiert den  $\text{CTDI}_{\text{FDA}}$ , mit den Intervallgrenzen  $z_0 = 7h$  (Food and Drug Administration, 2011). Für die optimierte

Dosismessung für  $h < 7$  mm ist der  $\text{CTDI}_{100}$  mit  $z_0 = 50$  mm definiert worden. Messtechnisch wird für den  $\text{CTDI}_{100}$  der Luftkerma  $K_a(z)$  mit einem zylindrischen Plexiglasphantom (Kopf- oder Rumpfphantom) zentral und peripher erfasst. Dabei ist  $K_a(z)$  die kinetische Energie der erzeugten Ladungsträger pro Masse, welche in der Luft des Detektors durch ionisierende Strahlung freigesetzt wird, angegeben in Gy (Nagel, 2002, S. 75). Die  $\text{CTDI}_{100}$ -Zentrumsmessung ohne Phantom wird als  $\text{CTDI}_{\text{Luft}}$  bezeichnet (Nagel, 2002, S. 8). Durch Addition der lagegewichteten  $\text{CTDI}_{100}$ -Messwerte aus dem Zentrum c und der Peripherie p des verwendeten Phantoms wird mit

$$\text{CTDI}_w = \frac{1}{3} \text{CTDI}_{100,c} + \frac{2}{3} \text{CTDI}_{100,p} \quad (2.26)$$

der gewichtete  $\text{CTDI}_w$  bestimmt (Leitz et al., 1995). An der Konsole von Standard-CT-Systemen wird mit  $\text{CTDI}_w$  ein pitchkorrigierter nach DIN EN 60601-2-44:2017-03 genormter

$$\text{CTDI}_{\text{vol}} = \frac{\text{CTDI}_w}{P}, \quad \text{mit } P = \frac{s_{\text{rot}}}{Nh}, \quad (2.27)$$

berechnet, bezeichnet als Volumen-CT- oder effektiver CT-Dosisindex. Für Mehrzeilen-CT-Messungen berücksichtigt der  $\text{CTDI}_{\text{vol}}$  den Effekt der Überlappung mehrerer Dosisprofile (Deutsches Institut für Normung, 2017). Der Pitchfaktor P für Mehrzeilen-CT-Systeme ist der Quotient aus dem Tischvorschub pro Röhrenrotation  $s_{\text{rot}} = v_t T_{\text{rot}}$  und der Gesamtkollimation  $Nh$ . Dabei ist  $v_t$  die Tischvorschubgeschwindigkeit und  $T_{\text{rot}}$  die Dauer einer Röhrenrotation. Durch die Pitchkorrektur wird der Einfluss von  $P \neq 1$  auf die Dosis berücksichtigt (Nagel, 2002, S. 10).

### Dosis-Längen-Produkt

Das Dosis-Längen-Produkt (DLP) ist eine integrale Dosisgröße und ein Maß für die Gesamtdosisexposition des abgescannten Körperabschnitts. Für Mehrschicht-CT-Systeme formuliert sich das DLP mit

$$\text{DLP}_v = \text{CTDI}_v nhN. \quad (2.28)$$

Das Suffix v variiert für entsprechende Definitionen des  $\text{CTDI}$  bzw. des DLP. Geeignete lokale Dosisgrößen zur Berechnung des  $\text{DLP}_v$  sind der  $\text{CTDI}_w$  oder der  $\text{CTDI}_{\text{Luft}}$ . Zudem ist  $n$  die Anzahl der Röhrenumläufe bei der Spiral-CT-Technik bzw. die Anzahl der Tischvorschübe bei der sequentiellen CT-Technik,  $h$  ist die nominelle Schichtdicke und  $N$  die Anzahl der simultan gescannten Schichten. Das DLP wird in mGy cm angegeben (Nagel, 2002, S. 42). Die Bestimmung des DLP aus dem Volumen-CT-Dosisindex  $\text{CTDI}_{\text{vol}}$  erfolgt über

$$\text{DLP}_w = \text{CTDI}_w nhN = \text{CTDI}_{\text{vol}} L. \quad (2.29)$$

Für den mittleren Teil von Gl. (2.29) wird der pitchkorrigierte  $\text{CTDI}_w$  der rechten Seite aus Gl. (2.27) in Gl. (2.28) eingesetzt. Um keine doppelt berücksichtigten Faktoren zu erhalten, muss die Pitchkorrektur des  $\text{CTDI}_w$  durch Multiplikation mit P revidiert werden (Nagel, 2002, S. 12).



Alternativ kann das DLP mit dem Produkt des Röhrenstroms  $I$ , mit der Gesamtscanzeit  $T$  für einen Körperabschnitt sowie dem normierten  ${}_n\text{CTDI}_V$ -Wert bestimmt werden aus

$$\text{DLP}_{I_V} = {}_n\text{CTDI}_V I T h N, \quad \text{mit } {}_n\text{CTDI}_V = \frac{\text{CTDI}_V}{Q}. \quad (2.30)$$

Für die Methode können die folgenden lokalen Dosisgrößen zur Anwendung kommen:  $\text{CTDI}_{\text{Luft}}$ ,  $\text{CTDI}_{\text{FDA}}$  und  $\text{CTDI}_{100}$ . Für das CT-Experiment bezüglich der  $\text{CTDI}_V$ -Messwerte wird ein konstantes Röhrenstrom-Zeit-Produkt  $Q$  von 100 mAs bzw. 1 mAs geschaltet. Der  ${}_n\text{CTDI}_V$  wird durch das Verhältnis des  $\text{CTDI}_V$  und  $Q$  erhalten (Nagel, 2002, S. 40).

### Effektive Dosis

Die effektive Dosis  $E$  unterliegt als eine personenbezogene Dosisgröße der Strahlenschutzempfehlung ICRP 103 (Bundesamt für Strahlenschutz, 2007). Dabei ist

$$E = \sum_{i=1}^k w_{T,i} D_{T,i}, \quad \text{mit } D_T = w_R D_{T,R}, \quad (2.31)$$

die Summe von  $i = 1, \dots, k$  gewichteten Organdosen  $D_{T,i}$  für die Anzahl von  $k$  festgelegten Organen bzw. Geweben (T; engl.: Tissue). Der Gewebewichtungsfaktor  $w_T$  gilt unabhängig von Lebensalter und Geschlecht. Der relative Beitrag einzelner Organe bzw. Gewebe an dem durch stochastische Effekte zustande kommenden Strahlengesamtschaden wird durch  $w_T$  quantifiziert. Die Werte für  $w_T$  sind einsehbar in (Bundesamt für Strahlenschutz, 2007, S. 61).  $D_T$  ist die Äquivalentdosis in einem Organ oder Gewebe für eine bestimmte Strahlung (R; engl.: Radiation). Für ein Organ oder Gewebe ist  $D_T$  gleich der Multiplikation des Strahlungswichtungsfaktors  $w_R$  mit der mittleren Energiedosis  $D_{T,R}$ . Explizit ist  $D_T$  der Quotient der auf ein Organ oder Gewebe T übertragenen mittleren Energie  $\epsilon_T$  und der zugehörigen Masse  $m_T$  von T (Bundesamt für Strahlenschutz, 2007). Der Betrag von  $w_R$  beruht auf der relativen biologischen Wirksamkeit einer Strahlung. Wertetabellen hat das Bundesamt für Strahlenschutz veröffentlicht (Bundesamt für Strahlenschutz, 2007, S. 60). Die SI-Einheit von  $E$  bzw.  $D_T$  ist Sievert,  $\text{Sv} = \text{J/kg} = \text{m}^2/\text{s}^2$ .

Überdies kann  $E$  nach einer Methode der „European guidelines on quality criteria for computed tomography“ (Europäische Union, 2000) aus

$$E = \text{DLP}_w f \quad (2.32)$$

geschätzt werden. Der Konversionsfaktor  $f$  bezieht sich auf einen Körperabschnitt und wird geschlechterunspezifisch in Abhängigkeit vom Lebensalter gewählt. Die Werte von  $f$  resultieren aus dem Vergleich einer effektiven Dosis  $E$  aus Monte-Carlo-simulierten  $D_T$ -Koeffizienten mit  $\text{DLP}_w$ -Werten entsprechender klinischer CT-Experimente. Die Public Health England (London, Vereinigtes Königreich) stellte die Datenbasis für die Monte-Carlo-Simulationen zur Verfügung. Eine Übersicht von  $f$  hat die American Association of Physicists in Medicine veröffentlicht (American Association of Physicists in Medicine, 2008, S. 13).

### 2.1.5 Ansätze zur Dosisreduzierung

Die CT–Hersteller–Realisierungen zur Dosisreduktion sind weitreichend. Im Folgenden werden ausschließlich Ansätze der Firma Siemens Healthineers (Forchheim, Deutschland) vorgestellt. Ein dosissparender DL–Softwareansatz zur CT–Bildrekonstruktion der Firma GE Healthcare (Chicago, Vereinigte Staaten von Amerika) kann eingesehen werden in (Hsieh et al., 2019).

Die Reduktion der Strahlendosis durch geeignete Soft– und Hardware kann seitens des Anwenders durch Einhaltung von Strahlenschutzmaßnahmen erhöht werden. Letztere beziehen sich auf die Patientenvorbereitung, Zentrierungs– und Lagerungsaspekte. Dabei gilt es bei einer aktivierten Röhrenstrommodulation (*I*–Mod) metallische Gegenständen im Scanbereich zu vermeiden (Euler et al., 2014). Die exakte Patientenzentrierung im Isozentrum verhindert Bildqualitätseinbußen durch den ineffizienten Einsatz von röhrennahen, auf das Isozentrum ausgerichteten, Formfiltern. Überdies werden erhöhte Dosisbeiträge bei eingeschalteter *I*–Mod verhindert (Li et al., 2007). Durch die Über–Kopf–Lagerung der Arme bei der CT des Thorax bzw. des Abdomens wird die Photoabsorption im Untersuchungsbereich gesenkt sowie der Röhrenstrom bei aktiver *I*–Mod. Weitere klinische Ansätze sind die Scanlängenreduktion, die Auswahl geeigneter Scanparameter, kombinierte Bildgebungsmethoden, der Einsatz von Strahlenschutzmitteln und die Einsparung von Teiluntersuchungen mit Dual–Energie–CT–Techniken (Euler et al., 2014).

Zu den Software–Implementationen zählen Modulationstechniken und die IR–Rekonstruktion (Siemens Healthineers, 2020). Die *I*–Mod wird mit dem Ziel eingesetzt, die Dosis für den Patienten bei gleichbleibender Bildqualität zu senken. Hierbei wird der Röhrenstrom an die Schwächungs– und Formeigenschaften einer zu untersuchenden Körperregion angepasst. CARE Dose4D ist eine *xyz*–Modulation und kombiniert die gelisteten Ansätze.

- Die Längs– bzw. *z*–Modulation. Durch den Vergleich des CT–Topogramms mit dem eines Standardmodells erfolgt in *z*–Richtung die Anpassung des Röhrenstroms an die Formeigenschaften einer Körperregion. Anatomische Formvarianten innerhalb einer Untersuchungsregion können hinsichtlich erhöhter Dosisbeiträge kompensiert werden. Der zu vergleichende Bildqualitätsparameter bei der Modulation ist ein Qualitäts–Referenz–Röhrenstrom–Zeit–Produkt.
- Die Winkel– bzw. *xy*–Modulation. Basierend auf einer Echtzeitmessung während des CT–Experiments wird in der *xy*–Ebene für jede Winkeleinstellung der Quellen–Detektor–Kombination das Schwächungsprofil hinter dem Patienten bestimmt. Anhand der erhobenen Profile wird für die nachfolgende Rotation, verzögert um eine halbe Rotation, die Anpassung des Röhrenstroms vorgenommen.

Ein organbasierter *I*–Mod Ansatz ist X–CARE. Hierbei erfolgt in der kranialen und thoralen Bildgebung eine Senkung des Röhrenstroms für einen ventralen Rotationswinkel von  $120^\circ$  zum Schutz der Augenlinsen und des Brustdrüsengewebes (Euler et al., 2014). Mit CARE kV moduliert der Hersteller die Röhrenspannung anhand eines topogrammbasierten

Patientendurchmesserprofils, der Systemleistung und des gewählten CT-Dosisprotolls. Die IR-Rekonstruktion arbeitet effizient bei CT-Messungen mit geringer Anzahl von detektierten Röntgenphotonen und geringen Signalpegeln (Beister et al., 2012). Grundlegend wird zwischen algebraischer Rekonstruktion und modellbasierter IR-Rekonstruktion unterschieden. Die zuletzt entwickelte ADMIRE-Implementation reduziert für ein festes CT-Dosisprotokoll die Patientendosis um bis zu 60% gegenüber der FBP-Rekonstruktion bei vergleichbarer Bildqualität (Siemens Healthineers, 2020).

Dosissparende Hardwareansätze bilden sowohl die dynamische Kollimatorsteuerung Adaptive Dose Shield zur Vermeidung von Dosisbeiträgen durch verlängerte Scanbereiche (engl.: over-ranging) bei der Spiral-CT-Technik als auch signaloptimierte Detektortechnologien (Siemens Healthineers, 2020).

## 2.2 Methodische Grundlagen der Statistik

### 2.2.1 Maßzahlen einer Stichprobe

Das arithmetische Mittel, sog. Mittelwert,  $\bar{x}$  einer Stichprobe vom Umfang  $n$  und den Stichprobenwerten  $x_1, x_2, \dots, x_n$  ist der Quotient aus der Summe der Stichprobenwerte und  $n$ :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (2.33)$$

Zudem bestimmt  $\bar{x}$  als Lageparameter die durchschnittliche Lage von  $x_1, x_2, \dots, x_n$  auf einer Zahlengeraden (Papula, 2016, S. 486).

Die Standardabweichung  $s$  der Stichprobe wird gebildet aus der Quadratwurzel des Quotienten aus den Abweichungsquadraten  $v_i^2 = (x_i - \bar{x})^2$  und der Zahl  $n - 1$ ,

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n v_i^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}. \quad (2.34)$$

Quadrieren der Gl.(2.34) liefert die Varianz der Stichprobe  $s^2$ . Beide Maßzahlen repräsentieren Streuungsmaße von  $x_1, x_2, \dots, x_n$  um  $\bar{x}$ . Sind alle Werte der Stichprobe gleich, gilt  $s = 0$  bzw.  $s^2 = 0$  (Papula, 2016, S. 487).

Die Standardabweichung des Mittelwerts  $s_{\bar{x}}$  wird über den Ausdruck

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} \quad (2.35)$$

bestimmt. Die rechte Seite der Gl. (2.35) zeigt die Division von  $s$  durch die Quadratwurzel von  $n$ . Dabei wird  $s_{\bar{x}}$  als Maß für die Genauigkeit von  $\bar{x}$  betrachtet (Papula, 2009, S. 294).

### 2.2.2 $t$ -Test für zwei gepaarte Stichproben

Der  $t$ -Test zählt zu den statistischen signifikanzprüfenden Hypothesentestverfahren. Als parametrischer Lagetest vergleicht die Methode zwei Mittelwerte zu einem gewählten Signifikanzniveau  $\alpha$ . Die Modellannahmen sind die Normalverteilung der zugrunde liegenden Zufallsvariablen in der Grundgesamtheit und eine vorliegende Varianzhomogenität (Rasch, 2008, S. 59). Die zu berechnende, für das Testproblem sensible Prüfgröße ist die Teststatistik  $t$ . Die Prüfverteilung ist die  $t$ -Verteilung von Student (Hartung et al., 2009, S. 154 f.).

Im Fall eines  $t$ -Tests für zwei gepaarte Stichproben gelten die verfahrensspezifischen Modellannahmen: zwei gepaarte Stichproben mit dem Umfang  $n$  und den Wertepaaren  $(x_i, y_i)$ , die Normalverteilung der Differenzen  $d_i = x_i - y_i$  innerhalb der Grundgesamtheit und die Intervallskalierung der Zufallsvariablen zur Differenzenbildung (Rasch, 2008, S. 89 ff.; Weiß, 2013, S. 181).

Die Hypothesen des Verfahrens sind:

- $H_0$ : Es liegt kein hypothetischer Unterschied zwischen den Mittelwerten der Stichproben vor, die mittlere Differenz der gepaarten Stichprobenwerte  $\bar{d}$  unterliegt dem Zufall.
- $H_1$ : Es liegt ein hypothetischer Unterschied vor,  $\bar{d}$  unterliegt nicht dem Zufall.

Vorab wurde  $H_1$  als zweiseitige Fragestellung ohne Kenntnis einer möglichen Richtung des angenommenen Unterschiedes formuliert. Liegt hingegen Richtungkenntnis über den Unterschied zwischen den Stichprobenmittelwerten vor, kann eine einseitige Fragestellung angewendet werden (Weiß, 2013, S. 181).

Für die Berechnung von  $t$  wird zunächst  $\bar{d}$  der gepaarten Stichprobenwerte berechnet,

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n (x_i - y_i), \quad \text{mit } i = \{1, \dots, n\}. \quad (2.36)$$

Durch Einsetzen von  $\bar{d}$  aus Gl. (2.36) in die Formulierung

$$s_d = \sqrt{\frac{1}{n-1} \sum_{i=1}^n ((x_i - y_i) - \bar{d})^2}, \quad (2.37)$$

ergibt sich die Standardabweichung  $s_d$  der Differenzen  $d_i$  (Rasch, 2008, S. 90). Mit  $\bar{d}$  und  $s_d$  lässt sich  $t$  berechnen über

$$t = \frac{\sqrt{n}}{s_d} (\bar{d} - \delta_0). \quad (2.38)$$

Dabei ist  $\delta_0$  ein vorgegebener Wert, gegen den  $\bar{d}$  getestet wird. Für  $H_0$  gilt  $\delta_0 = 0$ . Der Nenner gibt den geschätzten Standardfehler des Mittelwerts der Grundgesamtheit an (Papula, 2009, S. 293 f.; Weiß, 2013, S. 181).

Zur Überprüfung von  $H_0$  werden die Quantile  $t_{n;\alpha}$  der Prüfverteilung für  $n-1$  Freiheitsgrade zu dem gewählten Signifikanzniveau  $\alpha$  bestimmt. Dabei wird  $H_0$  abgelehnt, wenn  $|t| > t_{n-1;1-\alpha/2}$  oder  $|t| > t_{n-1;1-\alpha}$  bei zweiseitiger bzw. einseitiger Fragestellung (Weiß, 2013, S. 181). Die Werte für  $t_{n;\alpha}$  lassen sich mit Gl. (2.41) ermitteln. Berechnungsoptionen stellen kommerzielle Statistik- und Analysesoftwarepakete dar. Ein Beispiel ist das Programm SPSS (International Business Machines Corporation, 2020), Version Statistics 27. Alternativ können mit  $\alpha$  und den Freiheitsgraden für die Prüfverteilung die Quantile aus den Tabellen zugehöriger Fachliteratur erhoben werden (Weiß, 2013, S. 297 ff.).

### 2.2.3 Friedman-Test

Der Friedman-Test als signifikanzprüfender Hypothesentest zu einem Signifikanzniveau  $\alpha$ , entspricht einer Rangvarianzanalyse für die Anzahl  $k$  abhängiger Stichproben (Proben) mit dem Umfang  $n$ . Die nicht parametrische Methode setzt für die Lageparameteruntersuchung keine bekannte Verteilung der Zufallsvariablen in der Grundgesamtheit voraus. Somit ist die Methode auch auf kleine  $n$  anwendbar. Die unterliegende Prüfverteilung weist unter  $H_0$  eine asymptotisch genäherte Chi-Quadrat-Verteilung  $\chi^2$  mit  $q-1$  Freiheitsgraden auf. Abweichungen bezüglich der Prüfverteilung existieren für kleine Werte von  $k$  bzw.  $n$  der Stichproben (Bortz und Lienert, 2008, S. 204). Die Prüfgröße ist die als Friedman-Wert bezeichnete approximierte Teststatistik  $\chi_r^2$ . Der  $\chi_r^2$ -Wert wird nicht direkt aus den Zufallsvariablen bestimmt, sondern aus den Rangsummen  $R_j$  eines vorab durchgeführten Rangordnungstests, Kap. 2.5.4. Das Datenschema für die vergebenen Ränge  $x_{ij}$  für eine Anzahl von  $n = N$  Individuen zeigt Tab. 2.1 mit  $i = 1, \dots, N$  Zeilen und  $j = 1, \dots, k$  Spalten (Bortz und Lienert, 2008, S. 206).

Tab. 2.1: Rangdatenschema für den Friedman-Test zur Rangvarianzanalyse von  $N$  Individuen unter  $k$  Proben.

		$k$ Proben			
		1	2	...	$k$
N Individuen	1	$x_{11}$	$x_{12}$	...	$x_{1k}$
	2	$x_{21}$	$x_{22}$	...	$x_{2k}$
	...	...	...	...	...
	$N$	$x_{N1}$	$x_{N2}$	...	$x_{Nk}$
Rangsummen		$R_1$	$R_2$	...	$R_k$

Die zugehörigen Hypothesen lauten (Bortz und Lienert, 2008, S. 203):

- $H_0$ : Die durchschnittlichen Rangsummen der  $N$  Individuen unter der Anzahl von  $k$  Proben unterscheiden sich nicht. Die Rangvarianzen unterliegen dem Zufall.
- $H_1$ : Die durchschnittlichen Rangsummen der  $N$  Individuen unter der Anzahl von  $k$  Proben unterscheiden sich mindestens für zwei Proben. Die Rangvarianzen unterliegen nicht dem Zufall.

Mit  $R_j$ ,  $n$  und  $k$  lässt sich die Teststatistik über

$$\chi_r^2 = \frac{12}{n k(k+1)} \sum_{j=1}^n R_j^2 - 3 n(k+1) \quad (2.39)$$

bestimmen. Im Fall verbundener Ränge muss  $\chi_r^2$  korrigiert werden. Die Korrekturvorschrift ist einsehbar in (Bortz und Lienert, 2008, S. 205). Für die Signifikanzprüfung wird das Quantil  $\chi_{r(q;\alpha)}^2$  der Prüfverteilung mit  $q-1$  Freiheitsgraden und dem festgelegten Signifikanzniveau  $\alpha$  über Gl. (2.41) bestimmt. Anschließend erfolgt der Vergleich der Teststatistik mit  $\chi_{r(q;\alpha)}^2$  bei einseitiger Fragestellung.  $H_0$  wird abgelehnt bzw.  $H_1$  angenommen für  $\chi_r^2 > \chi_{r(q-1;1-\alpha)}^2$  (Papula, 2009, S. 460). Überdies können kommerzielle Softwareansätze, Kap. 2.2.2, zur Berechnung von  $\chi_r^2$  bzw. Literaturtabellen für  $\chi_{r(q;\alpha)}^2$ , einsehbar in (Weiß, 2013, S. 304 f.), verwendet werden.

## 2.2.4 Überschreitungswahrscheinlichkeit

Als Alternative zur Überprüfung von  $H_0$  bei signifikanzprüfenden Hypothesentests kann die Überschreitungswahrscheinlichkeit  $p$  herangezogen werden. Dabei ist  $p$  die Wahrscheinlichkeit, unter  $H_0$ , den Prüfgrößenwert oder einen in Richtung der Alternative extremeren Wert zu erhalten. Das Verfahren vergleicht den  $p$ -Wert mit dem Signifikanzniveau  $\alpha$ . Ist  $p$  kleiner als  $\alpha$ , wird  $H_0$  abgelehnt und  $H_1$  angenommen (Fahrmeir, 2004, S. 419).

Um  $p$  zu bestimmen, erfordert dies Kenntnisse über die unterliegende Prüfverteilung bzw. Wahrscheinlichkeitsdichtefunktion  $f(x)$  der Testsituation. Dabei ist  $f(x)$  die Wahrscheinlichkeitsdichte einer stetigen Zufallsvariablen  $X$  mit dem Wertebereich  $-\infty < X < \infty$ . Die Wahrscheinlichkeit  $P$ , dass  $X$  einen Wert kleiner oder gleich einer reellen Zahl  $x$  annimmt ist

$$P(X \leq x) = \int_{-\infty}^x f(x') dx', \quad (2.40)$$

mit  $x'$  als Integrationsvariable. Zudem ist  $P(X \leq x)$  die Fläche unter  $f(x)$ . Die Gesamtfläche im Intervall  $-\infty < x < \infty$  ist gleich 1 (Weiß, 2013, S. 106). Weitere Eigenschaften der Dichtefunktion können in (Papula, 2016, S. 330) eingesehen werden. Der Annahmebereich  $\varepsilon$  für  $H_0$  ist gegeben mit  $\varepsilon = 1 - \alpha$ . Durch Gleichsetzen von  $\varepsilon$  mit der Wahrscheinlichkeit  $P$ ,

dass  $X$  einen Wert zwischen  $-x$  und  $x$  annimmt,

$$P(-x \leq X \leq x) = \int_{-x}^x f(x') dx' = \varepsilon, \quad (2.41)$$

wird durch Umformen und Auflösen das Quantil  $x_{(n;\alpha)}$  für das gewählte Signifikanzniveau  $\alpha$  aus  $f(x)$  mit  $n$  Freiheitsgraden erhalten (Papula, 2016, S. 528). Dabei gilt Gl. (2.41) für symmetrische und nicht symmetrische  $f(x)$  und liefert für einseitige Fragestellungen das  $(1 - \alpha)$ -Quantil. Bei zweiseitiger Fragestellung kann für symmetrische  $f(x)$  das  $(1 - \frac{\alpha}{2})$ -Quantil durch Gleichsetzen des zugehörigen Werts für  $\varepsilon$  bestimmt werden (Weiß, 2013, S. 181).

Abschließend erfolgt unter Betrachtung von Gl. (2.40) und Einsetzen von  $x_{(n;\alpha)}$  in

$$P_{H_0}(X > x_{(n;\alpha)}) = 1 - P_{H_0}(X \leq x_{(n;\alpha)}) = p \quad (2.42)$$

die Berechnung von  $p$ . In der realen Anwendung ist  $X$  diskret, die äquivalenten Formulierungen zur stetigen Zufallsvariablen sind dargestellt in (Fahrmeir, 2004, S. 306).

## 2.2.5 Kappa-Koeffizient nach Cohen

Der Kappa-Koeffizient  $\kappa$  nach (Cohen, 1960) wird zur Untersuchung der Übereinstimmung kategorialer Beurteilungen von  $N$  Objekten herangezogen. Dabei repräsentiert  $\kappa$  ein Gütemaß für den Grad der Beurteilungsübereinstimmung (Urteilskonkordanz). Die Merkmale, nach denen kategorial beurteilt wird, können dabei zwei- oder mehrfach gestuft sein (Bortz und Lienert, 2008, S. 310). Unterschieden wird zwischen der Interrater-Reliabilität, d.h. der Urteilskonkordanz der kategorialen Beurteilungen von zwei Bewertern sowie nur eines Bewerters zu unterschiedlichen Zeitpunkten, der sog. Intrarater-Reliabilität (Kwiecien et al., 2011).

Die Bewertungen der  $N$  Objekte bilden die Häufigkeiten  $f_{ij}$  verteilt auf  $k$  Merkmalskategorien einer Kontingenztabelle mit  $i = 1, \dots, k$  Zeilen bzw.  $j = 1, \dots, k$  Spalten. Der tabellarische Aufbau für die Prüfung der Interrater-Reliabilität der Bewerter A und B bzw. für die Intrarater-Reliabilität eines Urteilers zum Zeitpunkt A und B zeigt Tab. 2.2, S. 36. Alle der Hauptdiagonalen zugehörigen  $f_{ij}$  der Kontingenztabelle repräsentieren die konkordanten Bewertungen, allen außerdiagonalen  $f_{ij}$  sind diskordante Bewertungen zuzuordnen. Die Randhäufigkeiten  $f_{\bullet j}$  in jeder Spalte bzw.  $f_{i \bullet}$  in jeder Zeile werden mit

$$f_{\bullet j} = \sum_{i=1}^k n_{ij} \quad \text{bzw.} \quad f_{i \bullet} = \sum_{j=1}^k n_{ij} \quad (2.43)$$

durch Aufsummieren der entsprechenden  $f_{ij}$  gebildet (Kwiecien et al., 2011).

Tab. 2.2: Cohen's  $\kappa$  Kontingenztabelle für  $k \geq 2$  Merkmalskategorien zur Bestimmung der Interrater-Reliabilität bzw. Intrarater-Reliabilität.

		Bewerter/ Zeitpunkt A				Summe	
		$k \geq 2$					
Bewerter/ Zeitpunkt B	$k \geq 2$	1	2	...	$k$		
		1	$f_{11}$	$f_{12}$	...	$f_{1k}$	$f_{1\bullet}$
		2	$f_{21}$	$f_{22}$	...	$f_{2k}$	$f_{2\bullet}$
		...	...	...	...	...	...
		$k$	$f_{k1}$	$f_{k2}$	...	$f_{kk}$	$f_{k\bullet}$
Summe		$f_{\bullet 1}$	$f_{\bullet 2}$	...	$f_{\bullet k}$	$N$	

Die Summe der  $f_{ij}$  entlang der Hauptdiagonalen dividiert durch  $N$  ergibt mit

$$B = \frac{1}{N}(f_{11} + f_{22} + \dots + f_{kk}) \quad (2.44)$$

den Anteil der übereinstimmenden Beurteilungen  $B$ . Die zufällig zu erwartenden konkordanten Bewertungen  $E$  berechnen sich durch das Aufsummieren der paarweise multiplizierten Zeilen und Spalten mit  $i = j = 1, \dots, k$ , dividiert durch  $N^2$ ,

$$E = \frac{1}{N^2}(f_{\bullet 1}f_{1\bullet} + f_{\bullet 2}f_{2\bullet} + \dots + f_{k\bullet}f_{\bullet k}). \quad (2.45)$$

Dabei geben die multiplizierten Werte im Zähler an, wie viele Übereinstimmungen der jeweiligen Merkmalskategorie zu erwarten wären, wenn die Bewerter rein zufällig geurteilt hätten bzw. die Bewertungen stochastisch unabhängig wären. Durch Einsetzen von Gl. (2.44) und Gl. (2.45) in die Beziehung

$$\kappa = \frac{B - E}{1 - E}, \quad (2.46)$$

wird  $\kappa$  bestimmt. Aus Gl. (2.46) geht hervor, dass  $\kappa$  den Anteil an Übereinstimmungen quantifiziert, der über den zufällig zu erwartenden Anteil hinausgeht (Weiß, 2013, S. 266 f.). Des Weiteren ist  $\kappa = 1$  mit  $B = 1$  für ausschließlich übereinstimmende Bewertungen über alle  $N$ . Im Fall von  $\kappa = 0$  gilt  $E = B$ . Zudem kann  $\kappa$  negativ werden, dabei beträgt der größte negative Wert  $-1/(k - 1)$ . Bei negativem  $\kappa$  ist  $B < E$ . Sollten beide Bewerter alle  $N$  Objekte derselben Merkmalskategorie zuordnen, ist  $E = 1$  und  $\kappa$  nicht definiert (Bortz und Lienert, 2008, S. 311).

Für die Interpretation von  $\kappa$  existieren in der Literatur verschiedene Vorschläge. In Tab. 2.3, S. 37, wird ein Vorschlag aus (McHugh, 2012) für die medizinische Forschung vorgestellt.



Tab. 2.3: Interpretation des Cohen's  $\kappa$ -Koeffizienten für die Bewertung der Urteilskonkordanz kategorialer Bewertungen, vorgeschlagen für den medizinischen Forschungsbereich.

$\kappa$ -Wert	Übereinstimmungsgrad
0 – 0.20	keine
0.21 – 0.39	minimal
0.40 – 0.59	schwach
0.60 – 0.79	moderat
0.80 – 0.90	stark
> 0.90	nahezu perfekt

**Modifizierter Kappa-Koeffizient.** Cohen's  $\kappa$  weist Eigenschaften auf, die seine Interpretierbarkeit erschweren (Grouven et al., 2007). Für  $\kappa$  wird der Anteil an übereinstimmenden Bewertungen  $B$  quantifiziert, der über die zufällig zu erwartenden Übereinstimmungen  $E$  hinausgeht. Dabei wird  $E$  als Korrekturterm bezeichnet. Der Anlass zur kritischen Betrachtung der Zuverlässigkeit von zufallskorrigierten Übereinstimmungsprüfungen wird motiviert durch (Guggenmoos-Holzmann, 1993). Die Arbeit berichtet, dass die zufallskorrigierte Methode paradoxe und kontraintuitive Ergebnisse liefern kann. Mögliche Fehleinschätzungen von  $\kappa$  werden in (Wirtz und Kutschmann, 2007) im Rahmen der Analyse von Bewerterübereinstimmungen vorgestellt.

In (Kutschmann, 2020) wird über das Paradoxon variabler  $\kappa$ -Werte bei gleichem  $B$  in der Kontingenztabelle berichtet, wenn die Häufigkeiten  $f_{ij}$  außerdiagonal konstant verteilt sind, jedoch in der Hauptdiagonalen variieren. Zudem erfolgt bei  $B = 0$  eine Zufallskorrektur, da für die Berechnung von  $E$  in Gl. (2.45) die Randhäufigkeiten außerdiagonale  $f_{ij}$  besitzen. Über weitere Paradoxa, unter anderem hinsichtlich balancierter und nicht balancierter Randhäufigkeiten, wird in (Feinstein und Cicchetti, 1990) berichtet. Zur Vermeidung der vorgestellten Paradoxa wird in (Kutschmann, 2020) vorgeschlagen, den Korrekturterm  $E$  in Gl. (2.46) durch die Einführung der Größe

$$C = \frac{1}{N^2}(f_{11}^2 + f_{22}^2 + \dots + f_{kk}^2) \quad (2.47)$$

zu ersetzen. Alle  $f_{ij}$  im Zähler stammen aus der Hauptdiagonalen der Kontingenztabelle. Das Ergebnis nach Einsetzen von Gl. (2.47) in Gl. (2.46) ergibt den modifizierten Koeffizienten

$$\kappa_{\text{ney}} = \frac{B - C}{1 - C}. \quad (2.48)$$

Dabei gehen in die Berechnung von  $\kappa_{\text{ney}}$  nur die  $f_{ij}$  des ursprünglichen Korrekturterms  $E$  ein, die gleichzeitig auch Anteil von  $B$  sind. Daraus lässt sich ableiten, dass für  $B = 0$  keine Zufallskorrektur vorgenommen wird. Zudem nimmt der Koeffizient in keinem Fall negative Werte an. Analog zu  $\kappa$  ist  $\kappa_{\text{ney}}$  nicht definiert, wenn beide Bewerter für alle  $N$  Objekte dieselbe Merkmalskategorie wählen (Kutschmann, 2020).

### 2.2.6 Kappa-Koeffizient nach Fleiss

Der Grad der Urteilskonkordanz bezogen auf die Übereinstimmung kategorialer Beurteilungen von  $N$  Objekten mit  $m > 2$  Bewertern (verwendet unter der Bezeichnung: Interrater-Reliabilität) kann mit dem Kappa-Koeffizienten  $\kappa_{\text{Fleiss}}$  nach (Fleiss, 1971) bestimmt werden. Dabei werden die Urteile der Bewerter einer  $k$ -stufigen Merkmalskategorie zugeordnet.

Die kategorialen Bewertungen über  $N$  Objekte von  $m > 2$  Bewertern in  $k$  Merkmalskategorien werden in eine Häufigkeitstabelle mit  $i = 1, \dots, N$  Zeilen und  $j = 1, \dots, k$  Spalten eingetragen, dargestellt in Tab. 2.4. Dabei ergeben sich die Häufigkeiten  $n_{ij}$ . Wenn  $n_{ij}$  für ein Objekt der Zeile  $i$  und eine Merkmalskategorie der Spalte  $j$  gleich der Anzahl von  $m$  Bewertern ist, liegt ein vollkommen konkordantes Urteil vor. Verteilen sich die  $n_{ij}$  über alle  $k$  Merkmalskategorien, liegt eine geringe Urteilskonkordanz vor (Bortz und Lienert, 2008, S. 314). Für die Berechnung von  $\kappa_{\text{Fleiss}}$  wird der Anteil  $\tilde{p}_j$  aller Bewertungen einer Merkmalskategorie über

$$\tilde{p}_j = \frac{1}{Nm} n_{\bullet j}, \quad \text{mit } n_{\bullet j} = \sum_{i=1}^N n_{ij}, \quad (2.49)$$

bestimmt. Hierzu werden die Randhäufigkeiten  $n_{\bullet j}$  durch Aufsummieren aller  $n_{ij}$  einer Spalte  $j$  gebildet und durch das Produkt aus  $N$  Objekten und  $m$  Beurteilern dividiert. Weiter wird die Größe  $\hat{p}_i$  berechnet mit

$$\hat{p}_i = \frac{1}{m(m-1)} \sum_{j=1}^k n_{ij}(n_{ij} - 1). \quad (2.50)$$

Die rechte Seite von Gl. (2.50) gilt für ein Objekt der Zeile  $i$  und setzt den Anteil der übereinstimmenden Bewerterpaare mit der maximal möglichen Anzahl von übereinstimmenden Bewerterpaaren in Beziehung (Bortz und Lienert, 2008, S. 314 f.).

Tab. 2.4:  $\kappa_{\text{Fleiss}}$ -Häufigkeitstabelle zur Bestimmung der Urteilskonkordanz für  $m > 2$  Bewerter mit  $k \geq 2$  Merkmalskategorien und  $N$  Objekten.

		Bewerter $m > 2$				$\hat{p}_i$
		$k \geq 2$				
N Objekte		1	2	...	$k$	
	1	$n_{11}$	$n_{12}$	...	$f_{1k}$	$\hat{p}_1$
	2	$n_{21}$	$n_{22}$	...	$f_{2k}$	$\hat{p}_2$
	...	...	...	...	...	...
	$N$	$n_{N1}$	$n_{N2}$	...	$n_{Nk}$	$\hat{p}_N$
	Summe	$n_{\bullet 1}$	$n_{\bullet 2}$	...	$n_{\bullet k}$	
	$\tilde{p}_j$	$\tilde{p}_1$	$\tilde{p}_2$	...	$\tilde{p}_k$	

Durch Mitteln von  $\hat{p}_i$  in allen Zeilen  $i$  nach Gl. (2.33) ergibt sich die Wahrscheinlichkeit

$$\bar{P} = \frac{\sum_{i=1}^N \hat{p}_i}{N}. \quad (2.51)$$

Explizit ist  $\bar{P}$  die Wahrscheinlichkeit, mit der die Urteile von zwei beliebigen Bewertern im Mittel für ein zufällig gewähltes Objekt  $N$  übereinstimmen. Weiter wird durch Aufsummieren der Quadrate von  $\tilde{p}_j$  über die  $k$  Merkmalskategorien die Wahrscheinlichkeit

$$\bar{P}_E = \sum_{j=1}^k \tilde{p}_j^2 \quad (2.52)$$

bestimmt. Dabei ist  $\bar{P}_E$  die Wahrscheinlichkeit, mit der die Urteile von zwei beliebigen Bewertern für ein zufällig gewähltes Objekt  $N$  zufällig übereinstimmen (Fleiss, 1971).

Zuletzt erfolgt die Berechnung von  $\kappa_{\text{Fleiss}}$  unter Einsetzen der Wahrscheinlichkeiten aus Gl. (2.51) und Gl. (2.52) in

$$\kappa_{\text{Fleiss}} = \frac{\bar{P} - \bar{P}_E}{1 - \bar{P}_E}. \quad (2.53)$$

Ist  $\kappa_{\text{Fleiss}} = 1$  für  $\bar{P} = 1$  liegt eine vollständige Urteilsakkordanz vor. Keine Beurteilerübereinstimmung ergibt sich aus  $\kappa_{\text{Fleiss}} = 0$  und  $\bar{P} = \bar{P}_E$ . Für  $\bar{P} < \bar{P}_E$  kann  $\kappa_{\text{Fleiss}}$  negativ werden (Bortz und Lienert, 2008, S. 311).

Einen Vorschlag zur Interpretation von  $\kappa_{\text{Fleiss}}$  bringt die Arbeit von (Landis und Koch, 1977). Die Tab. 2.5 zeigt die  $\kappa_{\text{Fleiss}}$  zugeordneten Übereinstimmungsgrade.

Tab. 2.5: Interpretation des  $\kappa_{\text{Fleiss}}$ -Koeffizienten zur Bewertung der Übereinstimmung kategorialer Bewertungen.

$\kappa_{\text{Fleiss}}$ -Wert	Übereinstimmungsgrad
< 0.00	gering
0.00 – 0.20	schwach
0.21 – 0.40	mittelmäßig
0.41 – 0.60	moderat
0.61 – 0.80	substanziell
0.81 – 1.00	nahezu perfekt

## 2.3 Messphantome

### 2.3.1 Allgemeine Spezifikationen

Die CT-Experimente der vorliegenden Dissertation wurden ausschließlich mit kommerziellen Messphantomen durchgeführt. Begründet durch das weitreichende Untersuchungsspektrum sind zwei multifunktionale, anthropomorphe Lungenphantome mit artifiziellen Lungenrund-

herden eingesetzt worden. Insbesondere unterlagen die Materialien des Messequipments einer Kalibrierung auf gewebezugeordneten HU–Werten bei 120 kV. Die Werkstoffbasis der verwendeten Modelle bildeten Polyurethane und Epoxidharze. Weitere Spezifikationen sind den Modelldatenblättern in Anhang A.2 zu entnehmen.

### 2.3.2 Lungenrundherdmodelle

Die Simulation von Lungenrundherden erfolgte durch Kugel- und Spikulaemodelle (QRM GmbH, Möhrendorf, Deutschland; Kyoto Kagaku Co., Ltd., Kyoto, Japan), dargestellt in Abb. 2.6 (a) und (b). Eine Kompatibilität der artifiziellen Lungenrundherden bestand mit beiden anthropomorphen Lungenphantomen.

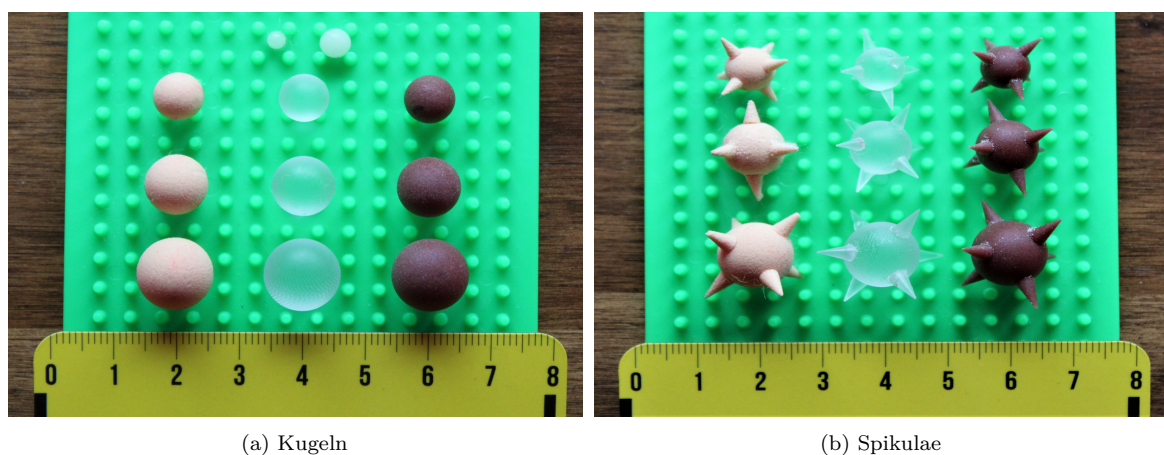


Abb. 2.6: Übersicht der Lungenrundherdmodelle als (a) Kugeln und (b) Spikulae. Die nominalen HU–Werte der Materialien beider Geometrietypen entsprechen der folgenden Farbzuzuordnung: 100/ – 50/ – 690 HU = opak/ braun/ beige.

Die Modelleigenschaften wurden auf Rat eines Radiologen (Kühl, 2018) und basierend auf einer Phantomstudie (Doo et al., 2014) zur Beurteilung der Qualität der Lungenrundherdvolumetrie in der Niedrigdosis–CT mit IR–Rekonstruktion ausgewählt. Die Klassifizierung der Lungenrundherdmodelle erfolgte anhand ihrer dedizierten, nominalen HU–Werte in drei Herdtypen: 100 HU = Referenzherd, –50 HU = solider Herd und –690 HU = subsolider Herd. Die kalibrierten Modelle unterlagen einer Herstellertoleranz von  $\pm 10$  HU auf die bei der Abnahmeprüfung gemessenen effektiven HU ( $HU_{\text{eff}}$ )–Werte. Das explizite Auswahlkriterium für die Referenzherdmodelle bestand in der Eigenschaft einer hohen Kantenschärfe in CT–Bildern und der damit verbundenen Eignung für objektive Bildqualitätsprüfungen. Die artifiziellen soliden Herde dienten der Simulation von Metastasen und sekundären Lungentumoren. Die subsoliden Herdmodelle simulierten frühe Formen von Lungentumoren (z. B. subsolide Herde eines bronchioloalveolären Karzinoms). Die Durchmesser der Kugeln sowie die Kerndurchmesser der Spikulae waren gleich. Überdies wurden zwei zusätzliche Kugeln mit kleinen Durchmessern verwendet. Eine Übersicht der Modelleigenschaften wird in Tab. 2.6, S. 41, vorgestellt. Die Listung beinhaltet ausschließlich absolute Durchmessermaße.

Tab. 2.6: Modelleigenschaften der artifiziellen Lungenrundherde für beide geometrischen Ausführungen.

Geometrie	HU bei 120 kV	HU <sub>eff</sub> bei 120 kV	Durchmesser mm	Herdtyp
Kugel	100	90	12, 10, 8, 5, 3	referenz
	-50	-49	12, 10, 8	solide
	-690	-682	12, 10, 8	subsolide
Spikulae	100	90	24, 20, 16	referenz
	-50	-53	24, 20, 16	solide
	-690	-682	24, 20, 16	subsolide

### 2.3.3 Lungman-N1-Phantom

Das Lungman-N1-Phantom PH-1 R16511 (Kyoto Kagaku Co., Ltd., Kyoto, Japan) simulierte den männlichen Thorax eines Erwachsenen, dargestellt in Abb. 2.7. Der Habitus replizierte die Anatomie durch synthetische Komponenten, bestehend aus einer Brustwand mit Knochen, einem Lungengefäßsystem mit Mediastinum (explizit: Luftröhre und Herz) sowie einem Zwerchfell-Abdomen-Block. Die Modelldimensionen beliefen sich auf: Länge,  $L = 430$  mm, Breite,  $B = 400$  mm, Höhe,  $H = 440$  mm und Umfang,  $U = 9400$  mm.



Abb. 2.7: Ventrale Ansicht des anthropomorphen Lungman-N1-Phantoms PH-1 R16511 zur Nachbildung der Thoraxanatomie eines erwachsenen Mannes. Die Brustwandreplikation mit synthetischen Knochen (links) kann der Kombination aus einem artifiziellen Lungengefäßsystem inklusive Mediastinum und einem Zwerchfell-Abdomen-Block (rechts) aufgesetzt werden.

In Vorbereitung auf die Scanprozedur wurden in die modellierte Lungenvaskulatur Kugel- bzw. Spikulaemodelle eingebracht, dargestellt in Abb. 2.8 (a) und (c), S. 42. Die gewählten Targetpositionen richteten sich nach Empfehlungen eines Radiologen hinsichtlich der zu erwartenden Positionen von Lungenrundherden bei einem Lungenkrebs-Screening (Kühl, 2018). Um die Reproduzierbarkeit der Positionen zu gewährleisten, wurden Markierungen in der synthetischen Lungenstruktur mit wasserfestem Marker gesetzt. Als Positionierungswerkzeug für den Materialeinsatz bzw. -austausch kam eine Pinzette zur Anwendung. Die Fixierung der

Herdmodelle erfolgte mit einem Styropor–Gelkleber. Als Lösungsmittel fungierte Waschbenzin. Für das erste Scan–Setup wurden ausschließlich Kugeln auf die gekennzeichneten Positionen geklebt. Das zweite Scan–Setup verwendete die Spikulae. Die Lagebeziehungen der Targets können jeweils der (b) ventralen und (d) dorsalen Konstruktionsansicht entnommen werden.

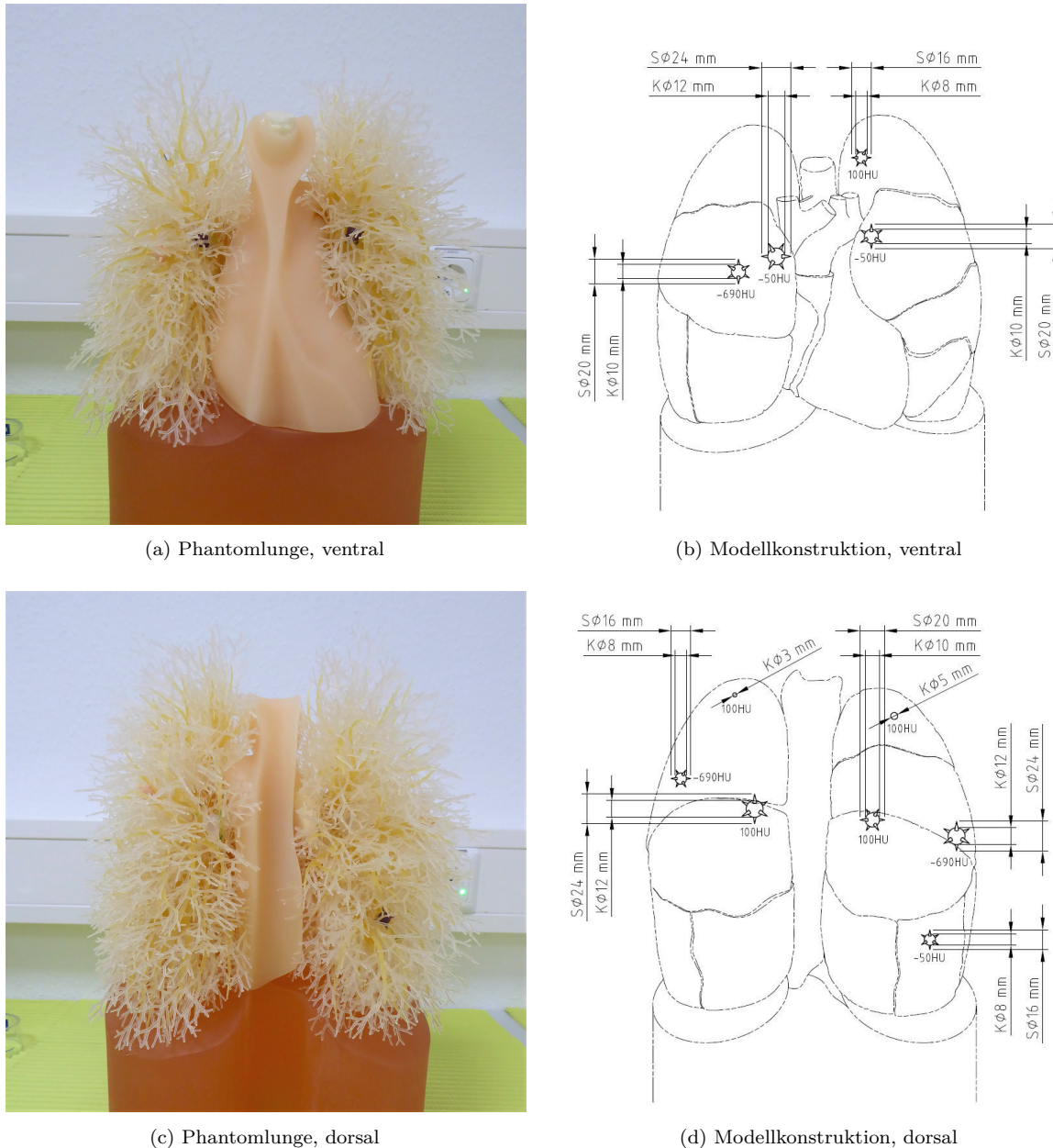


Abb. 2.8: Ansichten und maßstabsgerechte Konstruktion der manipulierten artifiziiellen Lungenvaskulatur des Lungman–N1–Phantoms. Die Targetpositionen sind innerhalb des Phantoms ausgehend von (a) der ventralen Modellansicht in (b) der zugehörigen Konstruktionszeichnung dargestellt. Der Konstruktionsansicht können die HU–Werte und die Durchmessermaße der Kugeln, K bzw. Spikulae, S entnommen werden. Gleiches gilt für (c) die dorsale Modellansicht mit (d) der zugehörigen Konstruktionszeichnung. Die schräggestellten Durchmessermaße sind ausschließlich Kugelpositionen zugeordnet.

### 2.3.4 QRM–Thorax–Phantom mit Extensionsringen

Das Lungenrundherd–QRM–Thorax–Phantom (QRM GmbH, Möhrendorf, Deutschland) simulierte die menschliche Lungenanatomie durch eine axial–extrudierte Modellscheibe (QRM–



Body), dargestellt in Abb. 2.9. Im Einzelnen modellierte der QRM-Body zwei mit Granulat befüllbare Lungenlappen, ein Mediastinum, eine Wirbelsäule sowie den Weichteilpart einer Brustwand. Verschraubbare Deck- und Basisplatten verschlossen den Modellkörper bei Materialeinsatz. Adaptiv verfügte das Phantom über zwei Extensionsringe, wobei entweder ein Gewebeäquivalent (Ring M) oder ein Fettäquivalent (Ring L) dem QRM-Body aufgesetzt werden konnte. Die Dimensionen des QRM-Body waren:  $L = 112 \text{ mm}$ ,  $B = 300 \text{ mm}$ ,  $H = 200 \text{ mm}$  und  $U = 825 \text{ mm}$ . Die Abmessungen für den Ring M beliefen sich auf  $L = 100 \text{ mm}$ ,  $B = 350 \text{ mm}$ ,  $H = 250 \text{ mm}$  und  $U = 980 \text{ mm}$  sowie für den Ring L auf  $L = 100 \text{ mm}$ ,  $B = 400 \text{ mm}$ ,  $H = 300 \text{ mm}$  und  $U = 1130 \text{ mm}$ .

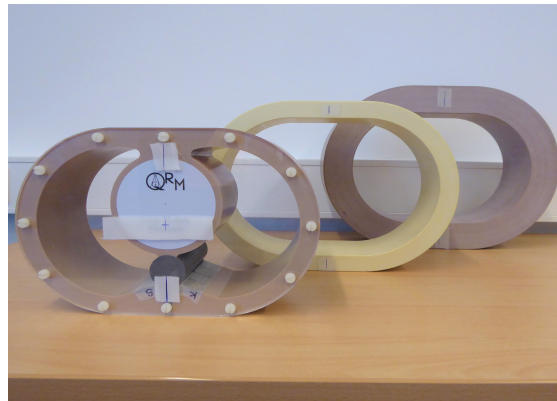
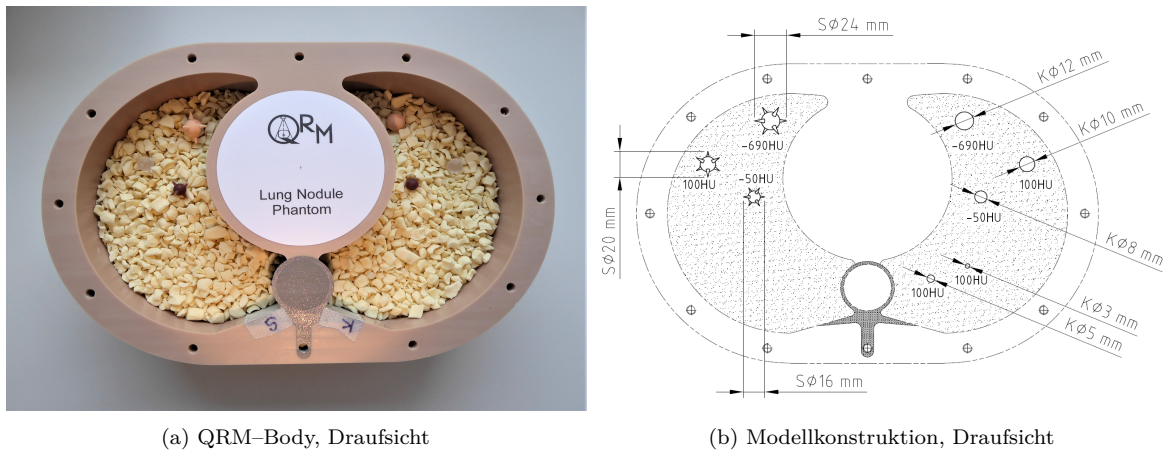


Abb. 2.9: Draufsicht des anthropomorphen Lungenrundherd-QRM-Thorax-Phantoms zur Nachbildung der menschlichen Lungenanatomie. Die axial-extrudierte, befüllbare Modellscheibe bildet den QRM-Body (vorne, links). Überdies können Extensionsringe, Ring M als Gewebeäquivalent (mittig) und Ring L als Fettäquivalent (hinten, rechts), adaptiert werden.

In Vorbereitung auf das CT-Experiment erfolgte eine Schichtung des Granulats in den modellierten Lungenlappen des QRM-Body über das innere Längenmaß von 100 mm. Bei den ausgehend von der Basisplatte vermessenen Granulat-Füllständen von 25 mm, 50 mm und 75 mm wurden jeweils in dem linken Modelllappen Kugeln und in dem rechten Modelllappen Spikulae platziert. Die Modelldraufsicht in Abb. 2.10 (a), S. 44, zeigt den Füllstand von 75 mm bei eingebrachten Targets. Die Wahl der artifiziellen Lungenrundherde bezog sich auf den absoluten Durchmesser und die nominalen HU-Werte. Hierbei galt die Bedingung, dass jeder mögliche Wert pro Eigenschaft in einem Modelllappen unabhängig voneinander mindestens einmal vorkam. Die Targetpositionierung erfolgte beliebig. Die Fixierung der Herdmodelle konnte durch Materialpressung bei verschraubter Deck- und Basisplatte gewährleistet werden. Für den Füllstand von 75 mm können die expliziten Targetdurchmesser sowie -positionen und nominalen HU-Werte der Konstruktionszeichnung aus (b) entnommen werden. Da die CT-Bilder der Granulat-Füllstände von 25 mm und 50 mm in den weiteren Bildqualitätsanalysen keine Anwendung fanden, sind die zugehörigen Konstruktionszeichnungen im Anhang A.3 einsehbar.



(a) QRM-Body, Draufsicht

(b) Modellkonstruktion, Draufsicht

Abb. 2.10: Draufsicht und maßstabgerechte Konstruktion des QRM-Body mit Granulatfüllung. Dargestellt sind die Targetpositionen bei einem Füllstand von 75 mm Granulat (a) innerhalb des Phantoms. Die Konstruktionszeichnung (b) beinhaltet die absoluten Durchmessermaße der Kugeln, K bzw. Spikulae, S sowie deren HU-Werte. Die Kugeln befinden sich im rechten und die Spikulae im linken Modelllungenlappen.

## 2.4 Akquisitions- und Rekonstruktionsmethodik

### 2.4.1 CT-Dosisprotokolle und Scanparameter

Die Rohdatenakquise erfolgte mit dem Dual-Source-CT Somatom Force (Siemens Healthineers, Forchheim, Deutschland) im Single-Source-Betrieb in der Strahlenklinik des Universitätsklinikums Essen (Essen, Deutschland). Die technischen Spezifikationen des Scanners befinden sich im Anhang A.4. Insgesamt wurden vier Konstanzprüfungen und zwei dosissparende Spiral-CT-Experimente durchgeführt. Pro Experiment bzw. Prüfung sind drei CT-Dosisprotokolle aus der folgenden Auflistung gewählt worden:

- Standard-Hochkontrast (SH; klinisch verwendet):  
CARE kV, Referenzspannung von 120 kV,  
CARE Dose4D, Referenz-Röhrenstrom-Zeit-Produkt von 72 mAs.
- Standard-Hochkontrast\* (SH\*; klinisch verwendet):  
Röhrenspannung von 120 kV, Röhrenstrom-Zeit-Produkt von 51 mAs.
- Low-Dose (LD):  
Röhrenspannung von 120 kV, Röhrenstrom-Zeit-Produkt von 40 mAs.
- Ultra-Low-Dose (ULD):  
Röhrenspannung von 120 kV, Röhrenstrom-Zeit-Produkt von 20 mAs.

Die definierten Scanparameter für beide Spiral-CT-Experimente und die dazu durchgeführten Konstanzprüfungen listet Tab. 2.7, S. 45, auf. Der Parameter Field of View (FOV) wurde den Abmaßen des jeweiligen Phantoms angeglichen.



Tab. 2.7: Scanparameter für die CT-Experimente mit beiden Lungenphantomen.

Parameter	Wert
FOV (Lungman-N1)	300 mm
FOV (QRM-Thorax)	430 mm
Rotationszeit	0,5 s
Verzögerungszeit	2 s
Kollimation	196 Detektorzeilen x 0,6 mm
Schichtdicke	1 mm
Pitch	1,2

Die Rekonstruktion der Rohdaten erfolgte mit drei Kernen: dem Standard-Lungen-BI57-Kern, dem weichen Br32-Kern und dem harten Br69-Kern. Alle rekonstruierten CT-Bilder wurden in dem Lungenparenchymfenster  $(-600, 1200)$  zur weiteren Untersuchung bereitgestellt.

## 2.4.2 CT-Experimente

### Konstanzprüfung an dem CT-Scanner

Vor und nach einem CT-Experiment erfolgte eine Konstanzprüfung zum Nachweis der stabilen Akquisitions- und Rekonstruktionsprozedur des kalibrierten CT-Scanners. Hierfür wurde das Waterphantom\_comp\_MxT Modeltype 10742380 (Siemens Healthineers, Forchheim, Deutschland) in eine zugehörige Haltevorrichtung am Tisch des CT-Systems eingespannt, dargestellt in Abb. 2.11 (a).



(a) Waterphantom\_comp\_MxT im CT.

(b) Waterphantom\_comp\_MxT, axial, Prüfbild 212

Abb. 2.11: Das Herstellerphantom Waterphantom\_comp\_MxT Modeltype 10742380 (a) innerhalb der CT-Gantry ist in einer am Tische fixierten Halterung eingespannt. Exemplarisch ist das axiale (b) Prüfbild 212 aufgeführt, mit: CT-Dosisprotokoll/ Kern/ Rekonstruktion/ Schichtdicke = LD/ BI57/ S<sub>3</sub>/ 1 mm.

Die Konstanzprüfung nutzte die drei CT-Dosisprotokolle und die Scanparameter des Lungenphantomscans am jeweiligen Messtag. Pro CT-Dosisprotokoll wurden drei identische Spiral-CT-Scans akquiriert. Für den Nachweis einer stabilen Akquisition erfolgte die Rekonstruktion

der neun Rohdatensätze mit den Algorithmen FBP und ADMIRE in der IR-Stärke  $S_3$ , dem Bl57-Kern und der Schrittweite bzw. Schichtdicke von 1 mm als axiale Schichten. Gleichermaßen wurde für den Nachweis einer stabilen Rekonstruktion jeweils der erste akquirierte Rohdatensatz pro CT-Dosisprotokoll dreimal rekonstruiert. Dabei entstanden zwei Bildproben für die Stabilitätsnachweise, Probe A für die Akquisition und Probe R für die Rekonstruktion, mit je 18 Prüfbildern. Sechs CT-Bilder kamen mit beiden Proben vor. Insgesamt resultierten 30 Prüfserien mit je 248 Einzelbildern. Für Prüfungszwecke wurde das beliebig gewählte und im Phantom zentral liegende Prüfbild 212 aus Abb. 2.11 (b), S. 45 verwendet.

### Lungenphantom-CT-Experimente

Um die Randbedingungen eines klinischen Lungen-CT-Scans zu simulieren, erfolgte die Lagerung beider Phantome auf dem CT-Tisch nach den Parametern: Rückenlage (engl.: supine) und kopfwärts zur Gantry (engl.: head first), dargestellt in Abb. 2.12(a) und (c).

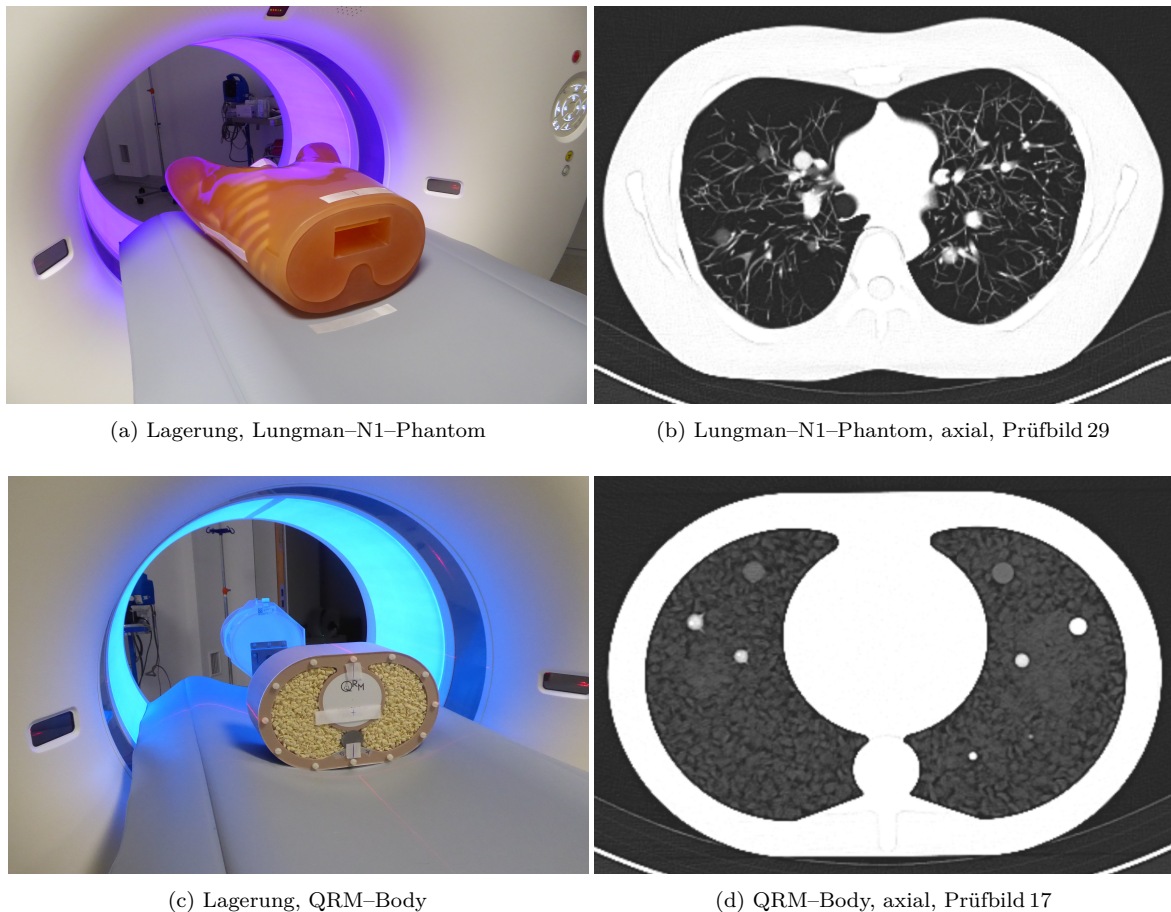


Abb. 2.12: Lagerungsübersicht des (a) Lungman-N1-Phantoms und des (c) QRM-Body. Die Positionierungen entsprechen den Lagerungsvorgaben eines klinischen Lungen-CT-Scans: Rückenlage und kopfwärts zur Gantry. Beide Phantome weisen Markierungen zur reproduzierbaren Positionierung auf. Es folgen die Beispieldarstellungen (b) des axialen Prüfbilds 29 mit dem Lungman-N1-Phantom bzw. (d) des axialen Prüfbilds 17 mit dem QRM-Body für: CT-Dosisprotokoll/ Kern/ Rekonstruktion/ Schichtdicke = LD/ Bl57/  $S_3$ / 5 mm.

**Lungman–N1.** Das CT–Experiment gliederte sich in ein Kugel– und ein Spikulaesetup. Die Scan– und die Rekonstruktionsprozedur waren für beide Setups gleich. Es erfolgte für die Protokolle SH, LD und ULD jeweils ein Spiral–CT. Die Rekonstruktion verwendete die Kerne Bl57, Br32 und Br69. Pro Kern wurden mit den Algorithmen FBP und ADMIRE in den IR–Stärken  $S_1$ ,  $S_3$ ,  $S_5$  sowie der Schrittweite bzw. Schichtdicke von 5 mm axiale Schichten rekonstruiert. Es entstanden 72 Bildserien mit je 62 Einzelbildern. Zur weiteren Analyse wurden alle Bildserien und das jeweilige, separierte Prüfbild 29, siehe Abb. 2.12 (b), S. 46, verwendet. Das Bildauswahlkriterium war die zentrale Lage im Bereich der synthetischen Bifurkation der Luftröhre. Überdies wurden alle Serien und Prüfbilder anonymisiert, randomisiert und die Rekonstruktionsparameter entfernt.

**QRM–Thorax.** Für jeden Modellaufbau (QRM–Body, QRM–Body mit Ring M, QRM–Body mit Ring L) beinhaltete das CT–Experiment den gleichen Scanvorgang. Es wurde jeweils ein Spiral–CT für die Protokolle SH\*, LD und ULD akquiriert. Die Rohdaten wurden mit dem Bl57–Kern, dem Algorithmus ADMIRE in den IR–Stärken  $S_1$ ,  $S_3$  bis  $S_5$  sowie der Schrittweite bzw. Schichtdicke von 5 mm als axiale Schichten rekonstruiert. In gleicher Weise erfolgte Rekonstruktion mit den Rohdaten des QRM–Body bei adaptiertem Ring L für die Kerne Br32 und Br69. Insgesamt entstanden 60 Bildserien mit je 25 Einzelbildern. Das Prüfbild 17, siehe Abb. 2.12 (d), S. 46, mit der höchsten Targetanzahl pro Schicht, wurde separiert. Ferner entstanden zwei Bildproben mit je 36 Prüfbildern (pro IR–Stärke neun CT–Bilder). Die Kriterien für die erste Probe (Probe 1) waren: QRM–Body mit Ring L, variable Kerne (Bl57, Br32, Br69). Für die zweite Probe (Probe 2) galt: Bl57–Kern, variabler Modellaufbau (QRM–Body ohne und mit je einem adaptierten Extensionsring, Ring M bzw. Ring L). Alle Prüfbilder wurden anonymisiert und die Rekonstruktionsparameter entfernt. Für Probe 2 ist der variable Modellaufbau über einen zentralen Bildausschnitt unkenntlich gemacht worden.

## 2.5 Methoden der Bildqualitätsanalyse

### 2.5.1 Soft– und Hardware

Die im Rahmen dieser Dissertation durchgeführte subjektive Bildqualitätsanalyse mit den Bildserien des Lungman–N1–CT–Experiments verwendete klinisch eingesetzte Befundungsmonitore, genormt nach DIN 6868–157:2014–11 (Deutsches Institut für Normung, 2014). Für die Beurteilungen der Bildserien wurden die auf den Monitoren vorhandenen Digital Imaging and Communications in Medicine–Umgebungen (DICOM; deutsch: Digitale Bildgebung und Kommunikation in der Medizin) verwendet. Weitere subjektive Analysen anhand der Bildproben des QRM–Thorax–CT–Experiments erfolgten mit der in Tab. 2.8, S. 48, gelisteten Hardware. Dabei stellte der Computer mit dem kommerziellen Programm MATLAB (The MathWorks, 2018), Version R2018a, grafische Benutzerschnittstellen (GUI; engl.: Graphical User Interface) und Routinen für die Analyse zur Verfügung.

Tab. 2.8: Auflistung eingesetzter Hardware mit Spezifikationen und Produktdetails.

Hardware	Spezifikationen	Produktdetails
PC	Motherboard	MB ASUS
	Chipsatz	Intel(R) Core(TM) i5-7200U CPU @ 2.50 GHz 2.71 GHz
	Arbeitsspeicher	8,00 GB
	Systemtyp	64-Bit Betriebssystem, x64-basierte Prozessor
	Grafikkarte	Intel(R) HD Graphics 620
	Peripherie	Tastatur, Maus
Monitor	Art	HP 27 " Curved LCD Monitor
	max. Auflösung	1920 (W) x 1080 (H)
	Produkt-Nr.	HSTND-9411-A

Die mit den Prüfbildern des Lungman-N1-CT-Experiments durchgeführten objektiven Bildqualitätsprüfungen unterlagen keiner Hardwarevorgabe. Ein einheitlicher Messrahmen wurde mit der frei verfügbaren 64 Bit-DICOM-Umgebung RadiAnt (Medixant, 2016), Version 4.6.9, gewährleistet. Zur Konstruktion bzw. Durchführung statistischer Analysen kam das Tabellenkalkulationsprogramm Excel (Microsoft Corporation, 2016), Version 2016, zur Anwendung. Im Zuge der Messqualitätsprüfungen wurde für Segmentierungs- und Bildmanipulationsaufgaben die Open-Source-Bildverarbeitungssoftware Fiji ImageJ (National Institutes of Health, 2019), Version 1.52p, eingesetzt.

## 2.5.2 Messqualitätsprüfungen

### Konstanzprüfung und statistische Untersuchung

Der Nachweis einer stabilen Akquisition und Rekonstruktion des kalibrierten Scanners am jeweiligen Tag eines CT-Experiments erfolgte anhand einer statistischen Untersuchung der bei den Konstanzprüfungen gemessenen objektiven Bildqualitätsparameter. Hierfür wurden im Rahmen einer Konstanzprüfung mit den Prüfbildern der Probe A bzw. Probe R, S. 46, SNR- und CNR-Messungen in definierten ROI-Regionen auf einer elliptischen Grundfläche von  $1 \text{ cm}^2 \pm 0,05 \text{ cm}^2$  durchgeführt. Die ROI-Positionen im Zentrum des Wasserphantoms ( $\text{ROI}_W$ ) und in Luft ( $\text{ROI}_L$ ) sind einsehbar in Anhang A.5. Die pro ROI detektierten Parameter beliefen sich auf den mittleren HU-Wert mit zugehöriger Standardabweichung sowie den Pixelwert. Die Beziehungen für die Berechnungen der Bildqualitätsmaße sowie der statistischen Maßzahlen sind einsehbar in Kap. 2.1.3 und Kap. 2.2.1. Mit jedem Prüfbild wurden pro ROI das SNR und das CNR (Referenz:  $\text{ROI}_L$ ) bestimmt. Über die FBP- bzw. modellbasiert IR-rekonstruierten (IR-Stärke  $S_3$ ) Prüfbilder pro CT-Dosisprotokoll sind die mittleren SNR- bzw. CNR-Werte pro ROI gebildet worden. Diese Mittelwerte wurden für die jeweils gleich rekonstruierten Prüfbilder beider Konstanzprüfungen an einem Messtag über einen einseitigen  $t$ -Test für gepaarte Stichproben verglichen (Gesamtanzahl  $t$ -Test: 36).

## Qualitätsprüfungen mit HU-Profilen

Via Segmentierung von HU-Profilen erfolgte die Prüfung auf eine korrekte Wiedergabe der nominalen HU der Lungenrundherdmodelle in den CT-Bildern des Lungman-N1-Kugelsetups bei variierender Rekonstruktion. Zur Identifikation von HU-Abweichungen einzelner Pixel innerhalb der abgebildeten größten Kugeln wurde eine manuelle Segmentierung auf einer Linie entlang des 12 mm großen Durchmessers vorgenommen. Die Messungen erfolgten für jeden HU-Modelltyp anhand von Prüfbildern des SH-Protokolls mit dem B157-Kern und den Algorithmen FBP bzw. ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$ . Die gewählten Prüfbilder zeigten den maximal dargestellten Durchmesser für die zu untersuchende Kugel: Bild 28 für 100 HU, Bild 27 für -50 HU und Bild 29 für -690 HU. Die segmentierten HU-Profilwerte wurden gemittelt und mit dem zugehörigen nominalen HU-Wert anhand eines einseitigen  $t$ -Tests für gepaarte Stichproben auf signifikante Abweichungen geprüft. Zur Untersuchung des Einflusses der Rekonstruktion auf die mittleren HU-Profilwerte pro Kugel wurden diese über die Rekonstruktion anhand des  $t$ -Tests gegeneinander getestet (Gesamtanzahl  $t$ -Test: 30). Um den Einfluss der Rekonstruktion auf die Streuung der HU-Werte um den Profilmittelwert  $\overline{HU}$  zu bewerten, ist die Standardabweichung der Stichprobe  $s$  bestimmt und visuell in Relation zu den HU-Profilen evaluiert worden.

### 2.5.3 Bildqualitätsanalyse mit Lungman-N1

Die subjektive Bildqualitätsanalyse wurde von drei Radiologen mit unterschiedlicher Berufserfahrung durchgeführt. Es variierte die Anzahl der Berufsjahre einschließlich der Befundungserfahrung von Lungen-CT-Aufnahmen. Die 72 Prüfserien des Lungman-N1-CT-Experiments, S. 47, sind hinsichtlich der Erkennbarkeit der artifiziellen Lungenrundherde anhand einer sechsstufigen Likert-Skala bewertet worden. Dabei repräsentierten die sechs Stufen Prozentbereiche mit zugeordneter Klassifizierung: 100 % = exzellent, 80–99 % = gut, 65–79 % = moderat, 50–64 % = ausreichend, 20–49 % = schlecht und 0–19 % = inakzeptabel. Die Auswertung umfasste einen visuellen Vergleich der Beurteilungen. Überdies identifizierten die Mediziner die Anzahl und Geometrie der Targets in beiden Lungenhälften des Phantoms.

Acht Medizintechnikingenieure führten in den 72 Prüfbildern, S. 47, in definierten ROI-Regionen SNR- bzw. CNR-Messungen durch. Die Grundfläche mit Toleranz sowie die gemessenen Parameter pro ROI und die formellen Beziehungen für weitere Berechnungen sind identisch mit denen aus Kap. 2.5.2. Es wurde jeweils eine ROI in Luft (Landmarke P), in die rechte Phantomlung (Landmarke RL) und in die linke Phantomlung (Landmarke LL) gesetzt, dargestellt in Abb. 2.13, S. 50.

Das SNR und CNR wurde aus den aufgenommenen Parametern der Landmarken RL und LL bestimmt. Dabei wurde die Landmarke P für das CNR als Referenzregion gewählt. Die Berechnung der mittleren SNR- bzw. CNR-Werte mit zugehöriger Standardabweichung des Mittelwerts erfolgte über alle Prüfprotokolle der Ingenieure. Die Ergebnisse sind nach Geometriesetup und Phantomlungenhälfte sortiert worden. Die feinere Gruppierung erfolgte

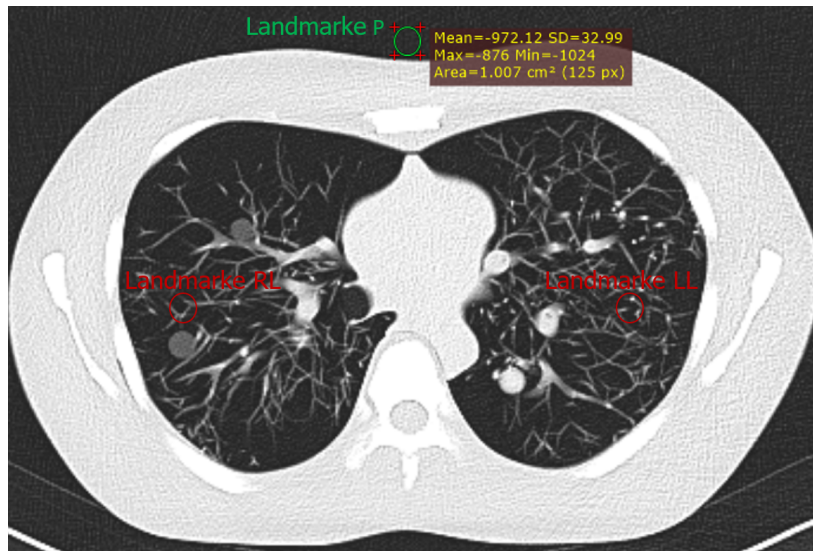


Abb. 2.13: Axiales Lungman-N1-Prüfbild 29 zur Messung objektiver Bildqualitätsmaße mit: CT-Dosisprotokoll/ Kern/ Rekonstruktion/ Schichtdicke = SH/ Bl57/ FBP/ 5 mm. Dabei wurden in der Landmarke RL bzw. LL der rechten und linken Phantomlung (rot) und der Landmarke P in Luft (grün) der mittlere HU-Wert (Mean, gelb), die zugehörige Standardabweichung (SD, gelb) sowie der Pixelwert (Area (px), gelb) gemessen.

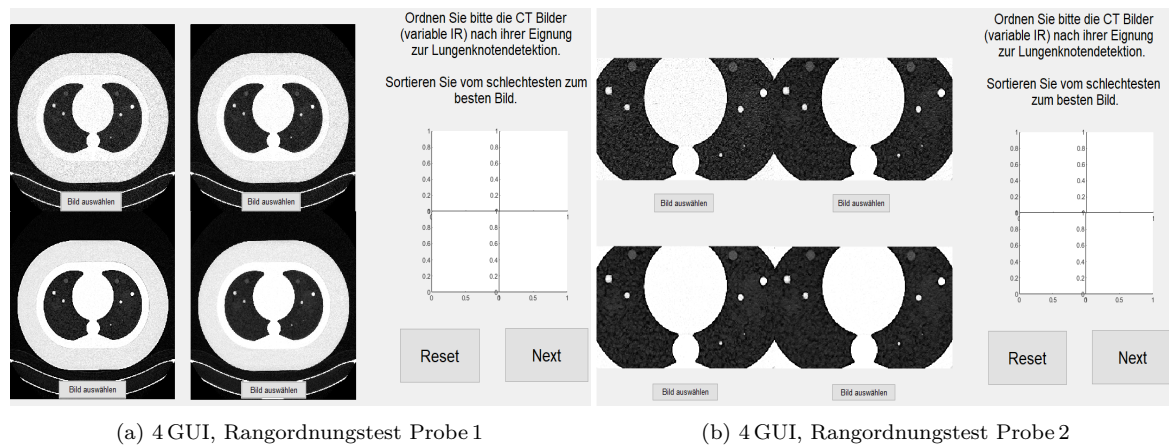
nach CT-Dosisprotokoll und den Rekonstruktionsalgorithmen FBP bzw. ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$ . Die Ergebnisse wurden visuell analysiert und die gruppierten Mittelwerte pro Kern mit einem einseitigen  $t$ -Test für gepaarte Stichproben auf signifikante Unterschiede untersucht (Gesamtanzahl  $t$ -Test: 432).

### 2.5.4 Bildqualitätsanalyse mit QRM-Thorax

Der Ansatz für die subjektiven Bildqualitätsanalysen mit Probe 1 und 2 des QRM-Thorax-CT-Experiments, S. 47, war eine der Sensorik zugeordnete Rangordnungstestmethode mit vorgeschlagenem Friedman-Test. Sechs Radiologen mit variierender radiologischer Erfahrung fungierten als Beobachter. Die 36 Prüfbilder je Probe wurden nach den modellbasierten IR-Stärken  $S_1$ ,  $S_3$  bis  $S_5$  in vier Unterproben mit neun Prüfbildern unterteilt. Diese Unterproben wurden für den Rangordnungstest mit Probe 1 bzw. 2 einer in der MATLAB-Umgebung implementierten Rangordnungstestroutine separat übergeben. Die Implementation erfolgte im Rahmen einer dissertationsbegleitenden Masterthesis (Nguyen, 2019). Für die Rangordnungstests stellte die Routine dem Beobachter eine Vierfeld-Benutzerschnittstelle (4 GUI) zur Verfügung. Mit Probe 1 und 2 zeigte der 4 GUI bei laufender Routine ein Prüfbild pro Unterprobe. Somit stellte die Routine neun Vierfeld-Bildsets je Rangordnungstest. Die Prüfbildpositionierung im 4 GUI sowie die Prüfbildauswahl aus der Unterprobe erfolgte randomisiert. Die Abb. 2.14, S. 51, stellt den 4 GUI für den Rangordnungstest mit (a) Probe 1 und (b) Probe 2 dar.

Die Radiologen wurden gebeten, die Prüfbilder des Vierfeld-Bildsets nach deren Eignung zur Lungenrundherddetektion (bezeichnet auf dem 4 GUI als Lungenknotendetektion) vom schlecht- zum bestgeeignetsten Bild hin zu sortieren. Anhand der Klickreihenfolge legten die Mediziner eine Rangfolge fest. Die Prozedur erfolgte für alle neun Vierfeld-Bildsets pro





(a) 4 GUI, Rangordnungstest Probe 1

(b) 4 GUI, Rangordnungstest Probe 2

Abb. 2.14: Vierfeld–Benutzerschnittstelle für (a) den Rangordnungstest mit Probe 1 und (b) mit Probe 2. Das Vierfeld auf der Schnittstelle zeigt ein Bildset aus zufällig gewählten und positionierten Prüfbildern der Unterproben. Die Aufgabe für den Beobachter ist rechts oben auf der Benutzerschnittstelle orientiert. Durch das Klicken auf **Bild auswählen** unterhalb der Prüfbilder werden diese selektiert und die Routine vergibt einen Rang. Wenn alle Ränge gesetzt wurden, kann mit **Next** das nächste Vierfeld–Bildset zur Bewertung aufgerufen werden. Bei falsch gesetztem Klick ist das Laden eines neuen Sets über **Reset** möglich.

Rangordnungstest. Jeder Test wurde dreimal durchgeführt. Abschließend sind die Testdaten in Excel exportiert und die vergebenen Ränge je Unterprobe bzw. IR–Stärke summiert worden. Mit den Rangsummen pro Testung wurde ein Friedman–Test zu dem Signifikanzniveau  $\alpha = 5\%$  durchgeführt (Gesamtanzahl Friedman–Test: 36).

Sowohl für den Rangordnungstest mit Probe 1 als auch mit Probe 2 erfolgten Konkordanzanalysen. Die Urteilkonkordanz aller Mediziner bzw. die Interrater–Reliabilität ist mit dem Fleiss‘ $\kappa$ –Test bestimmt worden. Dabei erfolgte der Test separat für jeden Rang (Gesamtanzahl Fleiss‘ $\kappa$ –Test: 24). Für die Urteilkonkordanz eines Mediziners zu unterschiedlichen Zeitpunkten bzw. die Intrarater–Reliabilität sind der Cohen’s  $\kappa$ –Test sowie ein modifizierter  $\kappa_{\text{ney}}$ –Test durchgeführt worden. Jeder Test erfolgte paarweise zwischen den drei einzelnen Rangordnungstestungen pro Probe und separat für jeden Rang (Gesamtanzahl  $\kappa$ –Test: 120; Gesamtanzahl  $\kappa_{\text{ney}}$ –Test: 120).

### 2.5.5 Analyse der Strahlenexposition

Die Bewertung der CT–spezifischen Dosisgrößen erfolgte mit dem scannerintegrierten Dosisreport des Lungman–N1–CT–Experiments. Für das Kugel– bzw. Spikulaesetup wurden pro CT–Dosisprotokoll das DLP sowie der  $\text{CTDI}_{\text{vol}}$  aufgenommen. Es folgte die Abschätzung der effektiven Dosis  $E$  mit dem thoraxspezifischen Konversionsfaktor  $f$  von  $0,014 \text{ mSv mGy}^{-1} \text{ cm}^{-1}$  nach Gl. (2.32). Die  $\text{CTDI}_{\text{vol}}$ –Werte wurden mit dem diagnostischen Referenzwert für Thorax–CT von 10 mGy (Bundesamt für Strahlenschutz, 2016) anhand eines einseitigen  $t$ –Tests für gepaarte Stichproben verglichen (Gesamtanzahl  $t$ –Test: 6). Dazu ist die prozentuale Minderung des  $\text{CTDI}_{\text{vol}}$  gegenüber dem diagnostischen Referenzwert bestimmt worden.

## 3 Ergebnisse

### 3.1 Qualitätsmessungen

#### Statistische Untersuchung mit der Konstanzprüfung

Die Akquisition und Rekonstruktion des kalibrierten CT-Scanners waren an jedem Messtag stabil. Die Ergebnisse des  $t$ -Tests für den statistischen Vergleich der mittleren SNR- und CNR-Werte der gleich rekonstruierten Prüfbilder aus den Konstanzprüfungen vor und nach einem Lungen-CT-Experiment legten in keinem Fall signifikante Unterschiede offen ( $p < 0,05$ ).

#### Evaluierung der HU-Profile

Eine Übersicht der HU-Profile mit eingezeichneten HU-Herstellertoleranzen ( $HU_{\text{eff}} \pm 10 \text{ HU}$ ) für die  $-690 \text{ HU}$ -Kugel (Durchmesser:  $12 \text{ mm}$ ) bringt Abb. 3.1 (a) bis (d), S. 53. Es wurden darin die für weitere Berechnungen herangezogenen HU-Profilwerte hellgrau hinterlegt und dem Kugellinneren zugeordnet. Dabei fanden sich Spitzen an den HU-Profilrändern. Die auf- bzw. absteigenden HU-Profilflanken wurden Messwerten umliegender Lungenstruktur zugeordnet. Mit zunehmendem modellbasierten IR-Stärkeniveau glättete sich der HU-Profilverlauf mit abnehmenden Spitzen. Am unebensten stellte sich das HU-Profil für die FBP-Rekonstruktion dar. Hier waren die wenigsten HU-Profilwerte innerhalb der HU-Herstellertoleranzen. Die meisten tolerierten Werte kamen mit der modellbasierten IR-Stärke  $S_5$  vor.

Die HU-Profile der Kugeln mit  $100 \text{ HU}$  und  $-50 \text{ HU}$  sind dem Anhang A.6 zu entnehmen. Mit beiden Kugeln fand sich sowohl ein geglätteter HU-Profilverlauf mit zunehmender modellbasierter IR-Stärke als auch abflachende Spitzen. Am rechten Rand des  $100 \text{ HU}$ -Profils zeigte sich keine Spitze. Für beide Kugeln befanden sich mit der FBP-Rekonstruktion die wenigsten HU-Profilwerte innerhalb der Toleranzen. Hingegen zeigten sich die meisten tolerierten Werte für die  $100 \text{ HU}$ -Kugel mit der modellbasierten IR-Stärke  $S_5$  und für die  $-50 \text{ HU}$ -Kugel mit jedem modellbasierten IR-Stärkeniveau.

In der Tab. 3.1, S. 53, werden die über die einzelnen HU-Profile, Abb. 3.1, S. 53 und Anhang A.6, berechneten Mittelwerte  $\overline{HU}$  mit der zugehörigen Standardabweichung der Stichprobe  $s$  über die Rekonstruktionsalgorithmen FBP und ADMIRE in den IR-Stärken  $S_1, S_3, S_5$  aufgeführt. Für die  $\overline{HU}$ -Profilwerte der  $-50 \text{ HU}$ -Kugel wurde anhand der Ergebnisse des  $t$ -Tests die korrekte Wiedergabe der nominalen HU über alle Algorithmen bestätigt. Hierbei gab es keine signifikanten Abweichungen der  $\overline{HU}$ -Werte zum Nominalwert ( $p < 0,05$ ). Dies zeigte sich bis auf eine Ausnahme auch für die  $-690 \text{ HU}$ -Kugel. Gegenteiliges fand sich über alle Algorithmen für die  $100 \text{ HU}$ -Kugel. Der  $t$ -Test zum Vergleich der  $\overline{HU}$ -Profilwerte je Kugel ergab über alle Algorithmen keine signifikanten Abweichungen ( $p < 0,05$ ), mit jeweils einer Ausnahme für die



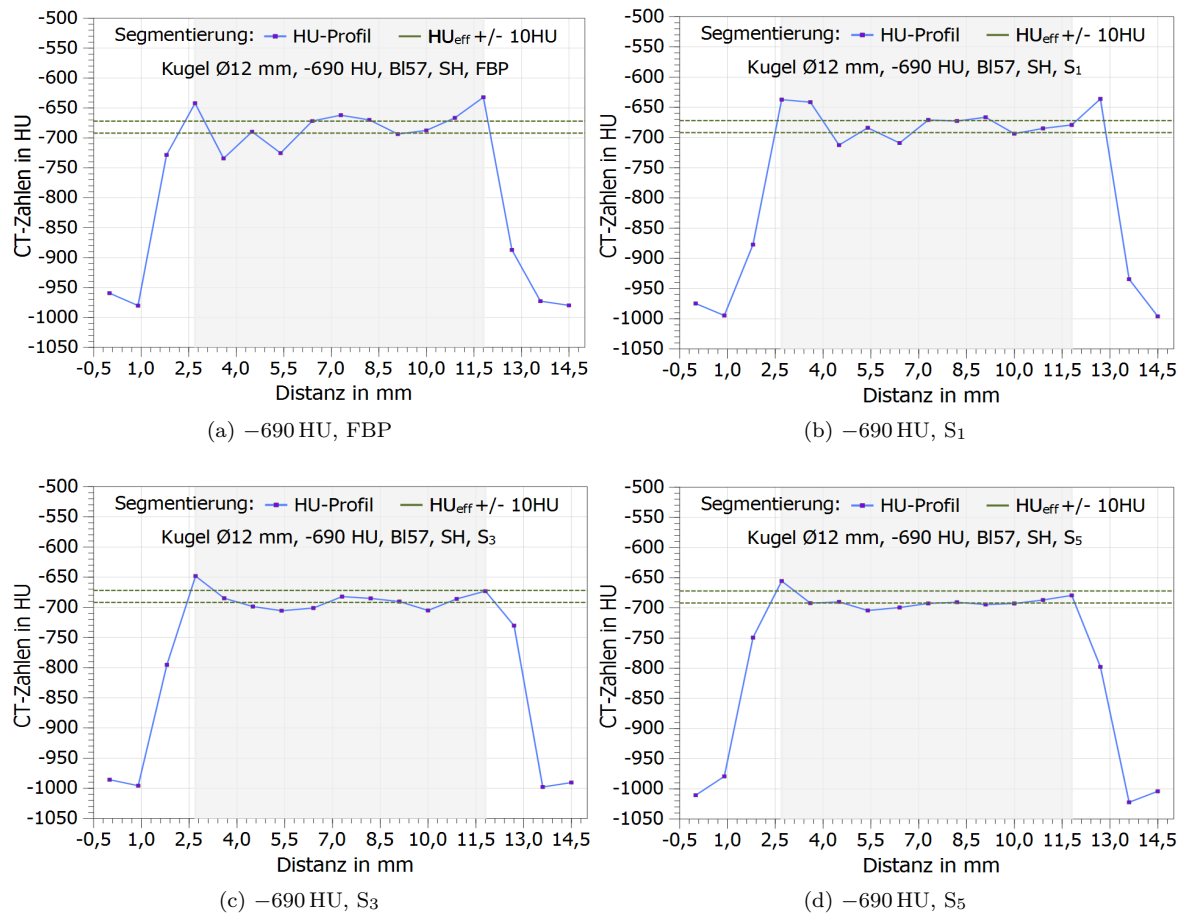


Abb. 3.1: Segmentierte HU-Profile aus den Lungman-N1-Kugelsetup-CT-Bildern für die  $-690$  HU-Kugel mit 12 mm Durchmesser. Die Darstellung erfolgt für das CT-Dosisprotokoll SH, den Standard-Lungen-BI57-Kern und die Rekonstruktionsalgorithmen (a) FBP sowie (b) bis (d) ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$ . Die HU-Profile werden über die CT-Zahlen in HU und über die Segmentierungslinie in mm abgebildet. Die grau hinterlegten HU-Profilwerte liegen innerhalb der Kugel. Für die Bildpixelgröße der verwendeten CT-Bilder von  $0,824 \text{ mm}^2$  lag eine Auflösung von  $1,101 \text{ Pixel/mm}$  vor.

Tab. 3.1: Mittelwert  $\overline{HU}$  und Standardabweichung der Stichprobe  $s$  aus den HU-Profilwerten innerhalb der Kugelmodelle (nominale HU: 100 HU,  $-50$  HU,  $-690$  HU) mit 12 mm Durchmesser. Die Ergebnisse werden über die Rekonstruktionsalgorithmen FBP und ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$  aufgeführt.

Rekonstruktions- algorithmus	HU-Profile der Kugeln mit 12 mm Durchmesser					
	100 HU		$-50$ HU		$-690$ HU	
	$\overline{HU}$	$s$	$\overline{HU}$	$s$	$\overline{HU}$	$s$
FBP	54,79	41,97	$-50,90$	63,44	$-679,54$	31,22
$S_1$	65,11	40,50	$-51,54$	59,32	$-677,37$	24,04
$S_3$	63,82	28,60	$-48,26$	50,13	$-687,33$	16,60
$S_5$	68,57	19,42	$-46,98$	42,67	$-688,91$	12,78

100 HU- und die  $-690$  HU-Kugel. Ein Einfluss der Rekonstruktion auf  $\overline{HU}$  konnte somit nicht festgestellt werden. Die Werte für  $s$  waren für die HU-Profile der FBP-Bilder am höchsten und verringerten sich mit zunehmendem modellbasierten IR-Stärkeniveau. Damit wurde der Einfluss der Rekonstruktion auf  $s$  identifiziert. Die höchsten bzw. geringsten  $s$ -Werte kamen mit dem  $-50$  HU- bzw.  $-690$  HU-Profil vor.

## 3.2 Lungman–N1

### 3.2.1 Likert–Skala

Die individuell wahrgenommene Erkennbarkeit der Lungenrundherdmodelle in den Prüfserien des Lungman–N1–CT–Experiments wird in Abb. 3.2, S. 55, anhand einer sechsstufigen Likert–Skala dargestellt. Dabei werden die subjektiven Ergebnisse über die Rekonstruktionsalgorithmen FBP und ADMIRE in den IR–Stärken  $S_1$ ,  $S_3$ ,  $S_5$  mit allen CT–Dosisprotokollen (SH, LD, ULD) für jeden Rekonstruktionskern (B157, Br32, Br69) abgebildet. Die Mediziner werden anhand ihrer variierenden radiologischen Erfahrung bezeichnet: 3 Jahre (Radiologe<sub>3</sub>), 6 Jahre (Radiologe<sub>6</sub>) und 32 Jahre (Radiologe<sub>32</sub>).

**Kugelsetup mit B157.** Die Beurteilungen reichten von „moderat“ bis „exzellent“ (Stufe 4–6), Abb. 3.2 (a). Dabei vergab Radiologe<sub>3</sub> mit jedem Rekonstruktionsalgorithmus „exzellent“, außer mit  $S_1$  und ULD „gut“. Radiologe<sub>6</sub> urteilte in allen Fällen mit FBP „gut“ und pro IR–Stärke zwischen „gut“ und „moderat“. Überdies bewertete Radiologe<sub>32</sub> mit jedem Rekonstruktionsalgorithmus „gut“ und „moderat“. Radiologe<sub>3</sub> beurteilte die Erkennbarkeit der Lungenrundherdmodelle höher als die anderen Radiologen und vergab bis auf eine Ausnahme seine individuelle Höchstbewertung. Letztere gab es für Radiologe<sub>6</sub> am meisten mit FBP. Radiologe<sub>32</sub> vergab die gleiche Anzahl individueller Höchstbewertungen für FBP und  $S_3$ . Die höchste Summe moderater Urteile zeigte sich mit  $S_5$ .

**Spikulaesetup mit B157.** Die subjektiven Ergebnisse waren zwischen „moderat“ und „exzellent“ (Stufe 4–6), Abb. 3.2 (b). Dabei bewertete Radiologe<sub>3</sub> mit FBP „exzellent“, mit  $S_1$  bzw.  $S_3$  „gut“ und „exzellent“ sowie mit  $S_5$  „moderat“ und „gut“. Radiologe<sub>6</sub> urteilte ausschließlich „gut“, nur mit  $S_5$  und SH „moderat“. Im Weiteren urteilte Radiologe<sub>32</sub> mit FBP „moderat“ bis „exzellent“, mit  $S_1$  bzw.  $S_5$  „moderat“ und „gut“ und mit  $S_3$  „gut“. Für Radiologe<sub>3</sub> ergaben sich die meisten individuellen Höchstbewertungen. Diese gab es pro CT–Dosisprotokoll mit FBP. Radiologe<sub>6</sub> beurteilte in jedem Fall gleich, nur mit  $S_3$  und SH niedriger. Radiologe<sub>32</sub> hatte das höchste und niedrigste Urteil mit FBP und bewertete über alle CT–Dosisprotokolle mit  $S_3$  „gut“. Die Summe moderater Urteile war mit  $S_5$  am höchsten.

**Kugelsetup mit Br32.** Die Urteile erstreckten sich von „ausreichend“ bis „gut“ (Stufe 3–5), Abb. 3.2 (c). Radiologe<sub>3</sub> bewertete mit FBP und  $S_5$  „moderat“ und mit anderen Algorithmen „gut“ und „moderat“. Zudem vergab Radiologe<sub>6</sub> mit FBP „moderat“ und pro IR–Stärkeniveau „ausreichend“ bis „moderat“. Radiologe<sub>32</sub> urteilte mit FBP „moderat“, mit  $S_1$  bzw.  $S_3$  jeweils „ausreichend“ bis „gut“ und mit  $S_5$  „ausreichend“ und „moderat“. Alle Mediziner urteilten mit FBP „moderat“. Radiologe<sub>3</sub> und Radiologe<sub>32</sub> hatten die individuellen Höchstbewertungen mit  $S_1$  und LD. Weitere zeigten sich für Radiologe<sub>3</sub> mit  $S_3$  und LD bzw. für Radiologe<sub>32</sub> mit  $S_3$  und ULD. Die individuellen Höchstbeurteilungen von Radiologe<sub>6</sub> kamen mit jedem Algorithmus vor. Die höchste Summe ausreichender Urteile entstand mit  $S_5$ .

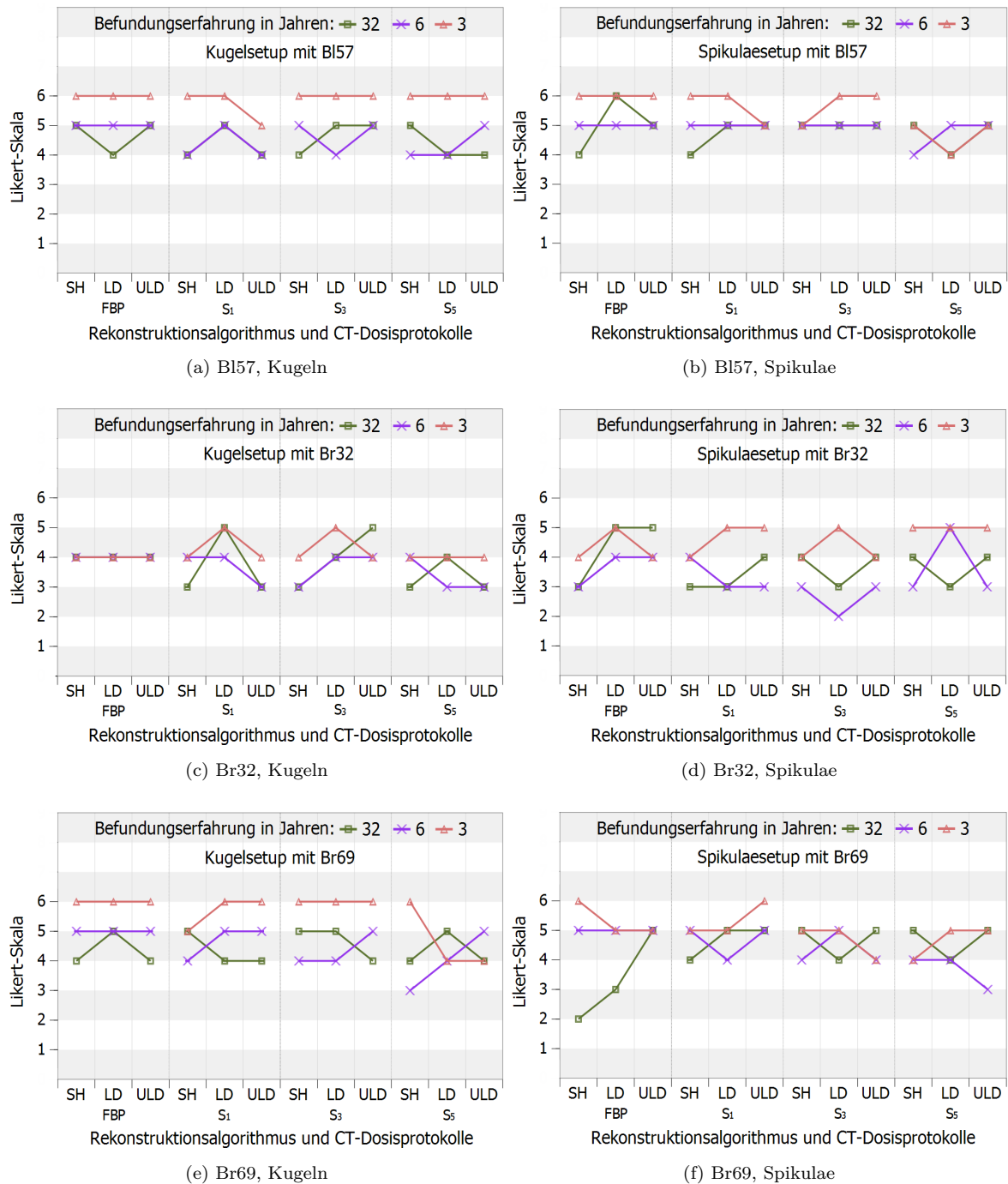


Abb. 3.2: Die individuell wahrgenommene Erkennbarkeit der Lungenrundherdmodelle anhand einer sechsstufigen Likert-Skala mit zugeordneter Klassifizierung: Stufe 6 mit 100 % = exzellent, Stufe 5 mit 80–99 % = gut, Stufe 4 mit 65–79 % = moderat, Stufe 3 mit 50–64 % = ausreichend, Stufe 2 mit 20–49 % = schlecht und Stufe 1 mit 0–19 % = inakzeptabel. Die radiologische Erfahrung folgt der Farbzurordnung: 32 Jahre (grün), 6 Jahre (lila) und 3 Jahre (rot). Die Darstellung der subjektiven Ergebnisse erfolgt im Kugel- (a), (c), (e) und Spikulaesetup (b), (d), (f) über die Algorithmen FBP bzw. ADMIRE in den IR-Stärken S<sub>1</sub>, S<sub>3</sub>, S<sub>5</sub> und alle CT-Dosisprotokolle (SH, LD, ULD) für den Standard-Lungen-BI57-Kern, den weichen Br32-Kern und den harten Br69-Kern.

**Spikulaesetup mit Br32.** Die subjektiven Ergebnisse reichten von „schlecht“ bis „gut“ (Stufe 2–5), Abb. 3.2 (d). Radiologe<sub>3</sub> bewertete mit FBP, S<sub>1</sub> und S<sub>3</sub> „moderat“ und „gut“ sowie mit S<sub>5</sub> „gut“. Die Urteile von Radiologe<sub>6</sub> waren mit FBP bzw. S<sub>1</sub> „ausreichend“ und „moderat“, mit S<sub>3</sub> „schlecht“ und „ausreichend“ sowie mit S<sub>5</sub> „ausreichend“ und „gut“. Radiologe<sub>32</sub> urteilte mit FBP „ausreichend“ und „gut“ sowie mit den IR-Stärkeniveaus jeweils „ausrei-

chend“ und „moderat“. Für Radiologe<sub>3</sub> kamen pro Algorithmus individuelle Höchstbewertungen vor. Diese waren für Radiologe<sub>6</sub> bei S<sub>5</sub> und LD. Radiologe<sub>32</sub> wies die individuellen Höchsturteile mit FBP und LD bzw. ULD auf. Für S<sub>1</sub> ergab sich die höchste Summe ausreichender Beurteilungen. Insbesondere bewertete Radiologe<sub>6</sub> mit S<sub>3</sub> und LD „schlecht“.

**Kugelsetup mit Br69.** Die Beurteilungen erstreckten sich von „ausreichend“ bis „exzellent“ (Stufe 3–6), Abb. 3.2 (e). Radiologe<sub>3</sub> bewertete mit FBP, S<sub>1</sub> und S<sub>3</sub> „exzellent“, außer mit S<sub>1</sub> und SH „gut“. Die Urteile mit S<sub>5</sub> waren „exzellent“ und „moderat“. Radiologe<sub>6</sub> vergab mit FBP „gut“, mit S<sub>1</sub> bzw. S<sub>3</sub> „moderat“ und „gut“ sowie mit S<sub>5</sub> „ausreichend“ bis „gut“. Es fanden sich für Radiologe<sub>32</sub> pro Algorithmus die Urteile „moderat“ und „gut“. Die meisten individuellen Höchstbewertungen und mindestens eine pro Algorithmus ergaben sich für Radiologe<sub>3</sub>. Dieser bewertete insgesamt höher als die anderen Radiologen. Radiologe<sub>3</sub> und Radiologe<sub>6</sub> vergaben für FBP ausschließlich individuelle Höchstbeurteilungen und Radiologe<sub>3</sub> zudem bei S<sub>3</sub>. Die Höchstbeurteilungen von Radiologe<sub>6</sub> und Radiologe<sub>32</sub> kamen mit jedem Algorithmus vor. Die Höchstsumme für moderate Bewertungen gab es mit S<sub>5</sub>. Speziell Radiologe<sub>6</sub> urteilte mit SH „ausreichend“.

**Spikulaesetup mit Br69.** Die Urteile fanden sich zwischen „schlecht“ und „exzellent“ (Stufe 2–6), Abb. 3.2 (f). Die Urteile von Radiologe<sub>3</sub> waren mit FBP bzw. S<sub>1</sub> „gut“ und „exzellent“, mit S<sub>3</sub> bzw. S<sub>5</sub> „moderat“ und „gut“. Radiologe<sub>6</sub> urteilte mit FBP „gut“, mit S<sub>1</sub> bzw. S<sub>3</sub> „moderat“ und „gut“ sowie mit S<sub>5</sub> „ausreichend“ bzw. „moderat“. Die Bewertungen von Radiologe<sub>32</sub> waren mit FBP „schlecht“, „ausreichend“ und „gut“. Weitere Bewertungen waren pro Algorithmus zwischen „moderat“ und „gut“. Die individuellen Höchstbewertungen für Radiologe<sub>3</sub> zeigten sich mit FBP und SH sowie S<sub>1</sub> und ULD. Für Radiologe<sub>6</sub> zeigten sich diese mit FBP ausschließlich, mindestens einmal mit S<sub>1</sub> bzw. S<sub>3</sub> und in keinem Fall mit S<sub>5</sub>. Radiologe<sub>32</sub> hatte mindestens eine individuelle Höchstbeurteilung pro Algorithmus. Die höchste Summe moderater Urteile entstand mit S<sub>3</sub> bzw. S<sub>5</sub>. Die zuletzt genannte IR-Stärke bekam mit ULD „ausreichend“ von Radiologe<sub>6</sub>. Insbesondere urteilte Radiologe<sub>32</sub> „ausreichend“ bzw. „schlecht“ mit FBP.

**Zusammenfassung der Ergebnisse.** Die erfahrenen Radiologen vergaben Höchstbeurteilungen für die Erkennbarkeit der artifiziellen Lungenrundherde in den FBP-Bildern mit dem Standard-Lungen-B157-Kern. Mit Letzterem bewertete der weniger erfahrene Radiologe in den FBP- und modellbasiert IR-rekonstruierten CT-Bildern in gleichem Maße mit höchsten Beurteilungen. Ähnliche subjektive Ergebnisse gab es mit dem Br69-Kern, mit wenigen Ausreißern in den FBP-Bildern und denen mit modellbasiertem IR-Stärkeniveau S<sub>5</sub>. Unabhängig vom CT-Dosisprotokoll wurden in den CT-Bildern mit dem weichen Br32-Kern bei individuell hohen Wahrnehmungsunterschieden deutlich schlechtere Beurteilungen gegenüber anderen Kernen vergeben.

Die Anzahl der individuellen Höchstbewertungen im Kugel- und Spikulaesetup über alle Prüfserien, explizit pro Algorithmus (FBP und ADMIRE in den IR-Stärken S<sub>1</sub>, S<sub>3</sub>, S<sub>5</sub>) und Kern (B157, Br32, Br69), stellt Tab. 3.2, S. 57, dar. Die beste subjektive Erkennbarkeit

Tab. 3.2: Die subjektiv wahrgenommene Erkennbarkeit der Lungenrundherde anhand der Anzahl der individuellen Höchstbewertungen. Die Ergebnisse werden für die Prüferien des Kugel- und Spikulaesetups über die Rekonstruktionsalgorithmen FBP und ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$  sowie den Standard-Lungen-BI57-Kern, den weichen Br32-Kern und den harten Br69-Kern aufgeführt.

Rekonstruktions- algorithmus	Anzahl der individuellen Höchstbewertungen						Summe
	Kugeln			Spikulae			
	BI57	Br32	Br69	BI57	Br32	Br69	
FBP	8	3	7	7	3	5	33
$S_1$	4	4	5	5	2	5	25
$S_3$	7	4	6	5	1	3	26
$S_5$	5	1	3	2	4	2	17
Summe	24	12	21	19	10	15	

für Kugeln und Spikulae ergab sich durch die Höchstsummen der gezählten individuellen Höchstbewertungen für die Prüferien mit dem BI57-Kern und der FBP-Rekonstruktion. Setup- und kernunabhängig fanden sich für die Rekonstruktionsalgorithmen die summierten individuellen Höchstbewertungsanzahlen je Zeile gemäß der Rangfolge: 33 mit FBP, 26 mit  $S_3$ , 25 mit  $S_1$  und 17 mit  $S_5$ . Die Rangfolge für die Kerne im Kugelsetup, unabhängig von der Rekonstruktion, anhand der summierten individuellen Höchstbewertungsanzahlen je Spalte, war: 24 mit BI57, 21 mit Br69 und 12 mit Br32. Dies zeigte sich im Spikulaesetup wie folgt: 19 mit BI57, 15 mit Br69 und 10 mit Br32. Insgesamt waren die Summen je Kern mit Kugeln höher als mit Spikulae. Bis auf einen Ausreißer fanden sich für die Höchstsummen der gezählten individuellen niedrigsten Bewertungen konsistente Ergebnisse.

Ferner identifizierten alle Radiologen in allen Prüferien die korrekte Geometrie der Lungenknotenmodelle. Die Anzahl der Spikulae wurden vollständig erfasst. Die gleiche Anzahl an Targets zählten die Radiologen im Kugelsetup. Kein Befunder detektierte die Kugeln mit 5 mm und 3 mm Durchmesser.

### 3.2.2 Signal-Rausch-Verhältnis

Die Darstellung des Mittelwerts  $\overline{\text{SNR}}$  über alle Prüfprotokolle der Medizintechnikingenieure mit zugehöriger Standardabweichung des Mittelwerts ist in Abb. 3.3, S. 58, einsehbar. Die Präsentation der Ergebnisse erfolgt für alle Kerne (BI57, Br32, Br69) sortiert nach Kugel- und Spikulaesetup sowie für die Landmarken LL und RL in der linken und rechten Phantomlung. Zudem sind die Mittelwerte nach CT-Dosisprotokoll (SH, LD, ULD) gruppiert und über die Algorithmen FBP und ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$  abgebildet.

**Landmarke LL.** Im Kugel- und Spikulaesetup, Abb. 3.3 (a) bis (b), zeigten sich pro CT-Dosisprotokoll für BI57 und Br69 die Minimalwerte des  $\overline{\text{SNR}}$  bei FBP und die Maximalwerte bei  $S_5$ . Für Br32 im Kugelsetup wurde das minimale  $\overline{\text{SNR}}$  in der Reihung SH/LD/ULD mit den IR-Stärken  $S_5/S_3/S_5$  erreicht. Die maximalen Werte ließen sich für  $S_1$ /FBP/FBP

identifizieren. Im Spikulaesetup führten in genannter Reihung  $S_3/S_3/S_5$  zu Minima und  $S_1/S_5$ /FBP zu Maxima des  $\overline{\text{SNR}}$ .

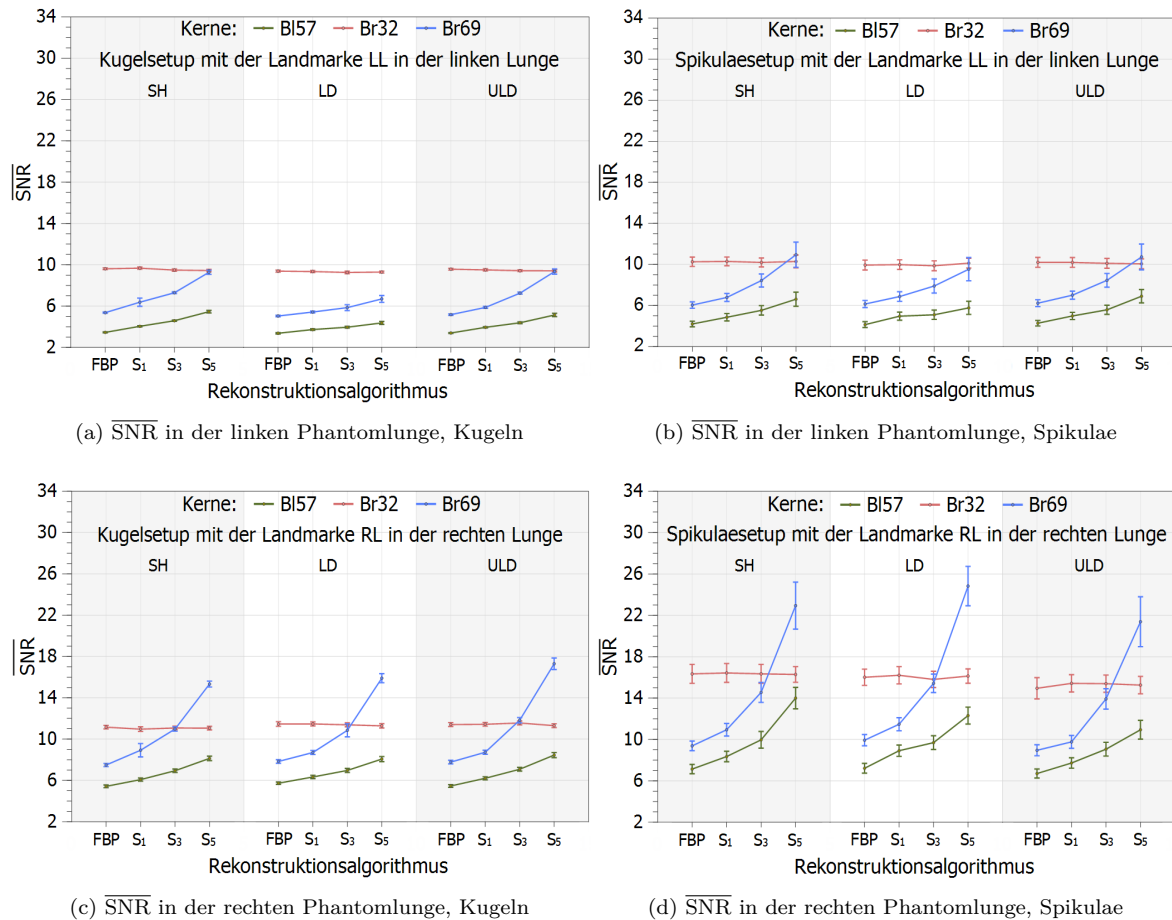


Abb. 3.3: Der Mittelwert  $\overline{\text{SNR}}$  über alle Prüfprotokolle der objektiven Bildqualitätsprüfung mit zugehöriger Standardabweichung des Mittelwerts. Die Ergebnisse werden für die Kerne: BI57 (standard, grün), Br32 (weich, rot) und Br69 (hart, blau) im Kugel- bzw. Spikulaesetup mit der Landmarke LL in der linken Phantomlunge (a) bis (b) und der Landmarke RL in der rechten Phantomlunge (c) bis (d) abgebildet. Die Darstellung erfolgt über die Rekonstruktionsalgorithmen FBP bzw. ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$  und gruppiert nach CT-Dosisprotokoll (SH, LD, ULD). Die Fehlerbalken zeigen die Standardabweichung des Mittelwerts.

Für die Kerne BI57 und Br69 gab es in beiden Setups je CT-Dosisprotokoll einen  $\overline{\text{SNR}}$ -Anstieg mit zunehmender IR-Stärke. Für den Br32-Kern zeigte sich das  $\overline{\text{SNR}}$  unabhängig vom CT-Dosisprotokoll nahezu unverändert. Hierbei fanden sich geringe Mittelwertsenkungen im Kugelsetup mit dem SH-Protokoll und in beiden Setups mit dem ULD-Protokoll. Das  $\overline{\text{SNR}}$  war für den Br32-Kern insgesamt am höchsten, ausser im Spikulaesetup mit dem SH- und dem ULD-Protokoll bei  $S_5$ . Die Standardabweichung des Mittelwerts stellte sich im Spikulaesetup höher dar als im Kugelsetup, ausser für Br69 mit dem SH-Protokoll bei  $S_1$ . Im Spikulaesetup waren die Standardabweichungen mit der FBP-Rekonstruktion am geringsten und erhöhten sich mit zunehmender IR-Stärke. Dies fand sich auch im Kugelsetup für die Kerne BI57 und Br69 mit den Protokollen LD bzw. ULD sowie für den Br32-Kern mit dem SH-Protokoll.

**Landmarke RL.** Für das Kugel- bzw. Spikulaesetup, Abb. 3.3 (c) bis (d), fanden sich für BI57 und Br69 über alle CT-Dosisprotokolle die Minima bzw. Maxima des  $\overline{\text{SNR}}$  bei FBP

bzw.  $S_5$ . Für Br32 gab es im Kugelsetup die Minimalwerte des  $\overline{\text{SNR}}$  gemäß der Reihenfolge SH/LD/ULD für die IR-Stärken  $S_1/S_5/S_5$ . Es fanden sich die Maximalwerte mit FBP/ $S_1/S_3$ . Im Spikulaesetup wurden die minimalen  $\overline{\text{SNR}}$ -Werte gemäß der benannten Reihenfolge mit  $S_5/S_3$ /FBP erreicht und die Maxima pro CT-Dosisprotokoll bei  $S_1$ .

Mit beiden Setups stieg pro CT-Dosisprotokoll für den B157- und den Br69-Kern das  $\overline{\text{SNR}}$  mit zunehmender IR-Stärke an. Hingegen blieben die Mittelwerte für Br32 unabhängig vom CT-Dosisprotokoll nahezu gleich. Dabei gab es mit zunehmender IR-Stärke im Kugelsetup mit dem LD-Protokoll und im Spikulaesetup mit den Protokollen SH und ULD geringe Mittelwertsenkungen. Zudem war für den Br32-Kern das  $\overline{\text{SNR}}$  gegenüber anderen Kernen erhöht, mit Ausnahmen für den Br69-Kern mit deutlich höheren Werten bei  $S_5$ . Es fand sich allgemein eine höhere Standardabweichung des Mittelwerts im Spikulae- als im Kugelsetup, außer für den Br69-Kern mit dem SH-Protokoll bei  $S_1$ . Im Spikulaesetup zeigten sich mit der FBP-Rekonstruktion die niedrigsten Standardabweichungen. Letztere erhöhten sich mit zunehmendem IR-Stärkeniveau. Dies ergab sich im Kugelsetup für den B157-Kern in allen CT-Dosisprotokollen und für den Br69-Kern mit dem ULD-Protokoll.

**Zusammenfassung der Ergebnisse.** Für das Kugel- und Spikulaesetup fand sich über beide Phantomlungenhälften mit dem Standard-Lungen-B157-Kern und dem harten Br69-Kern pro CT-Dosisprotokoll ein  $\overline{\text{SNR}}$ -Anstieg mit zunehmender IR-Stärke. Unabhängig vom CT-Dosisprotokoll blieb in beiden Phantomlungenhälften und Setups mit dem Br32-Kern das  $\overline{\text{SNR}}$  weitgehend unverändert. Die Mittelwerte aus den FBP-Rekonstruktionen mit dem B157- und dem Br69-Kern lagen immer unter denen der IR-Rekonstruktionen. Gleiches ließ sich bis auf drei Ausnahmen für den Br32-Kern beobachten.

Tab. 3.3: Differenzen der Minimal- und Maximalwerte des  $\overline{\text{SNR}}$  für Landmarke LL in der linken Phantomlung oder für Landmarke RL in der rechten Phantomlung im Kugel- und Spikulaesetup. Die Werte werden über alle CT-Dosisprotokolle (CTD, explizit: SH, LD, ULD) sowie den Standard-Lungen-B157-Kern, den weichen Br32-Kern und den harten Br69-Kern gelistet.

CTD	Differenzen der Maximal- und Minimalwerte von $\overline{\text{SNR}}$											
	Landmarke LL						Landmarke RL					
	Kugeln			Spikulae			Kugeln			Spikulae		
	B157	Br32	Br69	B157	Br32	Br69	B157	Br32	Br69	B157	Br32	Br69
SH	2,00	0,23	3,91	2,41	0,10	4,90	2,71	0,19	7,84	6,86	0,15	13,56
LD	1,00	0,13	1,65	1,63	0,25	3,39	2,34	0,19	8,06	5,09	0,40	14,90
ULD	1,74	0,16	4,17	2,63	0,15	4,51	2,99	0,26	9,52	4,23	0,48	12,43

Eine quantitative Betrachtung des Einflusses der Rekonstruktion auf das  $\overline{\text{SNR}}$  stellt Tab. 3.3 dar. Dabei werden die Differenzen der Minimal- und Maximalwerte für die Landmarken LL und RL im Kugel- bzw. Spikulaesetup über alle Kerne und CT-Dosisprotokolle dargestellt.

Über beide Landmarken und Setups fanden sich für den Br69-Kern die höchsten Differenzen und für den Br32-Kern die niedrigsten. Bis auf wenige Ausnahmen kamen für die Landmarken LL und RL mit Spikulae höhere Differenzen vor als mit Kugeln. Allgemein waren die Differenzen pro Kern und CT-Dosisprotokoll in der rechten Phantomlungenhälfte größer als in der linken.

Die Ergebnisse des  $t$ -Tests für den jeweiligen statistischen Vergleich der gruppierten  $\overline{\text{SNR}}$ -Werte aus Abb. 3.3, S. 58, erbrachten für den Bl57- und den Br69-Kern im Kugelsetup mit beiden Landmarken ausschließlich signifikant unterschiedliche Mittelwerte ( $p < 0,05$ ). Dies zeigte sich für das Spikulaesetup mit wenigen Ausnahmen. Es gab keine signifikant unterschiedlichen Mittelwerte für den Br32-Kern.

### 3.2.3 Kontrast-Rausch-Verhältnis

Die Übersicht in Abb. 3.4, S. 61, stellt den Mittelwert  $\overline{\text{CNR}}$  über alle Protokolle der objektiven Bildqualitätsprüfungen mit zugehöriger Standardabweichung des Mittelwerts vor. Die Ergebnisse werden für alle Kerne (Bl57, Br32, Br69) und geordnet nach Kugel- und Spikulaesetup für die Landmarken LL und RL in der linken bzw. rechten Phantomlung abgebildet. Darüber hinaus wurden die Ergebnisse gruppiert nach CT-Dosisprotokoll (SH, LD, ULD) sowie über die Rekonstruktionsalgorithmen FBP und ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$ .

**Landmarke LL.** Im Kugel- und Spikulaesetup, Abb. 3.4 (a) bis (b), gab es für Bl57 pro CT-Dosisprotokoll  $\overline{\text{CNR}}$ -Maxima bei FBP und die  $\overline{\text{CNR}}$ -Minima bei  $S_5$ . Für Br69 fanden sich in beiden Setups Maximalwerte in der Reihenfolge SH/LD/ULD mit  $S_5$ /FBP/ $S_5$ . Die Minimalwerte kamen gemäß der genannten Reihenfolge im Kugelsetup bei FBP/ $S_5$ / $S_1$  und im Spikulaesetup bei  $S_1$ / $S_5$ / $S_1$  vor. Nach gleicher Ordnung gab es für Br32 die Maxima für das  $\overline{\text{CNR}}$  im Kugelsetup bei  $S_3$ / $S_3$ / $S_1$  und die Minima bei  $S_1$ /FBP/ $S_3$ . Im Spikulaesetup führten  $S_3$ /FBP/FBP zu  $\overline{\text{CNR}}$ -Maxima, die Minima lagen ausschließlich bei  $S_5$ .

In beiden Setups zeigte sich für den Bl57-Kern eine Senkung des  $\overline{\text{CNR}}$  mit zunehmender IR-Stärke. Gegenteiliges ließ sich für den Br69-Kern in beiden Setups feststellen, mit Ausnahmen für das LD-Protokoll. Hierzu stellte sich das  $\overline{\text{CNR}}$  für den Br32-Kern über alle CT-Dosisprotokolle weitgehend unverändert dar. Dabei senkten sich im Spikulaesetup die Mittelwerte in geringem Maße mit zunehmendem IR-Stärkeniveau. Insgesamt fanden sich höhere Mittelwerte für den Br32-Kern als für andere Kerne. Die Standardabweichung des Mittelwerts war in jedem Fall im Kugelsetup höher als im Spikulaesetup. Eine Korrelation zwischen den IR-Stärkeniveaus und der Standardabweichungen des Mittelwerts ließ sich nicht identifizieren.

**Landmarke RL.** Mit dem Kugel- bzw. Spikulaesetup, Abb. 3.4 (c) bis (d), zeigten sich für Bl57 pro CT-Dosisprotokoll die maximalen  $\overline{\text{CNR}}$ -Werte bei FBP und die minimalen Werte bei  $S_5$ . Für Br69 fanden sich in beiden Setups über alle CT-Dosisprotokolle  $\overline{\text{CNR}}$ -Maxima bei  $S_5$ , ausser im Kugelsetup bei LD/ $S_3$ . Gemäß der Reihung SH/LD/ULD führten im Kugelsetup



FBP/ $S_1$ / $S_1$  zu  $\overline{\text{CNR}}$ -Minima, diese fanden sich im Spikulaesetup mit  $S_1$ /FBP/ $S_1$ . Mit Br32 gab es pro CT-Dosisprotokoll über beide Setups die maximalen Werte bei  $S_1$ , mit einer Ausnahme im Spikulaesetup für LD/FBP. Nach eingeführter Reihenfolge lagen im Kugelsetup die minimalen Werte bei  $S_5$ /FBP/ $S_3$  und im Spikulaesetup bei  $S_5$ / $S_1$ /FBP.

Für den Br157-Kern ließ sich in beiden Setups eine Senkung des  $\overline{\text{CNR}}$  mit zunehmender IR-Stärke beobachten. Gegensätzlich verhielten sich die Mittelwerte für den Br69-Kern mit jeweils einer Ausnahme im Kugelsetup mit dem SH- und dem LD-Protokoll. Mit dem Br32-Kern blieben die  $\overline{\text{CNR}}$ -Werte unabhängig vom CT-Dosisprotokoll nahezu gleich. Dabei gab es eine geringfügige Mittelwertsenkung mit zunehmender IR-Stärke im Kugelsetup mit dem SH-Protokoll und im Spikulaesetup mit den Protokollen SH und ULD. Insgesamt existierten für den Br32-Kern die höchsten Mittelwerte. Die Standardabweichung des Mittelwerts war im Spikulaesetup höher als im Kugelsetup, außer für den Br69-Kern mit dem SH-Protokoll bei  $S_1$ . Es fand sich keine Korrelation zwischen den IR-Stärkeniveaus und der Standardabweichung des Mittelwerts.

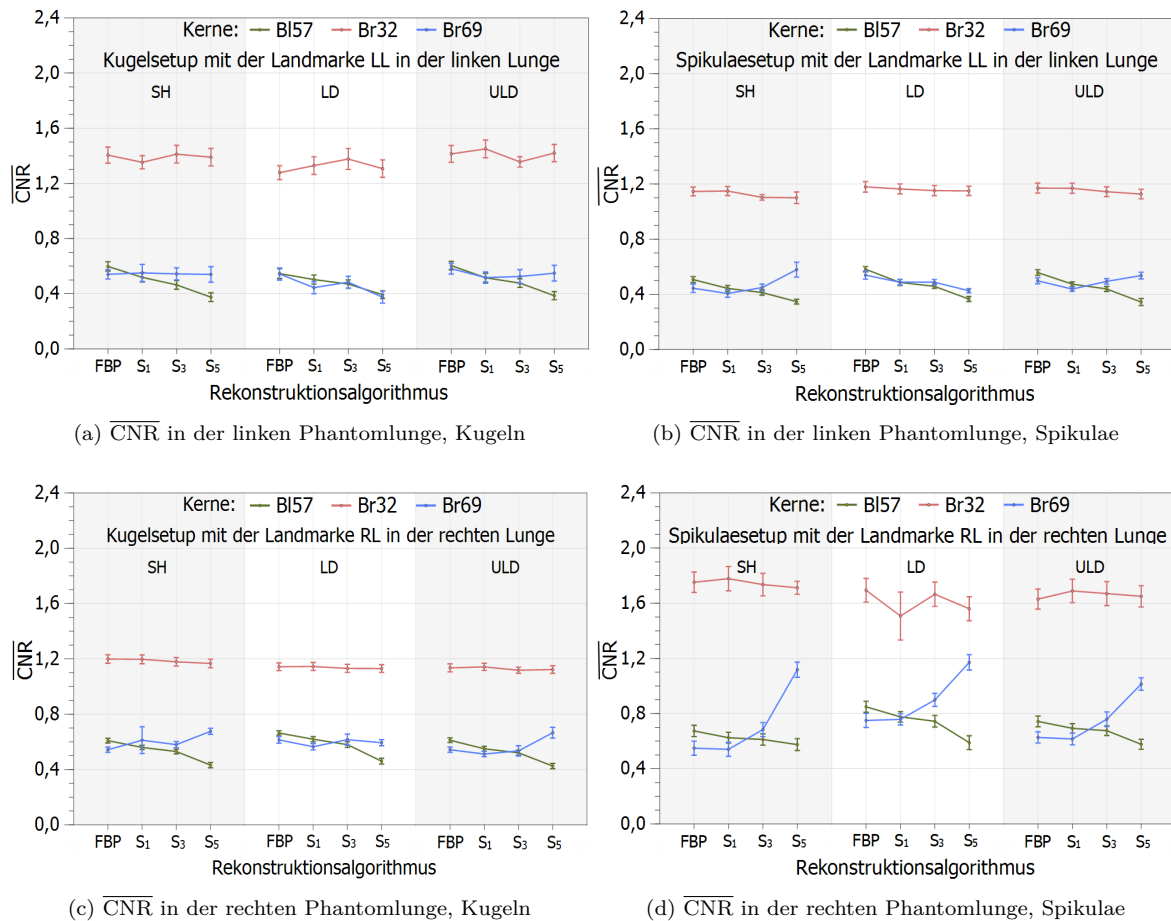


Abb. 3.4: Darstellung des Mittelwerts  $\overline{\text{CNR}}$  über alle Protokolle der objektiven Bildqualitätsprüfung mit zugehöriger Standardabweichung des Mittelwerts. Die Abbildung stellt die Ergebnisse für den Standard-Lungen-Br157-Kern (grün), den weichen Br32-Kern (rot) und den harten Br69-Kern (blau) dar. Ferner werden die Ergebnisse nach Kugel- bzw. Spikulaesetup mit der Landmarke LL in der linken Phantomlunge (a) bis (b) und der Landmarke RL in der rechten Phantomlunge (c) bis (d) aufgeführt. Überdies erfolgt die Übersicht über die Rekonstruktionsalgorithmen FBP bzw. ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$ ,  $S_5$  und gruppiert nach CT-Dosisprotokoll (SH, LD, ULD). Die Fehlerbalken zeigen die Standardabweichung der Mittelwerts.

**Zusammenfassung der Ergebnisse.** Mit dem Kugel- und Spikulaesetup kam es unabhängig vom CT-Dosisprotokoll in beiden Phantomlungenhälften mit dem Standard-Lungen-B157-Kern zu einer  $\overline{\text{CNR}}$ -Senkung mit zunehmendem IR-Stärkeniveau. Konträr verhielt es sich für den harten Br69-Kern mit Ausnahmen in drei LD-Protokollen. Indes blieb das  $\overline{\text{CNR}}$  unabhängig vom CT-Dosisprotokoll für den weichen Br32-Kern nahezu unverändert. Für den B157-Kern waren die Mittelwerte aus den FBP-Rekonstruktionen höher als die der IR-Rekonstruktionen. Gegensätzliches zeigte sich mit dem Br69-Kern mit Ausnahmen in der linken Phantomlung im LD-Protokoll. Bis auf wenige Ausnahmen fanden sich für den Br32-Kern ebenso niedrigere  $\overline{\text{CNR}}$ -Werte in den FBP-rekonstruierten Bildern gegenüber den IR-rekonstruierten Bildern.

In Tab. 3.4 ist der Einfluss der Rekonstruktion auf das  $\overline{\text{CNR}}$  durch die Differenzen der mini- und maximalen Mittelwerte für die Landmarken LL und RL im Kugel- bzw. Spikulaesetup über alle Kerne und CT-Dosisprotokolle dargestellt. Mit einer Ausnahme gab es für Landmarke LL mit Spikulae und für Landmarke RL mit Kugeln über alle CT-Dosisprotokolle die höchsten Differenzen mit dem B157-Kern und die niedrigsten mit dem Br32-Kern. Gleiches zeigte sich für den B157-Kern mit Landmarke LL und Kugeln. Es fanden sich die niedrigsten Differenzen mit dem Br69-Kern mit einer Ausnahme für das LD-Protokoll. In völligem Widerspruch verhielt sich Landmarke RL mit Spikulae. Die höchsten Differenzen gab es für den Br69-Kern und die niedrigsten für den Br32-Kern. Hier zeigte sich ein deutlich erhöhter Wert mit dem LD-Protokoll. Insbesondere gab es über das gesamte Setup mehrere Ausreißer. Mit Landmarke LL kamen mit Kugeln überwiegend höhere Differenzen vor als mit Spikulae. Gegenteiliges ergab sich bis auf zwei Ausnahmen mit Landmarke RL. Im Vergleich beider Phantomlungenhälften war die Anzahl der Höchstdifferenzen über alle Kerne und CT-Dosisprotokolle nahezu gleich.

Tab. 3.4: Differenzen der Minimal- und Maximalwerte des  $\overline{\text{CNR}}$  für Landmarke LL in der linken Phantomlung bzw. für Landmarke RL in der rechten Phantomlung im Kugel- und Spikulaesetup. Die Werte werden über alle CT-Dosisprotokolle (CTD, explizit: SH, LD, ULD) sowie den Standard-Lungen-B157-Kern, den weichen Br32-Kern und den harten Br69-Kern gelistet.

CTD	Differenzen der Maximal- und Minimalwerte von $\overline{\text{CNR}}$											
	Landmarke LL						Landmarke RL					
	Kugeln			Spikulae			Kugeln			Spikulae		
	B157	Br32	Br69	B157	Br32	Br69	B157	Br32	Br69	B157	Br32	Br69
SH	0,22	0,06	0,01	0,16	0,05	0,17	0,18	0,03	0,13	0,10	0,07	0,58
LD	0,15	0,10	0,17	0,22	0,03	0,11	0,20	0,02	0,05	0,26	0,19	0,42
ULD	0,22	0,09	0,06	0,21	0,04	0,10	0,19	0,02	0,15	0,17	0,04	0,40

Mit dem  $t$ -Test für die statistische Signifikanzuntersuchung ( $p < 0,05$ ) der jeweils gruppierten  $\overline{\text{CNR}}$ -Werte aus Abb. 3.4, S. 61, fanden sich für den B157-Kern überwiegend signifikant unterschiedliche Mittelwerte. Für Br69-Kern zeigten sich diese in der Hälfte

aller Fälle. Dabei kam mit beiden Landmarken im Kugelsetup ein überwiegender Anteil nicht signifikant unterschiedlicher Mittelwerte vor. Letztere gab es für den Br32-Kern mit Landmarke LL in beiden Setups überwiegend und mit Landmarke RL ausschließlich.

### 3.2.4 Fallbetrachtung

Die objektiven Bildqualitätsmaße aus den CT-Bildern mit der besten Erkennbarkeit der artifiziellen Lungenrundherde aus der Analyse mit der Likert-Skala, siehe Tab. 3.2, S. 57, wurden zu einer Fallbetrachtung herangezogen. Für die Simulation klinischer Bedingungen erfolgt die fallspezifische Darstellung in Abb. 3.5 mit Spikulae. Die Präferenzen der Radiologen wurden durch die FBP-Rekonstruktion und die modellbasierte IR-Rekonstruktion mit ADMIRE in der IR-Stärke  $S_3$  sowie den Standard-Lungen-BI57-Kern berücksichtigt. Die gegenübergestellten Verteilungen der SNR-Werte zeigt Abb. 3.5 (a) sowie der CNR-Werte (b), aus allen Prüfprotokollen der objektiven Bildqualitätsprüfung in der rechten und linken Phantomlunge mit den Landmarken RL und LL über alle CT-Dosisprotokolle.

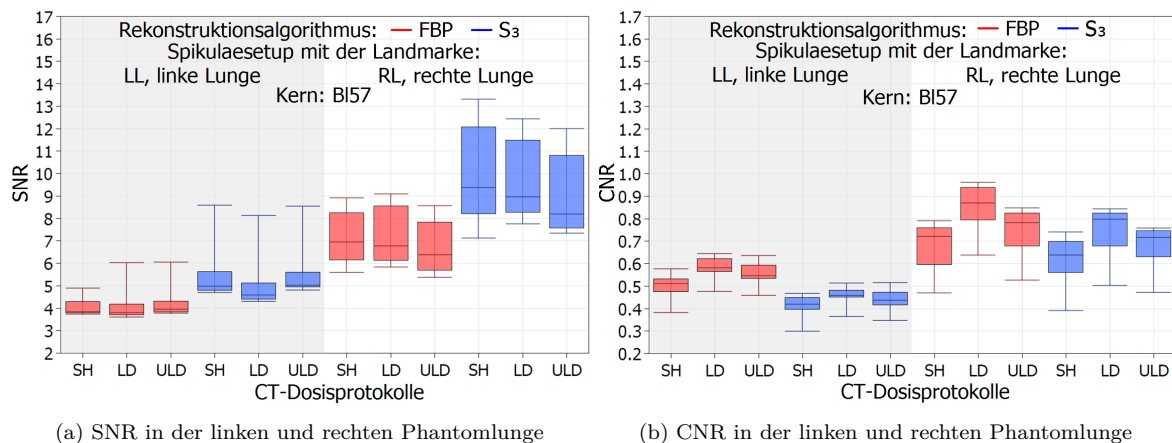


Abb. 3.5: Gemessene SNR- (a) und CNR-Verteilungen (b) aus den Rekonstruktionen mit dem BI57-Kern sowie der FBP-Rekonstruktion (rot) und ADMIRE in der IR-Stärke  $S_3$  (blau) über alle CT-Dosisprotokolle im Spikulaesetup. Die Parameter basieren auf den Resultaten der Likert-Skalabewertungen, respektive der höchsten Summe der individuell vergebenen Höchstbewertungen. Die Spikulaemodelle simulieren hierbei klinische Bedingungen. Die Verteilungen beinhalten die gewählten Parameter aus den Prüfungsprotokollen der objektiven Bildqualitätsprüfung für die Landmarken RL und LL in der rechten und linken Phantomlunge.

Für das SNR gab es pro Landmarke über alle CT-Dosisprotokolle breitere Verteilungen mit  $S_3$  als mit FBP. Alle Verteilungen waren rechtsschief. Hierbei lag der Median unterhalb des halben Interquartilsabstandes, d.h. der Differenz zwischen dem oberen und unteren Quartil. Es existierten für das CNR mit der Landmarke LL über alle CT-Dosisprotokolle breitere Verteilungen mit FBP als mit  $S_3$ . Mit der Landmarke RL kam eine breitere Verteilung für FBP und das ULD-Protokoll vor, in allen anderen Fällen zeigte sich diese mit  $S_3$ . Überdies ergaben sich mit der Landmarke LL nur rechtsschiefe Verteilungen mit einer Ausnahme. Letztere stellte sich mit FBP und dem SH-Protokoll in einer linksschiefen Verteilung dar, hier lag der Median über dem halben Interquartilsabstand. Mit der Landmarke RL ließen sich ausschließlich linksschiefe Verteilungen feststellen.

**Zusammenfassung der Ergebnisse.** Es waren sowohl die SNR- als auch die CNR-Verteilungen pro CT-Dosisprotokoll mit der Landmarke RL breiter als mit der Landmarke LL. Beide Darstellungen wiesen überlagerte Verteilungsbereiche zwischen den Rekonstruktionsalgorithmen auf. Hierbei gab es höhere SNR- bzw. geringere CNR-Werte in den Verteilungen für die modellbasierte IR-Rekonstruktion der Stärke  $S_3$  gegenüber den entsprechenden objektiven Parametern in den Verteilungen für die FBP-Rekonstruktion. Exemplarisch werden in Tab. 3.5 für die Landmarke LL aus Abb. 3.5, S. 63 die Mittelwerte  $\overline{\text{SNR}}$  und  $\overline{\text{CNR}}$  der aufgeführten SNR- und CNR-Verteilungen gelistet. Die Ergebnisse werden über die Rekonstruktionsalgorithmen FBP und ADMIRE in der IR-Stärke  $S_3$  sowie über alle CT-Dosisprotokolle vorgestellt.

Tab. 3.5: Auflistung der  $\overline{\text{SNR}}$ - und  $\overline{\text{CNR}}$ -Werte aus den objektiven Bildqualitätsprüfungen der CT-Bilder des Lungman-N1-CT-Experiments mit dem Spikulaesetup und dem Standard-Lungen-B157-Kern. Die Mittelwerte werden für die Landmarke LL der linken Phantomlung über die Rekonstruktionsalgorithmen FBP und ADMIRE in der IR-Stärke  $S_3$  über alle CT-Dosisprotokolle (CTD, explizit: SH, LD, ULD) aufgezeigt.

CTD	Objektive Bildqualitätsparameter im Spikulaesetup Landmarke LL und Kern B157			
	$\overline{\text{SNR}}$		$\overline{\text{CNR}}$	
	FBP	$S_3$	FBP	$S_3$
SH	4,20	5,53	0,51	0,41
LD	4,14	5,10	0,58	0,46
ULD	4,78	5,58	0,56	0,44

Über alle CT-Dosisprotokolle konnte ein  $\overline{\text{SNR}}$ -Anstieg für die CT-Bilder mit der modellbasierten IR-Stärke  $S_3$  gegenüber den FBP-Rekonstruktionen festgestellt werden. Explizit lag dieser Anstieg je CT-Dosisprotokoll bei: 31,60 % mit SH, 23,16 % mit LD und 16,84 % mit ULD. Gegenteilig hierzu fanden sich die  $\overline{\text{CNR}}$ -Minderungen: 18,54 % mit SH, 21,31 % mit LD und 21,33 % mit ULD.

### 3.3 QRM-Thorax

Die in diesem Kapitel vorgestellten Rangordnungstests wurden von sechs Radiologen mit unterschiedlicher Berufserfahrung durchgeführt. Entsprechend werden sie mit nachgestellter Anzahl an Berufsjahren wie folgt gekennzeichnet: Radiologe<sub>1</sub>, Radiologe<sub>1.5</sub>, Radiologe<sub>2.5</sub>, Radiologe<sub>6</sub>, Radiologe<sub>26</sub> und Radiologe<sub>32</sub>.

### 3.3.1 Rangordnungstest mit Probe 1

Die kumulativen und individuellen Ergebnisse der mittleren Rangsummen ( $\bar{R}$ ) des Rangordnungstests für die Eignungsbeurteilung der Prüfbilder aus Probe 1 zur Lungenrundherddetektion stellt Abb. 3.6 dar. Dabei wird die von den Radiologen subjektiv empfundene Eignung durch  $\bar{R}$  über den Rekonstruktionsalgorithmus ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$  bis  $S_5$  abgebildet.

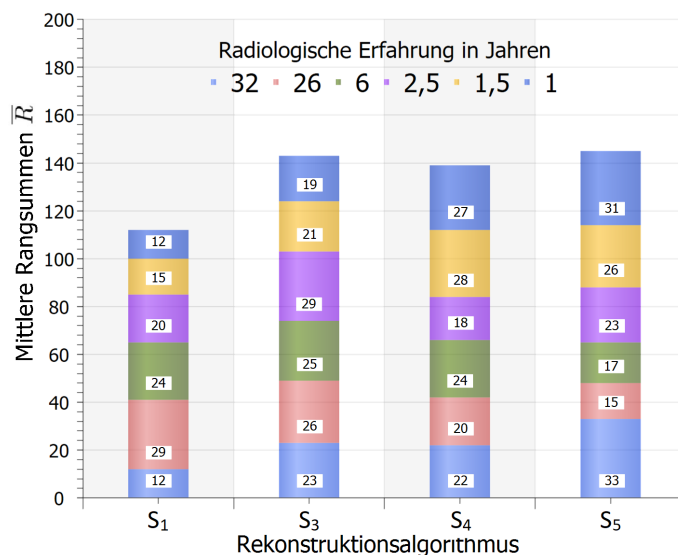


Abb. 3.6: Die von den Radiologen beurteilte Eignung der Prüfbilder des Rangordnungstests mit Probe 1 zur Lungenrundherddetektion anhand der mittleren Rangsummen  $\bar{R}$ . Die Probenparameter beliefen sich auf einen festen Phantomdurchmesser (QRM-Body mit Ring L) und die Variation der Kerne (BI57, Br32, Br69). Die Darstellung zeigt die kumulativen und individuellen Ergebnisse über die Rekonstruktionsalgorithmen ADMIRE  $S_1$ ,  $S_3$  bis  $S_5$ . Die radiologische Erfahrung in Jahren folgt der Farbkodierung: 1 Jahr (dunkelblau), 1,5 Jahre (gelb), 2,5 Jahre (lila), 6 Jahre (grün), 26 Jahre (rot) und 32 Jahre (hellblau).

Das kumulative Ergebnis erbrachte, dass die IR-rekonstruierten Prüfbilder mit mittlerer bis hoher IR-Stärke,  $\geq S_3$ , höher bewertet wurden als mit  $S_1$ . Dabei fand sich die Höchstbewertung von  $\bar{R} = 145$  mit  $S_5$ . Die niedrigste Gesamtbewertung war  $\bar{R} = 112$  mit  $S_1$ . Die kumulative Bewertung für  $S_3$  lag bei  $\bar{R} = 143$  und für  $S_4$  bei  $\bar{R} = 139$ .

Für die individuellen Ergebnisse folgt die detaillierte Leistung des niedrigsten und höchsten  $\bar{R}$  mit zugehörigem Algorithmus: Radiologe<sub>1</sub>,  $\bar{R} = 12$  mit  $S_1$  und  $\bar{R} = 31$  mit  $S_5$ , Radiologe<sub>1,5</sub>,  $\bar{R} = 15$  mit  $S_1$  und  $\bar{R} = 28$  mit  $S_4$ , Radiologe<sub>2,5</sub>,  $\bar{R} = 18$  mit  $S_4$  und  $\bar{R} = 29$  mit  $S_3$ , Radiologe<sub>6</sub>,  $\bar{R} = 17$  mit  $S_5$  und  $\bar{R} = 25$  mit  $S_3$ , Radiologe<sub>26</sub>,  $\bar{R} = 15$  mit  $S_5$  und  $\bar{R} = 29$  mit  $S_1$  und Radiologe<sub>32</sub>,  $\bar{R} = 12$  mit  $S_1$  und  $\bar{R} = 33$  mit  $S_5$ . Die Urteile von Radiologe<sub>1</sub> und Radiologe<sub>32</sub> folgten dem kumulativen Ergebnis, außer die höhere Beurteilung von  $S_4$  gegenüber  $S_3$  von Radiologe<sub>1</sub>. Radiologe<sub>1,5</sub> entsprach dem kumulativen Ergebnis nur mit  $S_1$ , Radiologe<sub>2,5</sub> und Radiologe<sub>6</sub> taten dies in keinem Fall. Die Höchstbewertungen von Radiologe<sub>2,5</sub> und Radiologe<sub>6</sub> lagen bei  $S_3$ . Im Widerspruch zum kumulativen Ergebnis waren die Beurteilungen von Radiologe<sub>26</sub> mit  $S_1$  und  $S_5$ . Übereinstimmungen gab es mit  $S_3$  und  $S_4$ .

Die individuellen Friedman-Test-Ergebnisse für die einzelnen Rangordnungstestungen mit Probe 1 werden in Tab. 3.6, S. 66, vorgestellt. Die Prüfung auf signifikante Rangvarianzen

erfolgte durch den Vergleich des Friedman–Werts  $\chi_r^2$  mit dem Quantil  $\chi_{r(q;\alpha)}^2 = 7,81$  (Signifikanzniveau  $\alpha = 5\%$ ). Zufällige Rangfolgen wurden mit  $\chi_r^2 < 7,81$  identifiziert.

Tab. 3.6: Individuelle Friedman–Werte  $\chi_r^2$  der einzelnen Rangordnungstestungen (Testung 1 bis 3) mit Probe 1.

Radiologische Erfahrung in Jahren	Friedman–Wert $\chi_r^2$ Rangordnungstest mit Probe 1		
	Testung 1	Testung 2	Testung 3
1	16,33	14,20	14,73
1,5	10,74	13,93	4,60
2,5	7,27	3,27	8,87
6	3,93	9,40	5,40
26	3,80	8,60	13,67
32	21,93	16,20	23,27

Radiologe<sub>1</sub> und Radiologe<sub>32</sub> hatten in keinem Fall zufällige Rangfolgen. Hingegen fanden sich Zufallsrangfolgen mit: Radiologe<sub>1,5</sub> für Testung 3, Radiologe<sub>2,5</sub> für Testung 1 und 2, Radiologe<sub>6</sub> für Testung 1 und 3 sowie Radiologe<sub>26</sub> für Testung 1. Insgesamt lagen in 12 von 18 Fällen keine zufälligen Rangfolgen vor.

Ein Vergleich der individuellen subjektiven Ergebnisse anhand der mittleren Rangsummen  $\bar{R}$  aus dem Rangordnungstest mit Probe 1 ist in Abb. 3.7, aufgeführt. Die Darstellung erfolgt über den Rekonstruktionsalgorithmus ADMIRE in den IR–Stärken  $S_1$ ,  $S_3$  bis  $S_5$ . Die Differenzen des höchsten und niedrigsten  $\bar{R}$  pro Algorithmus waren: 17 mit  $S_1$ , 10 mit  $S_3$  und  $S_4$  sowie 18 mit  $S_5$ . Es gab einen identischen  $\bar{R}$ –Wert bei  $S_1$  für Radiologe<sub>1</sub> und Radiologe<sub>32</sub>. Insgesamt fanden sich deutlich höhere Differenzen von  $\bar{R}$  mit  $S_5$  und  $S_1$ .

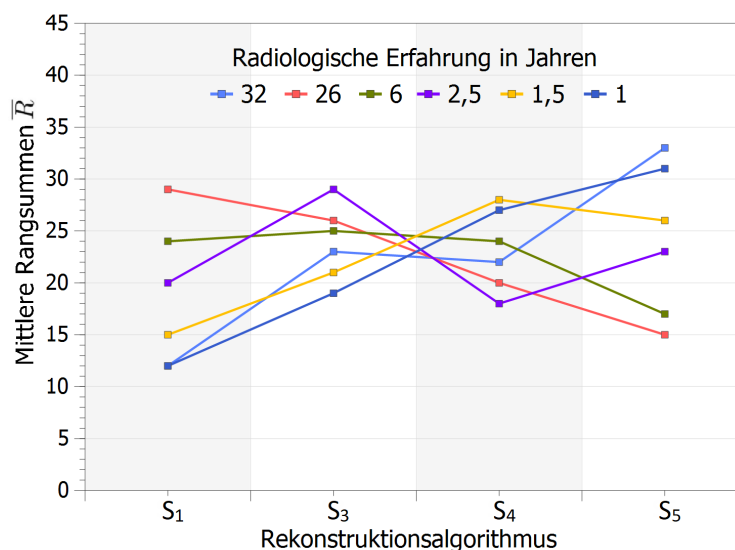


Abb. 3.7: Individuell wahrgenommene Eignung der Prüfbilder des Rangordnungstests mit Probe 1 für die Lungenrundherddetektion anhand der mittleren Rangsummen  $\bar{R}$ . Die Ergebnisse werden über den Rekonstruktionsalgorithmus ADMIRE in den IR–Stärken  $S_1$ ,  $S_3$  bis  $S_5$  dargestellt. Die Farbzuordnung der radiologischen Erfahrung in Jahren ist: 1 Jahr (dunkelblau), 1,5 Jahre (gelb), 2,5 Jahre (lila), 6 Jahre (grün), 26 Jahre (rot) und 32 Jahre (hellblau).

Die Ergebnisse der Interrater-Reliabilitätsanalyse mit dem Fleiss'- $\kappa$ -Test werden in Tab. 3.7 durch den  $\kappa_{\text{Fleiss}}$ -Koeffizienten präsentiert. Der  $\kappa_{\text{Fleiss}}$ -Koeffizient wird für die einzelnen Rangordnungstestungen mit Probe 1 über alle Rangpositionen aufgeführt. Für Rang 1 und 2 gab es ausschließlich negative Koeffizienten,  $\kappa_{\text{Fleiss}} < 0$ . Letzteres zeigte sich für Rang 3 und 4 mit jeweils zwei Testungen. Positive Koeffizienten resultierten in dem  $\kappa_{\text{Fleiss}}$ -Intervall  $[0,00; 0,20]$  für Rang 3 mit Testung 1 und für Rang 4 mit Testung 2. Zusammenfassend ergab sich eine geringe bis schwache Urteilskonkordanz der Radiologen.

Tab. 3.7: Ergebnisse der Interrater-Reliabilitätsanalyse mit dem Fleiss'- $\kappa$ -Test durch den  $\kappa_{\text{Fleiss}}$ -Koeffizienten für die einzelnen Rangordnungstestungen (Testung 1 bis 3) mit Probe 1 über alle Rangpositionen (Rang 1 bis 4).

Testung	$\kappa_{\text{Fleiss}}$			
	Rang 1	Rang 2	Rang 3	Rang 4
1	-0,014	-0,067	0,002	-0,019
2	-0,073	-0,057	-0,035	0,038
3	-0,067	-0,032	-0,089	-0,052

Die Abb. 3.8 stellt die Ergebnisse der Intrarater-Reliabilitätsanalysen aller Radiologen anhand des modifizierten  $\kappa_{\text{ney}}$ -Tests für den Rangordnungstest mit Probe 1 dar. Abgebildet ist die Funktion  $B(\kappa_{\text{ney}})$  des Anteils der übereinstimmenden Beobachtungen  $B$  von dem Koeffizienten  $\kappa_{\text{ney}}$ . Ferner wird die Inzidenz von  $\kappa_{\text{ney}}$  für eine Gesamtanzahl von  $n = 72$   $\kappa_{\text{ney}}$ -Testungen vorgestellt.

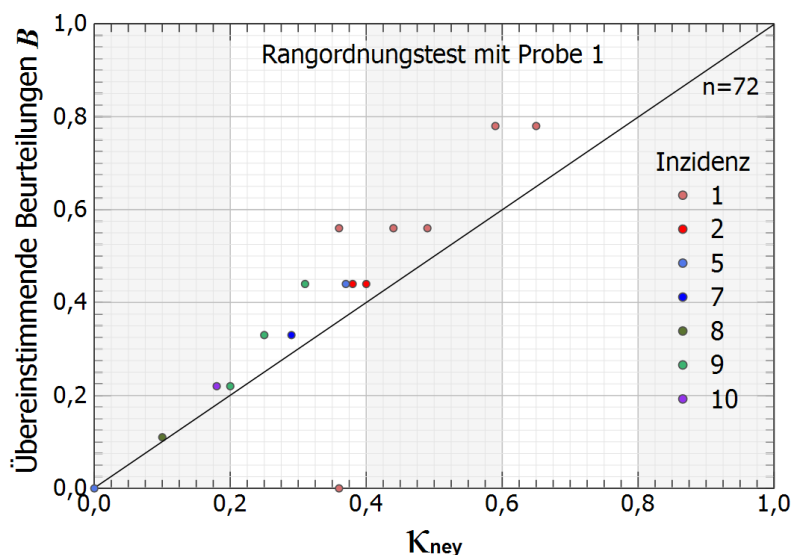


Abb. 3.8: Ergebnisse der Intrarater-Reliabilitätsanalyse mit dem modifizierten  $\kappa_{\text{ney}}$ -Test für den Rangordnungstest mit Probe 1 durch die übereinstimmenden Beobachtungen  $B$  als Funktion des  $\kappa_{\text{ney}}$ -Koeffizienten,  $B(\kappa_{\text{ney}})$ . Die Darstellung erfolgt über alle Radiologen. Für eine Gesamtanzahl von  $n = 72$   $\kappa_{\text{ney}}$ -Testungen folgt die Inzidenz für  $\kappa_{\text{ney}}$  der Farbkodierung: 1/ 2/ 5/ 7/ 8/ 9/ 10 = hellrot /dunkelrot /hellblau /dunkelblau /grün /türkis /lila.

Die  $\kappa_{\text{ney}}$ -Koeffizienten lagen zwischen 0,00 und 0,65 und die Werte von  $B$  zwischen 0,00 und 0,78. Es fanden sich die aufsummierten Inzidenzen pro Übereinstimmungsgrad mit zugehörigem  $\kappa_{\text{ney}}$ -Intervall in Klammern: 32 für „keine“  $[0,00; 0,20]$ , 34 für „minimal“  $[0,21; 0,39]$ , 5 für

„schwach“ [0,40; 0,59] und 1 für „moderat“ [0,60; 0,79]. Insgesamt lag die Urteilskonkordanz zwischen „keine“ und „moderat“. Die höchste Inzidenzsumme kam mit dem Übereinstimmungsgrad „minimal“ vor. Der Übereinstimmungsgrad „moderat“ ergab sich mit einer Inzidenz von 1.

Einen Überblick über die  $\kappa$ - und  $\kappa_{\text{ney}}$ -Koeffizienten als individuelle Ergebnisse des Cohen's  $\kappa$ - und des  $\kappa_{\text{ney}}$ -Tests im Rahmen der Intrarater-Reliabilitätsanalyse für die einzelnen Rangordnungstestungen mit Probe 1 über jede Rangposition gibt Tab. 3.8. Die  $\kappa$ -Koeffizienten einer jeweiligen Testsituation waren immer niedriger als die für  $\kappa_{\text{ney}}$ . Zudem war  $\kappa$  in 40 von 72 Fällen negativ,  $\kappa < 0$ , und somit nicht beurteilbar, Tab. 2.3, S. 37. Der beste Übereinstimmungsgrad war „schwach“ mit  $\kappa = 53$ .

Tab. 3.8: Individuelle Ergebnisse der Intrarater-Reliabilitätsanalyse mit dem Cohen's  $\kappa$ - und dem modifizierten  $\kappa_{\text{ney}}$ -Test für die einzelnen Rangordnungstestungen (Testung 1 bis 3) mit Probe 1 über alle Rangpositionen (Rang 1 bis 4) anhand der Koeffizienten  $\kappa$  und  $\kappa_{\text{ney}}$ .

Radiologische Erfahrung in Jahren		Testung		$\kappa$ und $\kappa_{\text{ney}}$							
				Rangordnungstest mit Probe 1							
				Rang 1		Rang 2		Rang 3		Rang 4	
		$\kappa$	$\kappa_{\text{ney}}$	$\kappa$	$\kappa_{\text{ney}}$	$\kappa$	$\kappa_{\text{ney}}$	$\kappa$	$\kappa_{\text{ney}}$		
1	1 mit 2	0,00	0,36	-0,08	0,25	-0,08	0,29	-0,10	0,29		
	1 mit 3	-0,36	0,31	-0,27	0,00	0,12	0,31	-0,06	0,25		
	2 mit 3	0,40	0,59	0,02	0,29	-0,37	0,00	-0,02	0,25		
1,5	1 mit 2	0,53	0,65	-0,03	0,20	0,14	0,44	-0,15	0,25		
	1 mit 3	0,00	0,29	-0,35	0,00	-0,03	0,18	-0,26	0,10		
	2 mit 3	0,17	0,37	0,08	0,31	-0,26	0,10	-0,09	0,20		
2,5	1 mit 2	0,06	0,20	0,18	0,25	0,26	0,40	-0,26	0,20		
	1 mit 3	0,21	0,38	-0,24	0,18	0,00	0,18	0,13	0,37		
	2 mit 3	0,00	0,20	0,11	0,20	0,05	0,20	-0,04	0,25		
6	1 mit 2	0,02	0,18	0,08	0,31	0,09	0,20	-0,02	0,18		
	1 mit 3	0,21	0,40	-0,20	0,10	-0,23	0,00	-0,18	0,10		
	2 mit 3	-0,15	0,25	0,00	0,20	-0,19	0,18	-0,31	0,10		
26	1 mit 2	0,33	0,49	-0,09	0,10	0,02	0,29	0,20	0,37		
	1 mit 3	0,15	0,37	0,14	0,29	-0,33	0,00	-0,33	0,10		
	2 mit 3	-0,29	0,18	-0,21	0,18	-0,02	0,18	-0,07	0,31		
32	1 mit 2	-0,25	0,31	-0,24	0,18	0,10	0,38	-0,04	0,25		
	1 mit 3	-0,06	0,25	-0,04	0,10	-0,10	0,31	0,08	0,31		
	2 mit 3	-0,07	0,31	0,07	0,29	0,18	0,37	-0,16	0,36		

Die individuellen maximalen und minimalen  $\kappa_{\text{ney}}$ -Koeffizienten kamen für einige Testungen über mehrere (maximal drei) Rangpositionen vor. Die Inzidenzbetrachtung der Maximalkoeffizienten pro Rangposition über alle Radiologen und Testungen ergab bei Rang 1 eine Höchstinzidenz von 11 und bei Rang 2 die niedrigste Inzidenz von 1. Mit den Minimalkoeffizienten erreichte Rang 2 die Höchstinzidenz von 11. Dazu lag die niedrigste Inzidenz von 3 bei Rang 4. Die konsistenten Ergebnisse der Inzidenzbetrachtung für die mini- und maximalen



Koeffizienten bei Rang 2 repräsentierten die insgesamt schlechteste Urteilskonkordanz. Für die anderen Ränge zeigten sich keine konsistenten Ergebnisse.

Im Detail lagen die maximalen  $\kappa_{\text{ney}}$ -Koeffizienten pro Testung für Radiologe<sub>1</sub> und Radiologe<sub>1,5</sub> bei Rang 1, mit einem doppelten Wert für Radiologe<sub>1</sub> bei Testung 1 mit 3 und Rang 3. Für Radiologe<sub>2,5</sub> waren die Maximalwerte bei: Testung 1 mit 2 und Rang 3, Testung 1 mit 3 und Rang 1 sowie Testung 2 mit 3 und Rang 4. Die maximalen  $\kappa_{\text{ney}}$  für Radiologe<sub>6</sub> und Radiologe<sub>26</sub> fanden sich für jeweils zwei Testungen bei Rang 1. Es gab einen weiteren Maximalkoeffizienten für Radiologe<sub>6</sub> bei Testung 1 mit 2 und Rang 2, und für Radiologe<sub>26</sub> bei Testung 2 mit 3 und Rang 4. Für Radiologe<sub>32</sub> ergaben sich zwei maximale  $\kappa_{\text{ney}}$  bei Rang 3 sowie bei Rang 4 für Testung 1 mit 3. Die minimalen  $\kappa_{\text{ney}}$  pro Testung resultierten für Radiologe<sub>1</sub> und Radiologe<sub>1,5</sub> bei Rang 2, außer bei Testung 1 mit 3 bei Rang 3. Für Radiologe<sub>2,5</sub> waren die minimalen  $\kappa_{\text{ney}}$  bei: Testung 1 mit 2 und Rang 1 bzw. 4, Testung 1 mit 3 und Rang 3 bzw. 4 sowie Testung 2 mit 3 und Rang 1 bis 3. Die Minimalkoeffizienten für Radiologe<sub>6</sub> fanden sich bei: Testung 1 mit 2 und Rang 1 bzw. 4, Testung 1 mit 3 und Rang 2 sowie Testung 2 mit 3 und Rang 4. Die minimalen  $\kappa_{\text{ney}}$  für Radiologe<sub>26</sub> zeigten sich bei: Testung 1 mit 2 und Rang 2, Testung 1 mit 3 und Rang 3 sowie Testung 2 mit 3 bei Rang 1 bis 3. Überdies kamen für Radiologe<sub>32</sub> die Minimalkoeffizienten ausschließlich bei Rang 2 vor.

### 3.3.2 Rangordnungstest mit Probe 2

Die Abb. 3.9 stellt die kumulativen und individuellen Ergebnisse der Eignungsbeurteilungen der CT-Bilder aus dem Rangordnungstest mit Probe 2 zur Lungenrundherddetektion durch die mittleren Rangsummen ( $\bar{R}$ ) vor. Die subjektiven Ergebnisse der Radiologen werden über den Rekonstruktionsalgorithmus ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$  bis  $S_5$  abgebildet.

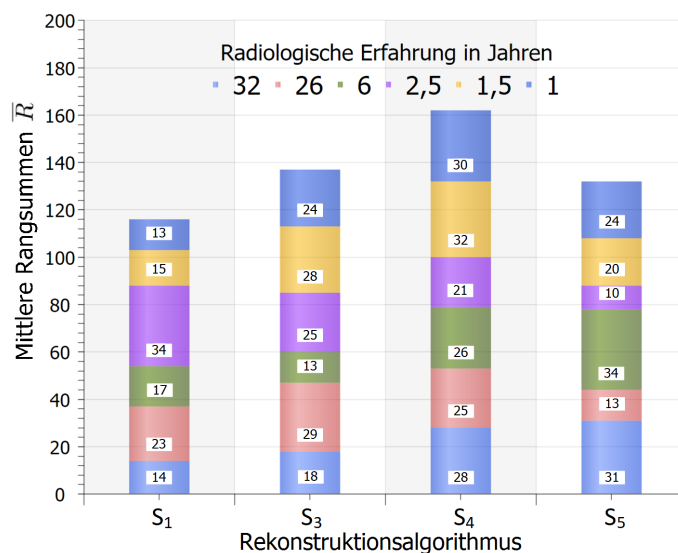


Abb. 3.9: Kumulative und individuelle Ergebnisse des Rangordnungstests mit Probe 2 über den Rekonstruktionsalgorithmus ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$  bis  $S_5$ . Die mittleren Rangsummen  $\bar{R}$  stellen die subjektiv empfundene Eignung der Prüfbilder zur Lungenrundherddetektion, bei festem Kern (Bl57) und variablem Phantomdurchmesser (QRM-Body ohne bzw. mit adaptiertem Ring M oder Ring L), dar. Die radiologische Erfahrung ist farblich gekennzeichnet: 1 Jahr (dunkelblau), 1,5 Jahre (gelb), 2,5 Jahre (lila), 6 Jahre (grün), 26 Jahre (rot) und 32 Jahre (hellblau).

Für Probe 2 resultierte das kumulative Ergebnis in der höheren Bewertung für mittlere bis hohe IR-Stärken,  $\geq S_3$ , gegenüber dem niedrigen IR-Stärkeniveau  $S_1$ . Die Höchstbewertung ergab sich für  $\bar{R} = 162$  bei  $S_4$ . Das niedrigste Ergebnis war  $\bar{R} = 116$  bei  $S_1$ . Zudem fanden sich die kumulativen Bewertungen  $\bar{R} = 137$  mit  $S_3$  und  $\bar{R} = 132$  mit  $S_5$ .

Eine detaillierte Auflistung der individuellen Ergebnisse anhand der niedrigsten und höchsten Werte für  $\bar{R}$  mit zugehörigem Algorithmus repräsentiert: Radiologe<sub>1</sub>,  $\bar{R} = 13$  mit  $S_1$  und  $\bar{R} = 30$  mit  $S_4$ , Radiologe<sub>1,5</sub>,  $\bar{R} = 15$  mit  $S_1$  und  $\bar{R} = 32$  mit  $S_4$ , Radiologe<sub>2,5</sub>,  $\bar{R} = 10$  mit  $S_5$  und  $\bar{R} = 234$  mit  $S_1$ , Radiologe<sub>6</sub>,  $\bar{R} = 13$  mit  $S_3$  und  $\bar{R} = 34$  mit  $S_5$ , Radiologe<sub>26</sub>,  $\bar{R} = 13$  mit  $S_5$  und  $\bar{R} = 29$  mit  $S_3$  und Radiologe<sub>32</sub>,  $\bar{R} = 14$  mit  $S_1$  und  $\bar{R} = 31$  mit  $S_5$ . Radiologe<sub>1,5</sub> urteilte dem kumulativen Ergebnis entsprechend. Für Radiologe<sub>1</sub> traf dies nur mit  $S_1$  und  $S_4$  zu. Gleiches erfolgte in jeweils einem Fall für Radiologe<sub>2,5</sub> mit  $S_3$ , für Radiologe<sub>6</sub> mit  $S_4$  und für Radiologe<sub>32</sub> mit  $S_1$ . Die Einschätzungen von Radiologe<sub>26</sub> standen in völligem Widerspruch zum kumulativen Ergebnis.

In Tab. 3.9 werden die individuellen Ergebnisse des Friedman-Tests für die einzelnen Rangordnungstestungen mit Probe 2 aufgeführt. Durch den Vergleich des Friedman-Werts  $\chi_r^2$  mit dem Quantil  $\chi_{r(q;\alpha)}^2 = 7,81$  (Signifikanzniveau  $\alpha = 5\%$ ) erfolgte die Untersuchung auf signifikante Rangvarianzen. Für die Testsituation  $\chi_r^2 < 7,81$  lag eine zufällige Rangfolge vor. Die Ergebnisse wiesen keine Zufallsrangfolgen auf, bis auf eine Ausnahme für Radiologe<sub>1</sub> mit Testung 1.

Tab. 3.9: Individuelle Friedman-Werte  $\chi_r^2$  der einzelnen Rangordnungstestungen (Testung 1 bis 3) mit Probe 2.

Radiologische Erfahrung in Jahren	Friedman-Wert $\chi_r^2$		
	Rangordnungstest mit Probe 2		
	Testung 1	Testung 2	Testung 3
1	5,93	14,47	13,40
1,5	11,40	23,13	21,67
2,5	24,60	18,73	16,60
6	24,33	13,67	16,60
26	12,60	12,47	15,80
32	10,33	18,73	21,40

In Abb. 3.10, S. 71, ist eine vergleichende Betrachtung der individuellen subjektiven Ergebnisse anhand der mittleren Rangsummen  $\bar{R}$  des Rangordnungstests mit Probe 2 über den Rekonstruktionsalgorithmus ADMIRE in den IR-Stärken  $S_1$ ,  $S_3$  bis  $S_5$  einsehbar. Die zum Vergleich herangezogene Differenz zwischen dem minimalen und maximalen  $\bar{R}$  für jeden Algorithmus ergab: 21 mit  $S_1$ , 16 mit  $S_3$ , 11 mit  $S_4$  und 24 mit  $S_5$ . In keinem Fall fanden sich gleiche  $\bar{R}$ -Werte pro Algorithmus. Mit  $S_5$  und  $S_1$  zeigten sich merklich höhere Differenzen.

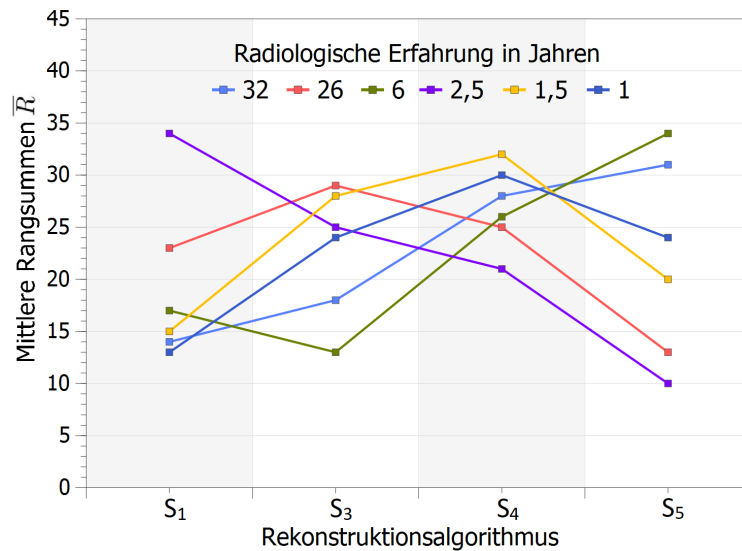


Abb. 3.10: Individuelle subjektive Ergebnisse des Rangordnungstests mit Probe 2 anhand der mittleren Rangsummen  $\bar{R}$ . Die Darstellung erfolgt über die Rekonstruktionsalgorithmen ADMIRE S<sub>1</sub>, S<sub>3</sub> bis S<sub>5</sub>. Die radiologische Erfahrung der Mediziner folgt der Farbzuordnung: 1 Jahr (dunkelblau), 1,5 Jahre (gelb), 2,5 Jahre (lila), 6 Jahre (grün), 26 Jahre (rot) und 32 Jahre (hellblau).

Mit Tab. 3.10 werden die Ergebnisse der Interrater-Reliabilitätsanalyse mit dem  $\kappa_{\text{Fleiss}}$ -Test anhand des  $\kappa_{\text{Fleiss}}$ -Koeffizienten über alle Rangpositionen und einzelne Rangordnungstestungen mit Probe 2 vorgestellt. Es lagen ausschließlich negative Koeffizienten vor,  $\kappa_{\text{Fleiss}} < 0$ , was einer geringen Urteilskonkordanz der Radiologen entsprach.

Tab. 3.10: Ergebnisse der Interrater-Reliabilitätsanalyse mit dem Fleiss'- $\kappa$ -Test durch den  $\kappa_{\text{Fleiss}}$ -Koeffizienten für die einzelnen Rangordnungstestungen (Testung 1 bis 3) mit Probe 2 über alle Rangpositionen (Rang 1 bis 4).

Testung	$\kappa_{\text{Fleiss}}$ Rangordnungstest mit Probe 2			
	Rang 1	Rang 2	Rang 3	Rang 4
1	-0,089	-0,083	-0,021	-0,019
2	-0,119	-0,092	-0,056	-0,105
3	-0,102	-0,089	-0,092	-0,137

Die Ergebnisse der Intrarater-Reliabilitätsanalysen aller Radiologen mit dem modifizierten  $\kappa_{\text{ney}}$ -Test für den Rangordnungstest mit Probe 2 sind in Abb. 3.11, S. 72, durch die Funktion  $B(\kappa_{\text{ney}})$  dargestellt. Dabei ist  $B$  der Anteil der übereinstimmenden Beobachtungen von dem Koeffizienten  $\kappa_{\text{ney}}$ . Die Inzidenz von  $\kappa_{\text{ney}}$  ist farbkodiert bei einer Gesamtanzahl von  $n = 72$   $\kappa_{\text{ney}}$ -Testungen.

Die  $\kappa_{\text{ney}}$ -Koeffizienten fanden sich zwischen 0,00 und 1,00 sowie die Werte von  $B$  zwischen 0,00 und 0,89. Es folgen Inzidenzsummen pro Übereinstimmungsgrad mit zugehörigem  $\kappa_{\text{ney}}$ -Intervall in Klammern: 19 für „keine“ [0,00; 0,20], 21 für „minimal“ [0,21; 0,39], 27 für „schwach“ [0,40; 0,59], 2 für „stark“ [0,80; 0,90] und 3 für „nahezu perfekt“,  $\kappa_{\text{ney}} > 0,90$ . Zusammenfassend variierte die Urteilskonkordanz zwischen „keine“ und „nahezu perfekt“. Die höchste Inzidenzsumme existierte mit dem Übereinstimmungsgrad „schwach“. Der Überein-

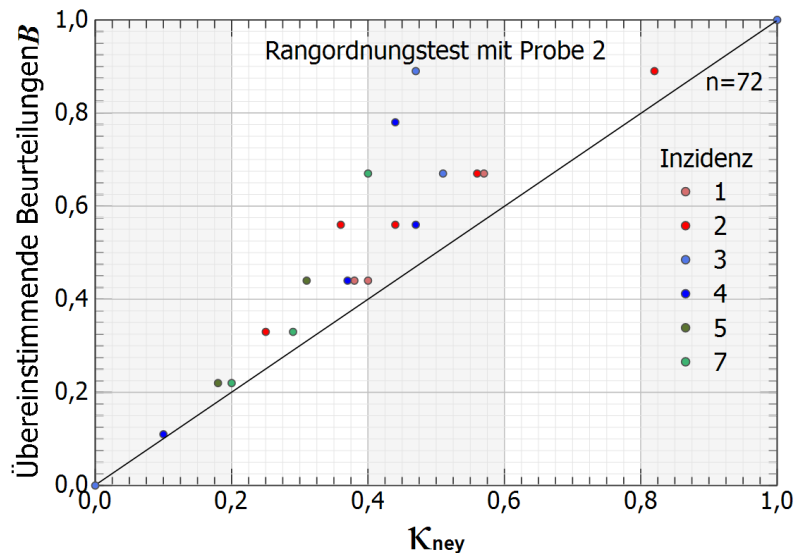


Abb. 3.11: Ergebnisse der mit dem modifizierten  $\kappa_{\text{ney}}$ -Test durchgeführten Intrarater-Reliabilitätsanalyse für den Rangordnungstest mit Probe 2 anhand der Funktion  $B(\kappa_{\text{ney}})$  des Anteils der übereinstimmenden Beobachtungen  $B$  von dem Koeffizienten  $\kappa_{\text{ney}}$ . Über alle Radiologen ergibt sich eine Gesamtanzahl von  $n = 72$   $\kappa_{\text{ney}}$ -Testungen. Die Inzidenz für  $\kappa_{\text{ney}}$  folgt der Farbkodierung: 1/ 2/ 5/ 7/ 8/ 9/ 10 = hellrot /dunkelrot /hellblau /dunkelblau /grün /türkis /lila.

stimmungsgrad „moderat“ mit dem  $\kappa_{\text{ney}}$ -Intervall  $[0,60; 0,79]$  kam nicht vor.

In Tab. 3.11, S. 73, werden die individuellen Ergebnisse der Intrarater-Reliabilitätsanalyse mit dem Cohen's  $\kappa$ - und dem modifizierten  $\kappa_{\text{ney}}$ -Test der einzelnen Rangordnungstestungen mit Probe 2 durch die Koeffizienten  $\kappa$  und  $\kappa_{\text{ney}}$  über alle Rangpositionen aufgeführt. Für dieselbe Testsituation wies  $\kappa$  niedrigere Koeffizienten auf als  $\kappa_{\text{ney}}$ . Eine Ausnahme gab es für den Fall  $\kappa = 1$ , diese Koeffizienten waren identisch zu denen mit  $\kappa_{\text{ney}}$ . Zudem fanden sich 19 von 72 nicht beurteilbare  $\kappa$ -Koeffizienten mit negativem Vorzeichen,  $\kappa < 0$ , Tab. 2.3, S. 37. Für  $\kappa = 1$  zeigte sich der beste Übereinstimmungsgrad mit „nahezu perfekt“.

Die individuellen Maximal- und Minimalkoeffizienten von  $\kappa_{\text{ney}}$  kamen teils auf mehreren (maximal drei) Rangpositionen pro einzelner Rangordnungstestung vor. Mit der Inzidenzbetrachtung der Maximalkoeffizienten pro Rangposition über alle Radiologen und Testungen fand sich die Höchstinzidenz von 12 bei Rang 1 sowie der niedrigste Inzidenzwert von 1 bei Rang 3. Konsistente Ergebnisse ließen sich mit der Höchstinzidenz von 11 der Minimalkoeffizienten bei Rang 3 und der niedrigsten Inzidenz von 1 bei Rang 1 feststellen. Somit konnte mit der Inzidenzbetrachtung die insgesamt beste Urteilskonkordanz bei Rang 1 und die insgesamt schlechteste bei Rang 3 festgestellt werden.

Im Einzelnen fanden sich die  $\kappa_{\text{ney}}$ -Maximalkoeffizienten pro Testung für Radiologe<sub>1</sub>, Radiologe<sub>6</sub> und Radiologe<sub>26</sub> bei Rang 1. Eine Ausnahme bildete ein doppelter Koeffizient für Radiologe<sub>1</sub> bei Testung 1 mit 3 und Rang 4. Für Radiologe<sub>1,5</sub> lagen die maximalen Koeffizienten bei Rang 2, bis auf eine Ausnahme bei Testung 2 mit 3 und Rang 2. Es fanden sich die Maximalwerte für Radiologe<sub>2,5</sub> bei: Testung 1 mit 2 und Rang 2 bzw. 4, Testung 1 mit 3 und Rang 1 sowie Testung 2 mit 3 und Rang 2. Zudem gab es maximale  $\kappa_{\text{ney}}$ -Koeffizienten für Radiologe<sub>32</sub> bei Rang 1, außer bei Testung 1 mit 3 und Rang 3 bzw. 4. Die Minimalkoeffizienten von  $\kappa_{\text{ney}}$

Tab. 3.11: Individuelle Ergebnisse der Intrarater-Reliabilitätsanalyse mit dem Cohen's  $\kappa$ - und dem modifizierten  $\kappa_{\text{ney}}$ -Test für die einzelnen Rangordnungstestungen (Testung 1 bis 3) mit Probe 2 über alle Rangpositionen (Rang 1 bis 4) anhand der Koeffizienten  $\kappa$  und  $\kappa_{\text{ney}}$ .

Radiologische Erfahrung in Jahren	Testung	$\kappa$ und $\kappa_{\text{ney}}$							
		Rangordnungstest mit Probe 2							
		Rang 1		Rang 2		Rang 3		Rang 4	
		$\kappa$	$\kappa_{\text{ney}}$	$\kappa$	$\kappa_{\text{ney}}$	$\kappa$	$\kappa_{\text{ney}}$	$\kappa$	$\kappa_{\text{ney}}$
1	1 mit 2	-0,15	0,31	-0,16	0,10	0,05	0,29	-0,29	0,10
	1 mit 3	-0,15	0,31	-0,02	0,20	0,07	0,29	0,13	0,31
	2 mit 3	1,00	1,00	0,15	0,37	-0,11	0,20	-0,62	0,00
1,5	1 mit 2	0,00	0,40	0,07	0,20	-0,20	0,25	0,28	0,47
	1 mit 3	-0,09	0,36	0,07	0,20	-0,13	0,18	0,29	0,51
	2 mit 3	0,00	0,47	1,00	1,00	0,05	0,20	0,07	0,29
2,5	1 mit 2	0,00	0,44	0,20	0,47	0,06	0,37	0,00	0,47
	1 mit 3	0,00	0,47	-0,07	0,37	-0,20	0,10	0,00	0,31
	2 mit 3	-0,17	0,40	0,25	0,47	0,26	0,40	0,22	0,44
6	1 mit 2	0,78	0,82	0,49	0,56	0,00	0,40	0,00	0,40
	1 mit 3	0,78	0,82	0,45	0,56	0,00	0,44	0,00	0,44
	2 mit 3	1,00	1,00	0,51	0,57	0,23	0,51	0,23	0,51
26	1 mit 2	0,18	0,40	0,14	0,29	0,03	0,18	0,17	0,38
	1 mit 3	0,16	0,40	-0,19	0,00	-0,15	0,20	0,17	0,37
	2 mit 3	0,08	0,44	-0,31	0,00	-0,03	0,20	0,04	0,31
32	1 mit 2	0,28	0,47	0,02	0,29	0,00	0,18	0,00	0,18
	1 mit 3	0,00	0,25	-0,03	0,18	0,00	0,40	0,31	0,40
	2 mit 3	0,00	0,44	0,13	0,29	0,03	0,29	-0,13	0,29

existierten für Radiologe<sub>1</sub> bei: Testung 1 mit 2 und Rang 3 bzw. 4, Testung 1 mit 3 und Rang 2 sowie Testung 2 mit 3 und Rang 4. Es gab zwei minimale Koeffizienten für Radiologe<sub>1,5</sub> bei Rang 3 und einen bei Testung 1 mit 2 und Rang 2. Für Radiologe<sub>2,5</sub> zeigten sich die Ergebnisse bei Rang 3 mit einem doppelten Minimalkoeffizienten bei Testung 2 mit 3 und Rang 1. Über alle Testungen für Radiologe<sub>6</sub> lagen bei Rang 3 und 4 die Minimalkoeffizienten. Letztere hatte Radiologe<sub>26</sub> in zwei Fällen bei Rang 2 sowie bei Testung 1 mit 3 und Rang 3. Die Ergebnisse mit Radiologe<sub>32</sub> waren bei: Testung 1 mit 2 und Rang 3 bzw. 4, Testung 1 mit 3 und Rang 3 sowie Testung 2 mit 3 und Rang 2 bis 3.

**Zusammenfassung der Ergebnisse.** Die Rangordnungstests mit den QRM-Phantom-Prüfbildern der Proben 1 und 2 wiesen nach, dass mittlere bis hohe Stärken der modellbasierten IR,  $\geq S_3$ , von den Radiologen für die Lungenrundherddetektion präferiert wurden. Insbesondere für Probe 2 mit dem klinisch verwendeten Standard-Lungen-BI57-Kern und variierendem Phantombdiameter stellte sich eine deutliche Präferenz für  $S_4$  dar. Dies bestätigten zudem separate Analysen der Prüfbilder einzelner Phantombdiameter innerhalb der Probe 2, einsehbar in Anhang A.7. Eine Korrelation zwischen der radiologischen Erfahrung und der subjektiven Wahrnehmung der CT-Bilder konnte für beide Proben

nicht festgestellt werden. Zudem erwies sich mit Probe 1 ein Drittel aller Rangfolgen als zufällig. Für Probe 2 gab es nur eine Zufallsrangfolge. Der Vergleich der individuellen mittleren Rangsummen stellte für beide Proben über alle Radiologen die größte Differenz der subjektiven Wahrnehmung der CT-Bilder für die IR-Stärke  $S_5$  fest.

Beide Rangordnungstests erbrachten für die Interrater-Reliabilität den Übereinstimmungsgrad „gering“, lediglich in zwei Fällen mit Probe 1 war dieser „schwach“. Die Intrarater-Reliabilität entsprach über alle Radiologen für den Rangordnungstest mit Probe 1 einem Übereinstimmungsgrad zwischen „keine“ und „moderat“. Letzterer lag für den Rangordnungstest mit Probe 2 zwischen „keine“ und „nahezu perfekt“. Die hierbei durchgeführte Inzidenzbetrachtung des maximalen und minimalen  $\kappa_{\text{ney}}$ -Koeffizienten pro Rang über alle Testungen und Radiologen identifizierte für den Rangordnungstest mit Probe 1 die insgesamt schlechteste Urteilskonkordanz auf Rang 2. Für die beste Urteilskonkordanz ergaben sich keine eindeutigen Ergebnisse. Zudem stellte sich mit Probe 2 die insgesamt beste Urteilskonkordanz bei Rang 1 und die insgesamt schlechteste bei Rang 3 dar.

### 3.4 Strahlendosisbewertung

Die CT-spezifischen Dosisgrößen des scannerintegrierten Dosisreports des Lungman-N1-CT-Experiments sowie die aus dem DLP über einen thoraxspezifischen Konversionsfaktor bestimmte effektive Dosis  $E$  sind in Tab. 3.12 über das Kugel- und Spikulaesetup pro CT-Dosisprotokoll aufgeführt.

Tab. 3.12: Scannerintegrierte CT-spezifische Dosisgrößen des Dosisreports für das Lungman-N1-CT-Experiment (CTDI<sub>vol</sub> und DLP) sowie die geschätzte effektive Dosis  $E$ . Die Ergebnisse sind für das Kugel- und Spikulaesetup über alle CT-Dosisprotokolle (CTD, explizit: SH, LD, ULD) aufgeführt worden.

Setup	CTD [mGy]	CT-spezifische Dosisgrößen und effektive Dosis $E$		
		CTDI <sub>vol</sub> [mGy]	DLP [mGy cm]	$E$ [mSv]
Kugel	SH	2,99	106,30	1,49
	LD	2,65	94,00	1,32
	ULD	1,31	46,50	0,65
Spikulae	SH	3,01	106,90	1,50
	LD	2,65	94,00	1,32
	ULD	1,31	46,50	0,65

Es fanden sich ausschließlich für das SH-Protokoll im Vergleich über beide Setups unterschiedliche Werte der Dosisgrößen. Der Vergleich mit dem diagnostischen Referenzwert für Thorax-CT von 10 mGy ergab für das Kugelsetup eine CTDI<sub>vol</sub>-Minderung von: 70,1 % mit

---

SH, 73,5 % mit LD und 86,9 % mit ULD. Für das Spikulaesetup fanden sich: 69,9 % mit SH, 73,5 % mit LD und 86,9 % mit ULD. Alle Minderungen zeigten sich gegenüber dem diagnostischen Referenzwert höchstsignifikant ( $p < 0,00001$ ), Kap. 2.2.4.

## 4 Diskussion

### Konstanzprüfung

An den Messtagen der Lungenphantom-CT-Scans lag für das kalibrierte CT-System eine stabile Akquisition und Rekonstruktion vor.

**Statistische Auswertung.** Die mittleren SNR- und CNR-Werte aus den gleich rekonstruierten Prüfbildern der Konstanzprüfung vor und nach dem Lungenphantom CT-Experiment wiesen keine signifikanten Unterschiede auf ( $t$ -Test,  $p < 0,05$ ). Damit konnten für die Phantomschans gleiche Mess- und Rekonstruktionsvoraussetzungen des CT-Systems nachgewiesen werden. Darüber hinaus bestätigen die konsistenten Ergebnisse der Konstanzprüfungen, dass über die Dauer des CT-Experiments die Ausrichtung des Positionslasers gleich blieb und die Fahrautomatik des CT-Tisches ohne merkbaren Positionsversatz arbeitete.

### Qualitätsanalyse mit HU-Profilen

Die nominalen HU wurden in den Lungman-N1-CT-Bildern für die Kugelmodelle mit  $-50$  HU und  $-690$  HU korrekt wiedergegeben. Dies bestätigte sich nicht für das  $100$  HU-Kugelmodell. Es konnte kein Einfluss der Rekonstruktion auf die mittleren HU-Profilwerte der Modelle festgestellt werden. Hingegen zeigte sich ein deutlicher Einfluss der Rekonstruktion auf die Standardabweichung der Stichprobe.

**Cupping-Artefakte.** Die HU-Profile wiesen erhöhte Werte in den Randbereichen auf, Abb. 3.1, S. 53 und Anhang A.6. Dieses Phänomen fand sich auch in einer Studie von (Ese et al., 2019) und wurde als Cupping-Artefakt identifiziert, welcher durch den Aufhärtungseffekt, Kap. 2.1.3, hervorgerufen wird. Im Randbereich einer Kugel findet gegenüber dem Zentrum eine geringere Strahlhärtung der polychromatischen Strahlung statt. Trifft diese geringer aufgehärtete Strahlung auf den Detektor, resultieren höhere CT-Zahlen für den Randbereich gegenüber dem Zentrum (Jayamani et al., 2019). In (Barrett und Keat, 2004) wurde der charakteristische HU-Profilverlauf eines Zylinderphantoms als schalenförmig (engl.: cupped) beschrieben. Die HU-Profile der  $100$  HU-Kugel wiesen für kein CT-Bild am rechten Rand Cupping-Artefakte auf. Die erhöhte Strahlhärtung des rechten Randbereichs begründet sich durch die Überlagerung der Kugel mit einer synthetischen Gefäßgabel. Für die  $-50$  HU-Kugel gab es über alle Profile geringe Formvarianten der Cupping-Artefakte. Eine mögliche Ursache liegt in der manuellen Segmentierungsmethode. Diese geht mit geringen Positionsvarianten der Segmentierungslinie einher. Eine geeignete Segmentierungsalternative wären moderne dreidimensionale computergestützte Quantifizierungsmethoden (Spekowius und Wendler, 2006, S. 397).



**Wiedergabe der nominalen HU.** Die korrekte Wiedergabe der nominalen HU in den Lungman-N1-CT-Bildern der  $-50$  HU- und  $-690$  HU-Kugel wurde durch den statistischen Vergleich ( $t$ -Test,  $p < 0,05$ ) mit dem zugehörigen mittleren HU-Profilwert nachgewiesen, Kap. 3.1. Hingegen fanden sich für die 100 HU-Kugel signifikante Abweichungen zum Nominalwert. Eine zuverlässige Lungenrundherdquantifizierung bei einer Screeningprozedur setzt präzise HU-Messungen voraus. Anhand der Volumensegmentierung von Lungenrundherden auf einer heterogenen Lungenumgebung wird in (Spekowitz und Wendler, 2006 S. 396 ff.) von einer segmentierungsabhängigen Schwankung der mittleren HU-Werte entlang des Kugelradius (vom Zentrum hin zum Rand) von  $\pm 100$  HU berichtet. Die einzelnen HU-Profilwerte der 100 HU-Kugel lagen in der genannten Schwankungsbreite mit zwei Ausnahmen in der FBP-Rekonstruktion. Dies legt nahe, dass auch die signifikanten HU-Abweichungen für die 100 HU-Kugel eine akzeptable Wiedergabe repräsentieren. Die Betrachtung der Ausnahmen motivierte die Untersuchung, ob die HU-Profilwerte in den CT-Bildern mit modellbasierter IR-Rekonstruktion präziser ausfallen als mit der FBP-Rekonstruktion.

**Mittlere HU-Profilwerte und Standardabweichung der Stichprobe.** Der statistische Vergleich ( $t$ -Test,  $p < 0,05$ ) der mittleren HU-Profilwerte pro Kugel zeigte unabhängig von der Rekonstruktion keine signifikanten Unterschiede, Kap. 3.1. Damit konnte kein Einfluss der Rekonstruktion auf die Profilmittelwerte festgestellt werden. Für jede Kugel glättete sich das HU-Profil mit zunehmender modellbasierter IR-Stärke. Die verwendete modellbasierte IR-Implementation führt mit zunehmendem IR-Stärkeniveau eine verstärkte Rauschunterdrückung durch (Ramirez-Giraldo et al., 2017). Die meisten Profilinebenheiten gab es mit der FBP-Rekonstruktion. Die Standardabweichung der Stichprobe wies hierzu konsistente Ergebnisse auf, mit Höchstwerten für die FBP-Rekonstruktion und geringeren Werten mit zunehmendem modellbasierten IR-Stärkeniveau. Dieser Effekt legte das Potential der modellbasierten IR-Rekonstruktion offen, eine präzise HU-Messung in Lungenrundherden mit HU-Profilen zu ermöglichen (König, Kühl et al., 2020).

**Limitationen.** Die verwendete Open-Source-Bildverarbeitungssoftware Fiji Image J (National Institutes of Health, 2019), Version 1.52p, war für eine reproduzierbare Segmentierung der HU-Profile nicht in jeder Hinsicht zufriedenstellend. Eine Segmentierungslinie durch ein Herdmodell eines Prüfbildes war nicht auf ein anderes Prüfbild übertragbar und musste anhand gezählter Pixel neu positioniert werden. Daher wurde auf die Segmentierung kleinerer Modelle verzichtet. Außerdem befand sich Klebstoff zur Fixierung im Bereich der Kugeln, welcher als zusätzliches Material Einfluss auf die Messungen im Kugelrandbereich haben könnte.

### Lungman–N1–Phantom

Die modellbasierte IR–Rekonstruktion stellt im Zusammenhang mit der Niedrigdosis–CT eine Alternative zur FBP–Rekonstruktion dar und ist somit für das Lungenkrebs–Screening geeignet.

**Subjektive Wahrnehmung der CT–Bilder bei variablen Kernen.** Die Erkennbarkeit der artifiziellen Lungenrundherde wurde von den erfahrenen Radiologen in den CT–Bildern des Lungman–N1–Phantoms mit FBP–Rekonstruktion höher bewertet als mit modellbasierter IR–Rekonstruktion und am höchsten mit dem Standard–Lungen–Bl57–Kern. Der weniger erfahrene Befunder beurteilte in den FBP– und modellbasierten IR–Bildern gleichermaßen mit Höchstbewertungen. Ähnliche Ergebnisse gab es für den harten Br69–Kern. Hingegen wurde die Erkennbarkeit der Kugeln und Spikulae mit dem weichen Br32–Kern unabhängig von der Rekonstruktion und dem CT–Dosisprotokoll deutlich schlechter eingeschätzt. Auffällig waren hierbei hohe individuelle Wahrnehmungsunterschiede der IR–Stärkeniveaus pro CT–Dosisprotokoll, Abb. 3.2, S. 55.

Die Ergebnisse implizieren, dass mit modellbasierter IR–Rekonstruktion eine bessere subjektiv empfundene Erkennbarkeit der Kugeln und Spikulae für den Bl57– und Br69–Kern erreicht wird als mit dem Br32–Kern. Die Ursache hierfür kann in den rauschreduzierenden und glättenden Eigenschaften weicher Kerne und dem damit einhergehenden Bildeindruck liegen (Buzug, 2014, S. 189 ff.). Zusätzlich könnte die Rauschreduktion der modellbasierten IR–Implementation (Ramirez–Giraldo et al., 2017) den kernbedingten geglätteten Bildeindruck verstärken.

Hinsichtlich der Diagnoseleistung in Lungen–CT–Bildern, lassen die Wahrnehmungsunterschiede mit dem Br32–Kern weniger konsistente Individualergebnisse für verschiedene Niedrigdosis–CT–Protokolle bei Anwendung weicher Kerne mit modellbasierten IR–Verfahren vermuten. Eine ähnliche Beobachtung machten Blazis et al. für die Detektionsleistung eines kommerziellen DL–CAD–Systems mit statistisch bzw. modellbasiert IR–rekonstruierten CT–Bildern bei variablen Rekonstruktionsparametern (Blazis et al., 2021). Für ein festes CT–Dosisprotokoll und die IR–Stärke  $S_4$  waren die System–Leistungsparameter für die weichen Kerne mit verschiedenen IR–Implementationen inkonsistenter als für andere Kerne. Dieser Effekt konnte durch einen geeigneten Rekonstruktionsparameter (Schichtdicke von 3 mm) reduziert werden. Dabei war die System–Sensitivität mit der modellbasierten IR–Rekonstruktion bei sonst ähnlichen Leistungsparametern gegenüber anderen Verfahren deutlich erhöht. In einer Analyse des Potentials von DL–ML–Systemen zur Lungenrundherderkennung berichten Chamberlin et al. von der exzellenten Übereinstimmung der Ergebnisse der Software mit denen von Radiologen (Chamberlin et al., 2021). Diese Beobachtungen lassen eine Effizienz der modellbasierten IR–Rekonstruktion mit weichen Kernen bei optimalen Rekonstruktionsparametern und CT–Dosisprotokollen für die Lungenrundherddetektion mit humaner und maschineller Befundung vermuten. Die CT–Dosisprotokolle und Rekonstruktionsparameter für ein Lungenkrebs–Screening zu begrenzen und zu optimieren ist somit naheliegend. Darüber hinaus machen diese Ergebnisse die Notwendigkeit einer vereinheitlichten Methodik im Lungenkrebs–Screening, d. h. Einigung auf standardisierte CT–Dosisprotokolle und Rekonstruktionsparameter, deutlich.

**Subjektive Wahrnehmung der CT-Bilder bei variabler Rekonstruktion.** Anhand einer sechsstufigen Likert-Skala wurde die subjektiv wahrgenommene beste Erkennbarkeit für Kugeln und Spikulae in den Lungman-N1-CT-Bildern mit der FBP-Rekonstruktion und dem Standard-Lungen-Bl57-Kern festgestellt. Geometrie- bzw. kernunabhängig und geordnet nach der Summe der individuellen Höchstbewertungen befand sich die IR-Stärke  $S_3$  innerhalb einer Rangfolge hinter der FBP-Rekonstruktion, Tab. 3.2, S. 57.

Die subjektiv beste Erkennbarkeit der artifiziellen Lungenrundherde in den FBP-Bildern ist vermutlich auf die Gewöhnung der Radiologen bei der Befundung zurückzuführen (König et al., 2019). Das modellbasierte IR-Verfahren auf dem Stärkeniveau  $S_3$  könnte sich möglicherweise als Alternative zur FBP-Rekonstruktion für ein Lungenkrebs-Screening eignen. Anhand statistisch IR-rekonstruierter Niedrigdosis-CT-Bilder der Lunge identifizierten Macri et al. das IR-Stärkeniveau  $S_3$  als effektive Rekonstruktionsmethode mit einer exzellenten bis guten anatomischen Bildqualität (Macri et al., 2016). Ferner empfiehlt der Hersteller der in dieser Dissertation verwendeten modellbasierten IR-Implementation das Niveau  $S_3$  als Standard-einstellung zur Rekonstruktion (Ramirez-Giraldo et al., 2017; Siemens Healthineers, 2020). Die Ergebnisse von Macri et al. beruhten auf den subjektiven Urteilen von zwei Radiologen mit unterschiedlicher Befundungserfahrung. Dabei wurde die anatomische Bildqualität in den CT-Bildern anhand großer Strukturen (z. B. Aorta, Spinalkanal etc.) auf einer fünfstufigen Likert-Skala bewertet. Die vorliegende Promotion verwendete eine sechsstufige Likert-Skala und setzte zur Bewertung kleine Targets an. Dazu wurden die Urteilkategorien enger gefasst, um Ausreißer besser zu identifizieren. Mit der IR-Stärke  $S_5$  gab es unabhängig von der Geometrie und den Kernen die schlechtesten Bewertungen für die Lungenrundherderkennbarkeit, wogegen Macri et al. die IR-Stärke  $S_5$  mit einem ULD-Protokoll als guten Kompromiss vorschlugen. Möglicherweise lag dies an der Wahl der weitgefassten Bewertungskriterien sowie an der Zusammenschau mit objektiven Bildqualitätsmaßen und Dosisparametern. Diese widersprüchlichen Ergebnisse motivierten die Untersuchung der Eignung einzelner modellbasierter IR-Stärkeniveaus für eine Screeningprozedur durch weitere subjektive Analysen mit Betrachtung objektiver Bildqualitätsmaße und der Strahlendosis. Es erfolgten hierbei Konkordanzanalysen für mehr als drei Radiologen um eine höhere Qualität der Statistik zu erzielen.

Die Kugeln mit den Durchmessern 5 mm und 3 mm wurden in keiner Prüfbildserie (Schichtdicke von 5 mm, Schrittweite von 5 mm) von keinem Radiologen detektiert. In der US-amerikanischen NLST-Studie sind Lungenrundherde mit einem Durchmesser größer oder gleich 4 mm als positiv definiert worden, in der europäischen NELSON-Studie erfolgte dies ab einem Durchmesser von rund 10 mm bei einem definierten Herdvolumen von  $500 \text{ mm}^3$  (Center for Statistical Sciences, 2011; Koning et al., 2018). Entscheidend für die Prognose eines Patienten ist die Erkennung einer Lungenkrebs-Erkrankung im frühen Stadium (Siegmund-Schultze, 2019). Die radiologische Diagnostik bei der Lungenrundherddetektion schließt die Betrachtung von coronaren CT-Schichten mit einer geringen Schichtdicke (z. B. 1 mm) ein (Kühl, 2018). Daher ist davon auszugehen, dass die kleinen Kugeln bei einer vollständigen Befundungsprozedur detektiert worden wären. Die Diagnoseleistung in modellbasiert IR-rekonstruierten CT-Bildern in Abhängigkeit der Schichtebene und -dicke wurde in dieser Dissertation nicht betrachtet.

Ferner verdeutlichen die genannten Ergebnisse die Notwendigkeit einer computerassistierten Diagnose bzw. einer KI-gestützten Lungenrundherderkennung als Zweitbefundung. Erste Erkenntnisse über die KI-gestützte Befundung demonstrierten eine verbesserte Genauigkeit und Sensibilität bei der Lungenrundherddetektion (Zhang et al., 2021).

**Objektive Bildqualitätsmaße.** Unabhängig von der Geometrie, dem CT-Dosisprotokoll und der Lage der Messregion im Phantom stieg das SNR mit zunehmender modellbasierter IR-Stärke und gegenüber der FBP-Rekonstruktion für den B157- und Br69-Kern signifikant an ( $p < 0,05$ ). Ebenso verhielt sich das CNR mit wenigen Ausreißern im LD-Protokoll ( $p < 0,05$ ). Für den Br32-Kern zeigten sich hingegen das SNR und CNR weitgehend unverändert, mit Ausreißern in der linken Lunge im Kugel- sowie in der rechten Lunge im Spikulaesetup, Abb. 3.3, S. 58 und Abb. 3.4, S. 61.

Der mit statistischer IR-Rekonstruktion erreichte SNR- und CNR-Anstieg wird von Macri et al. als Verbesserung der Bildqualitätsmaße interpretiert (Macri et al., 2016). Demgemäß verbesserte sich das SNR mit zunehmendem modellbasierten IR-Stärkeniveau für die Kerne B157 und Br69 sowie das CNR für den Br69-Kern. Eine Ursache für die Variation der Bildqualitätsmaße zwischen den artifiziellen Lungenhälften könnte das höhere Rauschen in der linken Messregion gewesen sein. Unterschiedliche Ergebnisse zwischen dem Kugel- und Spikulaesetup sind durch Positionsvariationen der Targets beim Geometrieaustausch zu vermuten. In einer Lungman-N1-Phantomstudie berichten Andersen et al., dass die modellbasierte IR-Methode das CNR mehr verbessert als andere Verfahren (Andersen et al., 2018). Als Hauptursache der Optimierung wurden die rauschreduzierenden Eigenschaften modellbasierter IR-Algorithmen aufgeführt. Diese Ergebnisse bestätigen sich in der vorliegenden Dissertation nur für das CNR mit dem Br69-Kern. Die hierbei fehlende Qualitätsverbesserung mit den LD-Protokollen in der linken Phantomlunge mit Kugeln und Spikulae lässt sich durch das dort höhere Rauschen sowie eine protokollinhärente geminderte Rauschreduktion erklären. Die Ursache in der rechten Phantomlunge im Kugelsetup bezieht sich auf einen systematischen Messfehler. Letzterer kam in der rechten Phantomlunge für mehrere Prüfer über alle Messungen unabhängig vom CT-Dosisprotokoll vor. Die Fehler konnten einer falsch positionierten Messregion (an benachbarter Modellvaskulaturstruktur) zugeordnet werden. Für den Br32-Kern wurde dies im Spikulaesetup nach der CNR-Berechnung deutlich sichtbar. Die CNR-Ausreißer in der linken Phantomlunge im Kugelsetup bezogen sich auf einzelne systematische Messfehler. Das weitgehend unveränderte SNR und CNR mit dem Br32-Kern könnte darauf beruhen, dass die hohen Rauschunterdrückungseigenschaften des weichen Kerns die des IR-Algorithmus übertreffen. Mit dem B157-Kern senkte sich das CNR signifikant mit zunehmender modellbasierter IR-Stärke und gegenüber der FBP-Rekonstruktion ( $p < 0,05$ ). Die Ursache hierfür war das hohe Rauschen in den Messregionen innerhalb der Phantomlunge.

Die Fallbetrachtung mit den CNR- und SNR-Verteilungen aus den CT-Bildern mit den besten Resultaten der Likert-Skala-Bewertungen erbrachte für beide Phantomlungenhälften pro CT-Dosisprotokoll überlagerte Bereiche für die FBP- und modellbasierte IR-Rekonstruktion (IR-Stärke  $S_3$ ), Kap. 3.2.4. Damit kann eine ähnliche objektive und subjektive Bildqualität der Rekonstruktionsmethoden angenommen werden und die IR-Stärke  $S_3$  als eine mögliche

Alternative zur FBP-Rekonstruktion. Die breiteren Verteilungen in der rechten Phantumlunge stellen die genannten systematischen Messfehler durch Positionsvariation der rechten Messregion dar. Deshalb wurden exemplarisch nur die mittleren SNR- und CNR-Werte der Verteilungen in der linken Phantumlunge gebildet und über die Rekonstruktionen verglichen. Mit dem IR-Stärkeniveau  $S_3$  verbesserte sich das SNR um 31,60 % mit dem SH-, um 23,16 % mit dem LD- und um 16,84 % mit dem ULD-Protokoll gegenüber der FBP-Rekonstruktion. Wogegen sich das CNR um 18,54 % mit dem SH-, um 21,31 % mit dem LD- und um 21,33 % mit dem ULD-Protokoll verschlechterte. Der Hersteller garantiert die Beibehaltung der Bildqualität bei Anwendung der modellbasierten IR-Implementation mit der IR-Stärke  $S_3$  nur bis zu einer Dosisreduktion von 60 % (Ramirez-Giraldo et al., 2017; Siemens Healthineers, 2020). Jedes der verwendeten Niedrigdosis-CT-Protokolle reduzierte die Dosis in einem höheren Maß.

**Bewertung der Strahlendosis.** Für das Scan-Setup mit Kugeln bzw. Spikulae konnte die Strahlendosis mit jedem CT-Dosisprotokoll gegenüber dem diagnostischen Referenzwert für Thorax-CT von 10 mGy (Bundesamt für Strahlenschutz, 2016) höchstsignifikant reduziert werden ( $p < 0,00001$ ), Tab. 3.12, S. 74. Die Dosisreduktion betrug für die CT-Dosisprotokolle gemäß der Reihung SH/ LD/ ULD: 70,1 %/ 73,5 %/ 86,9 % auf eine effektive Dosis von 1,49 mSv/ 1,32 mSv/ 0,65 mSv. Es fand sich eine geringere Minderung für das SH-Protokoll mit Spikulae von 69,1 % auf 1,50 mSv, was vermutlich auf Lagerungsunterschiede des Phantoms nach dem Targetwechsel zurückzuführen ist. Die erreichte effektive Dosis mit dem ULD-Protokoll belief sich auf 0,65 mSv. Miller et al. fanden in einer Studie zur Lungenrundherddetektion mit Ultra-Niedrigdosis-CT und modellbasierter IR-Rekonstruktion eine effektive Dosis in der Größenordnung einer Radiographie von 0,13 mSv (Miller et al., 2019). Dies entspricht für das ULD-Protokoll der vorliegenden Dissertation einer Dosis von fünf Radiographien. Es ist anzunehmen, dass die durchgeführten CT-Scans auch für Röhrenstrom-Zeit-Produkte unter 20 mAs sinnvoll wären. Ein anderer Definitionsvorschlag für Ultra-Niedrigdosis-CT ist eine effektive Dosis von unter 0,5 mSv (Macri et al., 2016). Tsalafoutas et al. kritisierten die Abschätzung der effektiven Dosis über Konversionsfaktoren bei Verwendung von Modulationstechniken. Moderne Dosisüberwachungssysteme bieten hierfür neue und genauere Methoden der Berechnung, welche die Röhrenstrommodulation entlang der Scanrichtung sowie ein inhärentes anthropomorphes Phantom berücksichtigen (Tsalafoutas et al., 2020). Das SH-Protokoll der vorliegenden Promotion wurde mit Modulationstechniken durchgeführt. Folglich könnte die Dosisbewertung mit einer Dosisüberwachungssoftware zweckmäßig sein.

**Limitationen.** Die subjektive Bildqualitätsanalyse mit den Lungman-N1-CT-Bildern wurde mit einer begrenzten Anzahl von Radiologen und für eine einzige Orientierung (axial) und Schichtdicke (5 mm) durchgeführt. Diese Einschränkung war auf den hohen Arbeitsaufwand und die zeitliche Verfügbarkeit von fachkundigen Lesern zurückzuführen. Zudem konnte kein genormter Befundungsmonitor sowie eine Software gestellt werden. Die Befunder verwendeten die eigenen genormten Komponenten. Des Weiteren wurde das CT-Experiment nur an einem CT-Scanner durchgeführt. Der Zugang zu modernen CT-Scannern mit implementierter modellbasierter IR-Rekonstruktion mit Niedrigdosis-CT-Protokollen war zum Zeitpunkt der

Dissertation begrenzt. Es ist davon auszugehen, dass in der Zukunft mehr CT-Scanner mit dem verwendeten Softwarebestand ausgerüstet sind. Die für diese Dissertation verwendete und frei verfügbare 16 Bit-DICOM-Umgebung RadiAnt (Medixant, 2016), Version 4.6.9, bot Einschränkungen bei der Positionierung der Messregionen. Die Prüflingenieurinnen waren gezwungen sich an anatomischen Strukturen in den Bildern zu orientieren. Eine Bildqualitätsanalyse für verschiedene Rekonstruktionsalgorithmen und CT-Dosisprotokolle erforderte mehrere CT-Scans am selben Individuum. Hierfür sind anthropomorphe Phantome zweckmäßig. Verfälschte Ergebnisse durch Atembewegungen wurden somit umgangen. Die Anzahl modellierter Gewebe und deren Morphologie waren begrenzt. Obwohl das verwendete Phantom exakt auf das Lungengewebe kalibriert war, könnten Patientenbilder weitere entscheidende Aspekte offenbaren.

### **QRM-Thorax-Phantom**

Mittlere bis hohe modellbasierte IR-Stärken eignen sich für die Detektionsaufgabe bei einem Lungenkrebs-Screening mit Niedrigdosis-CT am besten. Dazu zeigte sich eine deutlich erhöhte Präferenz für das IR-Stärkeniveau  $S_4$ . Eine Korrelation zwischen der subjektiv wahrgenommenen Eignung der modellbasiert IR-rekonstruierten CT-Bilder mit der radiologischen Erfahrung der Befunder wurde nicht festgestellt.

**Beurteilung der modellbasierten IR-Stärken.** Die beste Eignung zur Lungenrundherddetektion durch die gemittelten Rangsummen pro Rangordnungstest (je drei Testungen) ergab sich mit mittlerem bis hohem modellbasierten IR-Stärkeniveau, Kap. 3.3.1 und Kap. 3.3.2. Für die CT-Bilder der Probe 2 mit klinisch angewandten Parametern (Standard-Lungen-BI57-Kern, variabler Modelldurchmesser) wurde die modellbasierte IR-Stärke  $S_4$  deutlich bevorzugt. Mit den CT-Bildern der Probe 1 für variierende Kerne und einen festen Phantomdurchmesser (QRM-Body mit Ring L) gab es eine Präferenz für das IR-Stärkeniveau  $S_5$ .

Beide Präferenzen stellen kumulative Ergebnisse dar, mit welchen höhere IR-Stärken für ein Lungenkrebs-Screening bevorzugt wurden als bei der subjektiven Bildqualitätsanalyse mit den Lungman-N1-CT-Bildern. Die Likert-Skala-Bewertung erfolgte ohne die modellbasierte IR-Stärke  $S_4$ , gleichwohl wurde der klinisch angewandte BI57-Kern präferiert. Diese Bewertungen stammten von drei Radiologen mit unterschiedlicher Berufserfahrung. Für die Rangordnungstests wurden sechs Radiologen um ihr Urteil gebeten. Dies lässt eine bessere statistische Qualität der Ergebnisse annehmen. Daher folgt die Hypothese, dass das modellbasierte IR-Stärkeniveau  $S_4$  mit dem BI57-Kern in den Lungman-N1-CT-Bildern eine bessere subjektive Erkennbarkeit der Herdmodelle aufweisen könnte als  $S_3$ . Das IR-Niveau  $S_4$  könnte als beste Alternative zur FBP-Rekonstruktion für das Lungenkrebs-Screening mit Niedrigdosis-CT geeignet sein. Mit dem Friedman-Test fanden sich bis auf eine Ausnahme keine Zufallsrangfolgen für Probe 2 mit klinisch angewandten Parametern, Tab. 3.9, S. 70. Indes wurden in einem Drittel aller Fälle Zufallsrangfolgen für Probe 1 detektiert, Tab. 3.6, S. 66. Die kumulative Bewertung der IR-Stärke  $S_5$  aus Probe 1 stand in völligem Widerspruch zu den Ergebnissen der Likert-Skala. Ferner erbrachte der Vergleich der Individualergebnis-

se anhand der mittleren Rangsummen beider Rangordnungstests die größte Differenz der subjektiven Wahrnehmung der CT-Bilder für das IR-Stärkeniveau  $S_5$ , Abb. 3.7, S. 66 und Abb. 3.10, S. 71. Die Ursache für die Zufallsrangfolgen und das widersprüchliche Urteil kann in einem Trainingseffekt der Radiologen bezüglich des klinischen Standard-Bildeindrucks liegen (König, Guberina et al., 2020). Infolgedessen wäre die Bewertung des höchsten IR-Stärkeniveaus, respektive des meistveränderten Bildeindrucks, erwartungsgemäß inhomogen.

Bei dem Rangordnungstest mit Probe 2 variierte der Phantomdurchmesser. In Abhängigkeit vom Body-Mass-Index (BMI) wurde in (de Lorenzo, 2011, S. 58) empfohlen, das Röhrenstrom-Zeit-Produkt/Schicht von 10 mAs bei der Niedrigdosis-CT der Lunge für übergewichtige Patienten ( $BMI \geq 25 \text{ kg/m}^2$ ) anzuheben. Dadurch sollte eine ausreichende Bildqualität zur Diagnostik kleiner anatomischer Details gewährleistet werden. Euler et al. berichteten, dass bei der Anwendung der automatischen Röhrenstrommodulation gegenüber einem festen Röhrenstrom-Zeit-Produkt an übergewichtigen Patienten eine deutlich höhere Strahlendosis appliziert werden könnte (Euler et al., 2014). Demnach wäre die Wahl eines optimalen Röhrenstrom-Zeit-Produktes sinnvoll. Viry et al. untersuchten in einer aufgabenbasierten Phantomstudie den Einfluss von IR-Algorithmen auf die Niedrigkontrasterkennbarkeit in Abhängigkeit des effektiven Abdomeindurchmessers (Viry et al., 2018). Dabei kam ein QRM-Abdomen-Phantom (QRM GmbH, Möhrendorf, Deutschland) mit artifizieller Leber und Milz sowie Extensionsringen (Gewebe- und Fettäquivalent) zur Anwendung. Das Patientengewicht wurde mit dem effektiven Phantomdurchmesser anhand von 500 Patientendaten korreliert. Eine effektive Modellbreite von 350 mm (adaptiertes Fettäquivalent) entsprach hierbei einem Patientengewicht von ca. 100 kg. Das QRM-Thorax-Phantom dieser Dissertation hatte eine nahezu identische Kontur bei variantem Materialeinsatz und eine effektive Modellbreite von 400 mm bei sonst identischen effektiven Größenordnungen (adaptierter Ring L als Fettäquivalent). Die Spirale eines Thorax-CT beginnt unterhalb der Lungenspitzen und wird ggf. bei der Frage nach einem Tumor in Richtung der Leber aufgebildet (Möller und Reif, 2019, S. 272 ff.). Diese überlappende Lagebeziehung bei einem CT-Scan von Oberbauch und Lunge sowie die ähnlichen Phantomdimensionen des QRM-Abdomen- und des QRM-Thorax-Phantoms lässt eine ähnliche Simulation des Patientengewichts (bei adaptiertem Ring L) annehmen. Mitunter würde mit beiden Rangordnungstests ein übergewichtiger Patient simuliert. Analysen für einzelne Modelldurchmesser des Rangordnungstests mit Probe 2, Anhang A.7, konnten für einen adipösen Patienten keine abweichende subjektive Eignung der CT-Bilder zur Lungenrundherddetektion mit modellbasierter IR-Rekonstruktion feststellen. Die Vorbehalte gegenüber niedrigen Röhrenstrom-Zeit-Produkten bei übergewichtigen Patienten hinsichtlich der anatomischen Bildqualität ließen sich mit der vorliegenden Dissertation nicht bestätigen. Die modellbasierte IR-Rekonstruktion könnte möglicherweise ein Lungenkrebs-Screening mit Niedrigdosis-CT-Protokollen von übergewichtigen Patienten optimieren.

**Korrelation der subjektiven Ergebnisse mit der radiologischen Erfahrung.** Mit beiden Rangordnungstests konnte keine Korrelation zwischen der subjektiven Wahrnehmung der CT-Bilder und der radiologischen Erfahrung festgestellt werden. Bei dem Rangordnungstest mit Probe 1 urteilten der erfahrenste und der am wenigsten erfahrene Beurteiler

gleich und folgten dem kumulativen Ergebnis. Ein weiterer erfahrener Radiologe bewertete widersprüchlich. Ähnliche Gegensätze fanden sich für Probe 2. Die Tests wurden auf einer Vierfeld-Benutzerschnittstelle durchgeführt. Diese Schnittstelle erlaubte keine klinisch angewandten Bildgrößenverhältnisse. Die Radiologen bewerteten kleinere Formate als im Routinebetrieb. Zudem wurde kein festes Kriterium vorgegeben, an dem die Eignung der CT-Bilder zur Lungenrundherddetektion zu beurteilen war. Die Radiologen benannten unterschiedliche Bewertungskriterien. Die nicht festgestellte Korrelation könnte auf die individuelle fachliche Expertise der Radiologen zurückzuführen sein. Der Radiologe mit viel Mammographieerfahrung beurteilte anhand der Wiedergabe von Inhomogenitäten innerhalb der Kugeln und Spikulae, andere legten den Fokus auf die Abbildung der Modellränder. Wirtz et al. berichteten, dass für subjektive Beurteilungssituationen von Qualitätseigenschaften in medizinischen Bereichen bevorzugt auf dichotome kategoriale Urteile oder mehrfach gestufte kategoriale Urteile zurückgegriffen wird (Wirtz und Kutschmann, 2007). Demzufolge wäre die verwendete Rangordnungstestmethode entgegen der Gewöhnung von Radiologen bei der Qualitätsbewertung ausgewählt worden und könnte für die Variation in den individuellen Beurteilungen verantwortlich sein. Dieser Effekt würde zudem die angestrebte Korrelationsanalyse negativ beeinflussen. Mit der Likert-Skala als Bewertungsinstrument ließ sich eine Korrelation zwischen der radiologischen Erfahrung und der subjektiven Wahrnehmung der Radiologen erahnen. Weitere Forschung ist somit indiziert.

**Konkordanzanalysen.** Die Interrater-Reliabilität der Mediziner war mit beiden Rangordnungstests gering und mit Probe 1 (QRM-Body mit Ring L, Variation der Kerne) gering bis schwach, Tab. 3.7, S. 67 und Tab. 3.10, S. 71. Diese Urteilskonkordanz wurde mit dem Kappa-Koeffizienten  $\kappa_{\text{Fleiss}}$  nach (Fleiss, 1971) bestimmt und erwies sich hinsichtlich der Ergebnisse der mittleren Rangsummen beider Rangordnungstests als erwartungsgemäß. Die Intrarater-Reliabilität wurde substantiell an dem modifizierten Koeffizienten  $\kappa_{\text{ney}}$  (Kutschmann, 2020) ermittelt. Dabei zeigte sich für den Rangordnungstest mit Probe 1 keine bis hin zu einer moderaten Übereinstimmung, Abb. 3.8, S. 67. Für Probe 2 (Bl57, Variation der Modelldurchmesser) gab es keine bis hin zu einer nahezu perfekten Übereinstimmung, Abb. 3.11, S. 72. Die Inzidenz von  $\kappa_{\text{ney}}$  erbrachte weitgehend die meisten Koeffizienten beider Rangordnungstests in den Intervallen:  $[0,21; 0,39]$  mit dem Übereinstimmungsgrad „minimal“ und  $[0,40; 0,59]$  mit dem Übereinstimmungsgrad „schwach“. Der zur Bewertung herangezogene Interpretationsvorschlag wurde speziell für medizinische Forschungsergebnisse empfohlen und setzte kategorisch höhere  $\kappa$ -Koeffizienten voraus als in (Landis und Koch, 1977). Eine Inzidenzbetrachtung der maximalen und minimalen  $\kappa_{\text{ney}}$  pro Rang über alle Radiologen und Testungen identifizierte für konsistente Ergebnisse die beste und schlechteste Urteilskonkordanz. Der Rangordnungstest mit Probe 2 ergab die beste Urteilskonkordanz mit Rang 1 und die schlechteste mit Rang 3. Mit Probe 1 lagen für die Identifikation der besten Urteilskonkordanz keine konsistenten Ergebnisse vor. Die schlechteste Urteilskonkordanz fand sich bei Rang 2. Die Übersicht aller Ergebnisse deutete auf eine geringere Intrarater-Reliabilität für die Positionen Rang 2 und Rang 3 gegenüber Rang 1 und Rang 4 hin, Tab. 3.8, S. 68 und Tab. 3.11, S. 73. Dies suggeriert, dass die Entscheidung für das schlechteste und beste Prüfbild leichter fiel. Hier ist die Ursache in der ungewohnten Testmethode zu vermuten (König, Guberina et al., 2020; Wirtz und Kutschmann, 2007).



**Wahl der Intrarater-Reliabilitäts-Testmethode.** Eine sekundärer Ansatz bestand darin, den Kappa-Koeffizienten  $\kappa$  nach (Cohen, 1960) zur Untersuchung der Intrarater-Reliabilität zu verwenden, Kap. 2.2.5. Es fand sich für beide Rangordnungstests eine Vielzahl negativer Koeffizienten, welche nicht zur Bewertung herangezogen werden konnten. Außerdem kann der  $\kappa$ -Koeffizient paradoxe und kontraintuitive Ergebnisse liefern die seine Interpretation erschweren (Grouven et al., 2007). Um dem hohen Anspruch an medizinische Forschungsdaten gerecht zu werden, wurde die  $\kappa$ -Modifikation  $\kappa_{\text{ney}}$  zur Konkordanzanalyse eingesetzt. Die Berechnung von Cohen's  $\kappa$  erfolgt üblicherweise aus dem Anteil der übereinstimmenden Beurteilungen  $B$  und den zufällig zu erwartenden konkordanten Bewertungen  $E$  gemäß der Formel  $\kappa = (B - E)/(1 - E)$ . Anstatt  $E$  verwendet der  $\kappa_{\text{ney}}$ -Koeffizient die Modifikation  $C$ . Mit der Modifikation werden nur die Anteile der Randverteilungen der Cohen'schen Kontingenztabelle berücksichtigt, die auch einen Anteil an  $B$  haben (Kutschmann, 2020; König, Guberina et al., 2020). Hauptvorteile dieser Methode bestehen mitunter darin, dass keine  $\kappa$ -Paradoxa und negativen Koeffizienten auftreten, welche die Interpretation verkomplizieren.

**Limitationen.** Für das Experiment mit dem QRM-Phantom galten hinsichtlich der Analysen und des CT-Experiments die Einschränkungen des Lungman-N1-Phantoms, S. 81, gleichermaßen. Die Befundungsbedingungen waren aus folgenden Gründen nicht standardisiert. Erstens, der für die Rangordnungstests verwendete Monitor sowie die Lichtverhältnisse in den Befundungsräumen waren nicht genormt. Zweitens, das zur Untersuchung benötigte kommerzielle, Programm MATLAB (The MathWorks, 2018), Version R2018a, durfte wegen Datenschutzregeln nicht auf Rechnern mit angeschlossenen Befundungsmonitoren installiert werden. Drittens, eine Zoomfunktion zur Bildvergrößerung stand nicht zur Verfügung. Für den experimentellen Aufbau des QRM-Phantoms wurden die Targets über gepresstes Granulat in Lage und Ausrichtung fixiert. Für die Dauer des Experiments ist die Pressung nicht aufgehoben worden und alle Ringadaptionen wurden bei gleicher Targetlage vorgenommen. Dennoch stellten eine weitreichende fotografische Dokumentation, Höhenmarkierungen im Phantom und maßstabsgetreue Konstruktionszeichnungen die Reproduzierbarkeit des experimentellen Aufbaus sicher.

## Ausblick

Die vorliegende Dissertation stellte die Eignung der Niedrigdosis-Computertomographie mit modellbasierter IR-Rekonstruktion für das Lungenkrebs-Screening fest. Mit der Betrachtung objektiver und subjektiver Bildqualitätsaspekte, verschiedener Kerne sowie IR-Stärken erwies sich das modellbasierte IR-Stärkeniveau  $S_4$  potentiell am besten für die Rekonstruktion der CT-Screening-Bilder geeignet. Dies ließ sich auch auf computergestützte DL-CAD-Ansätze zur Lungenrundherddetektion mit weichen Kernen und modellbasierter IR-Rekonstruktion übertragen und diskutieren (Blazis et al., 2021). Offen blieb jedoch die Frage nach der Korrelation der subjektiven Wahrnehmung der Lungen-CT-Bilder und der radiologischen Erfahrung der Befunder.

Die in (Prosch und Ebner, 2019) chronologisch vorgestellten US-amerikanischen und europäi-

schen Lungenkrebs-Screening-Studien verdeutlichten, dass sich die Idee einer dosissparenden Screeningprozedur bereits seit zwei Jahrzehnten bei Radiologen manifestiert hat. Dem wird die modellbasierte IR-Rekonstruktion gerecht, und zwar in den in dieser Dissertation verwendeten Niedrigdosis-CT-Protokollen mit einer höchstsignifikanten Strahlendosisreduktion ( $p < 0,00001$ ) gegenüber dem diagnostischen Referenzwert für Thorax-CT (Bundesamt für Strahlenschutz, 2016). Beachtlich ist, dass die ersten IR-Verfahren als Niedrigdosis-CT-Ansatz nach dem Jahr 2010 in die klinische Routine eingeführt wurden. Seitdem haben sich Hersteller mit einer Vielzahl von IR-Algorithmen dem FBP-Problem des Bildqualitätsverlusts bei reduzierter Strahlendosis gestellt (Willemink und Noël, 2019).

Doch auch die Vorzüge moderner modellbasierter IR-Methoden (z. B. Strahlenschutzaspekte, optimierte Bildqualitätsparameter, Eignung zur Befundung durch DL-CAD-Systeme) könnten künftig durch den Einsatz neuer Rekonstruktionsverfahren mit integrierter künstlicher Intelligenz (KI) übertroffen werden. Der schnelle Entwicklungsfortschritt und das enorme Interesse an der KI auf dem Gebiet der radiologischen Bildgebung führte zu kontroversen Diskussionen verschiedener Fachdisziplinen bezogen auf das Potential und das Verständnis dieser Methoden (Barrett, 2019; Herold und Delorme, 2020).

Parallel wird an einem vielversprechenden Rekonstruktionsverfahren geforscht: Compressed Sensing (CS, deutsch: komprimiertes Abtasten) ist ein neuer Ansatz der Datenakquisition und -rekonstruktion, welcher mit unterabgetasteten Messdaten Bilder erzeugt. Für die Magnetresonanztomographie wurde diese Methode bereits in die klinische Routine implementiert. Hierbei lag der Hauptvorteil des CS-Verfahrens darin, den  $k$ -Raum (Rohdatenraum) nicht vollständig mit Messdaten füllen zu müssen und somit kürzere Scanzeiten zu erzielen (Geethanath et al., 2013). Die Simulationsstudie von Hou et al. belegte die Fähigkeit der CS-Algorithmen, als dosissparende CT-Rekonstruktion eine gute Bildqualität zu erreichen (Hou und Zhang, 2014). Zhang et al. verwendeten zur Simulation CS-inhärente mathematische Optimierungsansätze (z. B. rauschreduzierende Regularisierungen und Gewichtungsterme, die Lagrange-Methode) und einen niedrigen Strahlendosisparameter (Zhang et al., 2019). Die Ergebnisse waren äußerst hochwertige CT-Bilder bei optimierten Qualitätsparametern. In (Hayes et al., 2021) wurde über das CS-Potential für einen artefaktfreien und schnellen Spiral-CT-Scan berichtet. Bei gleichzeitiger Nutzung moderner CS- und DL-Techniken mit der FBP-Rekonstruktion ist für ein Mehrzeilen-CT-System mit einer Quellen-Detektor-Kombination ein Pitchfaktor von bis zu 4 erreichbar.

KI ist jedoch noch nicht in der Lage, die Komplexität eines Lungen-CT so zu behandeln, dass es zu einer sicheren Diagnose kommt. Doch die engumrissene Fragestellung bei der Lungenkrebs-Screeningprozedur hinsichtlich Detektion und Quantifizierung bildet ein weites Feld der gegenwärtigen Forschungsaktivitäten (Herold und Delorme, 2020). Hinzu kommen prädikative KI-Ansätze, welche auf Biomarkern aufbauen, die mitunter aus medizinischen Bilddaten stammen. Bereits die Rohdaten liefern Informationen (z. B. Textur und Homogenität), welche über mathematische Methoden objektivierbar sind. Ein Ansatz, der diese Zusammenhänge analysiert und klinische Targets prognostiziert, ist Radiomics (Prayer et al., 2020; Herold und Delorme, 2020). In (Ye et al., 2021) wurde über einen signifikanten Einfluss statistischer

IR-Stärkeniveaus mit Ultra-Niedrigdosis-CT auf die Quantifizierung von Lungenrundherden mit Radiomics berichtet. Eine entsprechende Untersuchung mit modellbasierten IR-Stärken ist im Hinblick auf ein dosissparendes Lungenkrebs-Screening empfehlenswert.

In der Zukunft könnte das CS-Rekonstruktionsverfahren in Verbindung mit DL- oder ML-Methoden gegenüber der modellbasierten IR-Methode als schnellere und dosissparende Technik für das Lungenkrebs-Screening eine exzellente Bildqualität bereitstellen. Es ist zu erwarten, dass sich für Radiologen und CAD-DL-Systeme sowie prädikative KI-Ansätze eine optimierte Ausgangsdatenbasis bietet, an der die Diagnose- bzw. Systemleistung zu messen sein wird.

## 5 Zusammenfassung

Diese Dissertation untersuchte die Eignung der Niedrigdosis-Computertomographie (LDCT) mit modellbasierter iterativer Rekonstruktion (MBIR) für das Lungenkrebs-Screening. Mit dem Scanner Somatom Force wurden Spiral-CT-Scans von zwei anthropomorphen Lungenphantomen akquiriert und einer umfangreichen Bildqualitätsanalyse unterzogen. Kalibrierte artifizielle Lungenrundherde (Hounsfield-Werte:  $-690/ -50/ 100$  HU) in Form von Kugeln und Spikulae (Durchmesser:  $12/ 10/ 8/ 5/ 3$  mm bzw.  $16/ 20/ 24$  mm) wurden auf markierte Positionen in den Phantomen fixiert. Jedes Phantom wurde mit einem Standard-Hochkontrast- und zwei LDCT-Protokollen gescannt: Low-Dose (LD), Ultra-Low-Dose (ULD). Zur Rekonstruktion wurde die gefilterte Rückprojektion (FBP), die MBIR (Stärke  $S_1$  bis  $S_5$ ) sowie der Standard-Lungen-BI57-, der weiche Br32- und der harte Br69-Kern verwendet.

Drei Radiologen mit unterschiedlicher Befundungserfahrung bewerteten die Erkennbarkeit der Kugeln und Spikulae mit einer sechsstufigen Likert-Skala. Erfahrene Radiologen bewerteten die Erkennbarkeit in den FBP-Bildern höher als mit MBIR und am höchsten mit BI57. Der weniger erfahrene Radiologe beurteilte FBP und MBIR gleichermaßen mit Höchstbewertungen. Es wurde das Signal- (SNR) bzw. Kontrast-Rausch-Verhältnis (CNR) gemessen. Unabhängig vom LDCT-Protokoll stieg mit der MBIR-Stärke das SNR für BI57 und Br69 sowie das CNR für Br69, wogegen SNR und CNR für Br32 weitgehend unverändert blieben (jeweils  $p < 0,05$ ).

Die geschätzte effektive Dosis, das Dosis-Längen-Produkt und der Volumen-CT-Dosisindex ( $CTDI_{vol}$ ) wurden aufgenommen. Die  $CTDI_{vol}$ -Werte wurden mit dem diagnostischen Referenzwert (DRW) für Thorax-CT, welcher  $10$  mGy beträgt, verglichen. Gegenüber dem DRW fiel der  $CTDI_{vol}$  für Kugeln als auch für Spikulae um  $73,5\%$  auf  $1,32$  mSv mit LD und um  $86,9\%$  auf  $0,65$  mSv mit ULD ( $p < 0,00001$ ).

Mit zwei Rangordnungstests beurteilten sechs Radiologen die Eignung der CT-Bilder für die Screeningprozedur. Die Auswertung wurde mit der sog. Friedman-Statistik vorgenommen. Die Radiologen bevorzugten mittlere bis hohe MBIR-Stärken. Für die CT-Bilder mit BI57 und variablem Phantomdurchmesser gab es eine deutliche Präferenz für  $S_4$  und keine Zufallsrangfolgen. Hingegen wurden diese in den Rangordnungen mit variierenden Kernen detektiert. Um die Interrater- und Intrarater-Reliabilität zu quantifizieren, wurden der Koeffizient  $\kappa_{\text{Fleiss}}$  sowie die Cohen's  $\kappa$ -Modifikation  $\kappa_{\text{ney}}$  verwendet. Die Interrater-Reliabilität war gering bis schwach. Die Intrarater-Reliabilität lag überwiegend zwischen schwach und minimal.

Die Ergebnisse legen nahe, dass die MBIR eine geeignete Alternative zur FBP ist und ein Lungenkrebs-Screening bei deutlich geringerer Strahlendosis und gleicher oder besserer Bildqualität ermöglicht. Besonders geeignet erwies sich die Stärke  $S_4$ . Die Korrelation zwischen subjektiver Wahrnehmung der CT-Bilder und radiologischer Erfahrung wurde nicht belegt.

## 6 Literaturverzeichnis

1. Afadzi, M., Fosså, K., Andersen, H.K., Aaløkken, T.M., Martinsen, A.C.T. (2020): Image quality measured from ultra-low dose chest computed tomography examination protocols using 6 different iterative reconstructions from 4 vendors, a phantom study. *J. Comput. Assist. Tomogr.* 44, 95–101.
2. Ali, N., Lifford, K.J., Carter, B., McRonald, F., Yadegarfar, G., Baldwin, D.R., Weller, D., Hansell, D.M., Duffy, S.W., Field, J.K., Brain, K. (2015): Barriers to uptake among high-risk individuals declining participation in lung cancer screening: A mixed methods analysis of the UK lung cancer screening (UKLS) trial. *Br. Med. J. Open.* 5, 1–9.
3. American Association of Physicists in Medicine. (2008): Report 096. The measurement, reporting, and management of radiation dose in CT. 2. Aufl. Berlin: Georg Thieme Verlag.
4. Andersen, H.K., Völgyes, D., Martinsen, A.C.T. (2018): Image quality with iterative reconstruction techniques in CT of the lungs—A phantom study. *Eur. J. Radiol. Open.* 5, 35–40.
5. Barrett, H.H. (2019): Is there a role for image science in the brave new world of artificial intelligence? *J. Med. Imag.* 7, 1–6.
6. Barrett, J.F., Keat, N. (2004): Artifacts in CT: Recognition and avoidance. *RadioGraphics.* 24, 1679–1691.
7. Becker, N., Motsch, E., Trotter, A., Heussel, C.P., Dienemann, H., Schnabel, P.A., Kauczor, H.-U., Maldonado, S.G., Miller, A.B., Kaaks, R., Delorme, S. (2020): Lung cancer mortality reduction by LDCT screening—results from the randomized german LUSI trial. *Int. J. Cancer.* 146, 1503–1513.
8. Beister, M., Kolditz, D., Kalender, W.A., (2012): Iterative reconstruction methods in x-ray CT. *Phys. Med.* 28, 94–108.
9. Blazis, S.P., Dickerscheid, D.B.M., Linsen, P.V.M., Martins Jarnalo, C.O. (2021): Effect of CT reconstruction settings on the performance of a deep learning based lung nodule CAD system. *Eur. J. Radiol.* 136, 1–6.
10. Bortz, J., Lienert, G.A. (2008): *Kurzgefasste Statistik für die klinische Forschung: Leitfaden für die verteilungsfreie Analyse kleiner Stichproben; mit 97 Tabellen sowie zahlreichen Formeln.* 3. Aufl. Heidelberg: Springer Medizin Verlag.
11. Bundesamt für Strahlenschutz. (2016): *Bekanntmachung der aktualisierten diagnostischen Referenzwerte für diagnostische und interventionelle Röntgenanwendungen.*
12. Bundesamt für Strahlenschutz. (2007): *Die Empfehlungen der Internationalen Strahlenschutzkommission (ICRP) von 2007: ICRP-Veröffentlichung 103.*

13. Bundesamt für Strahlenschutz. (2019): Früherkennung mittels CT: BfS kritisiert aktuelle Angebote. Abgerufen am 07.04.2021 von <https://www.bfs.de/SharedDocs/Pressemitteilungen/BfS/DE/2019/016.html>.
14. Buzug, T.M. (2014): Einführung in Die Computertomographie. Mathematisch–physikalische Grundlagen der Bildrekonstruktion. Berlin: Springer Verlag.
15. Center for Statistical Sciences. (2011): The national lung screening trial: Overview and study design. *Radiol.* 258, 243–253.
16. Chamberlin, J., Kocher, M.R., Waltz, J., Snoddy, M., Stringer, N.F.C., Stephenson, J., Sahbaee, P., Sharma, P., Rapaka, S., Schoepf, U.J., Abadia, A.F., Sperl, J., Hoelzer, P., Mercer, M., Somayaji, N., Aquino, G., Burt, J.R. (2021): Automated detection of lung nodules and coronary artery calcium using artificial intelligence on low–dose CT scans for lung cancer screening: accuracy and prognostic value. *BMC. Med.* 19, 1–14.
17. Cohen, J. (1960): A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46.
18. de Lorenzo, C.T. (2011): Low–Dose–Multidetektor–Computertomographie des Thorax: Qualität der Aufnahmen in Abhängigkeit von der Körpermasse. Dissertation. LMU München: Medizinische Fakultät.
19. De Man, B., Basu, S. (2003): Distance–driven projection and backprojection. *IEEE. Nucl. Sci. Symp. Conf. Rec.* 3, 1477–1480.
20. Deserno, T.M., Oberschelp, W., Pelikan, E., Repges, R. (1997): Bildverarbeitung für die Medizin: Grundlagen, Modelle, Methoden, Anwendungen. Berlin, Heidelberg, New York (usw.): Springer Verlag.
21. Deutsches Institut für Normung. (2014): DIN 6868–157:2014–11. Sicherung Der Bildqualität in röntgendiagnostischen Betrieben–Teil 157: Abnahme- und Konstanzprüfung nach RöV an Bildwiedergabesystemen in ihrer Umgebung.
22. Deutsches Institut für Normung. (2017): DIN EN 60601–2–44:2017–3. Medizinische elektrische Geräte–Teil 2–44: Besondere Festlegungen für die Sicherheit einschließlich der wesentlichen Leistungsmerkmale von Röntgeneinrichtungen für die Computertomographie (IEC 60601–2–44:2009; EN 60601–2–44:2009).
23. Doo, K.W., Kang, E.–Y., Yong, H.S., Woo, O.H., Lee, K.Y., Oh, Y.–W. (2014): Accuracy of lung nodule volumetry in low–dose CT with iterative reconstruction: An anthropomorphic thoracic phantom study. *Br. J. Radiol.* 87, 1–10.
24. Dössel, O., Buzug, T.M. (2014): Biomedizinische Technik–Medizinische Bildgebung. Berlin, Boston: De Gruyter.
25. Dremel, K. (2017): Modellbildung des Messprozesses und Umsetzung eines modellbasierten iterativen Lösungsverfahrens der Schnittbild–Rekonstruktion für die Röntgen–Computertomographie. Dissertation. JMU Würzburg: Fakultät für Physik und Astronomie.
26. Ese, Z., Qamhiyeh, S., Kreutner, J., Schaefer, G., Erni, D., Zylka, W. (2019): CT–extended hounsfield unit range in radiotherapy treatment planning for patients with implantable medical devices. *IFMBE Proceedings* 68, 599–603.

27. Euler, A., Szücs–Farkas, Z., Schindera, S. (2014): Möglichkeiten der Strahlenreduktion bei der CT des Körperstamms. *Radiol. up2date*. 14, 163–176.
28. Europäische Union. (2000): EUR 16262 EN:2000–2. European guidelines on quality criteria for computed tomography.
29. Fahrmeir, L., Künstler, R., Pigeot, I., Tutz, G. (2004): Statistik: Der Weg zur Datenanalyse. 5. Aufl. Berlin: Springer Verlag.
30. Food and Drug Administration. (2011): 21 CFR 1020:2011–04. Performance standards for ionizing radiation emitting products–section 1020.30: Diagnostic x–ray systems and their major components.
31. Feinstein, A.R., Cicchetti, D.V. (1990): High agreement but low kappa: I. The problems of two paradoxes. *J. Clin. Epidemiol.* 43, 543–549.
32. Fleiss, J.L. (1971): Measuring nominal scale agreement among many raters. *Psychol. Bull.* 5, 378–382.
33. Flohr, T., Stierstorfer, K., Raupach, R., Ulzheimer, S., Bruder, H. (2004): Performance evaluation of a 64–slice CT system with z–flying focal spot. *Rofo*. 176, 1803–1810.
34. Fu, B., Wang, G., Wu, M., Li, W., Zheng, Y., Chu, Lv, F. (2020): Influence of CT effective dose and convolution kernel on the detection of pulmonary nodules in different artificial intelligence software systems: A phantom study. *Eur. J. Radiol.* 126
35. Geethanath, S., Reddy, R., Konar, A.S., Imam, S., Sundaresan, R., Babu, R., Venkatesan, R. (2013): Compressed sensing MRI: a review. *Crit. Rev. Biomed. Eng.* 41, 183–204.
36. Grouven, U., Bender, R., Ziegler, A., Lange, S. (2007): Der Kappa–Koeffizient. *Dtsch. med. Wochenschr.* 132, e65–e68.
37. Guggenmoos–Holzmann, I. (1993): How reliable are chance–corrected measures of agreement? *Stat. Med.* 12, 2191–2205.
38. Han, M., Baek, J. (2020): A convolutional neural network–based anthropomorphic model observer for signal–known–statistically and background–known–statistically detection tasks. *Phys. Med. Biol.* 65.
39. Hartung, J., Elpelt, B., Klösener, K.–H. (2009): Statistik: Lehr– und Handbuch der angewandten Statistik. 15. Aufl. München: Oldenbourg Verlag.
40. Hassani, C., Ronco, A., Prosper, A.E., Dissanayake, S., Cen, S.Y., Lee, C. (2018): Forward–projected model–based iterative reconstruction in screening low–dose chest CT: Comparison with adaptive iterative dose reduction 3D. *Am. J. Roentgenol.* 211, 548–556.
41. Hayes, J.W., Montoya, J., Budde, A., Zhang, C., Li, Y., Lia, K., Hsieh, J., Chen, G.H. (2021): High pitch helical CT reconstruction. *IEEE. Trans. Med. Imag.* 1, 1.
42. Herold, C.J., Delorme, S. (2020): Brave new world. *Radiologe.* 60, 1–5.
43. Hofer, M. (2000): CT–Kursbuch: ein Arbeitsbuch für den Einstieg in die Computertomographie. 3. Aufl. Düsseldorf: Didamed Verlag.
44. Hou, W., Zhang, C. (2014): Analysis of compressed sensing based CT reconstruction with low radiation. *IEEE. Int. Symp. Intell. Signal Process. Commun. Syst.* 1, 291–296.

45. Hsieh, J., Lui, E., Nett, B., Tang, J., Thibault., J.–B., Sahney, S. (2019): A new era of image reconstruction: TrueFidelity. Technical white paper on deep learning image reconstruction. *GE Healthcare*. 1, 1–15.
46. Huesman, R.H., Gul, G.T., Budinger, T.F. (1977): Lawrence Berkeley National Laboratory (LBNL)–Report. Users manual: Donner algorithms for reconstruction tomography.
47. Hunger, T., Wanka–Pail, E., Brix, G., Griebel, J. (2021): Lung cancer screening with low–dose CT in smokers: A systematic review and meta–analysis. *Diagn.* 11, 1–15.
48. Infante, M., Cavuto, S., Lutman, F.R., Brambilla, G., Chiesa, G., Ceresoli, G., Passera, E., Angeli, E., Chiarenza, M., Aranzulla, G., Cariboni, U., Errico, V., Inzirillo, F., Bottoni, E., Voulaz, E., Alloisio, M., Destro, A., Roncalli, M., Santoro, A. (2009): A randomized study of lung cancer screening with spiral computed tomography: Three–year results from the DANTE trial. *Am. J. Respir. Crit. Care. Med.* 180, 445–453.
49. Infante, M., Lutman, F.R., Cavuto, S., Brambilla, G., Chiesa, G., Passera, E., Angeli, E., Chiarenza, M., Aranzulla, G., Cariboni, U., Alloisio, M., Incarbone, M., Testori, A., Destro, A., Cappuzzo, F., Roncalli, M., Santoro, A., Ravasi, G. (2008): Lung cancer screening with spiral CT: Baseline results of the randomized DANTE trial. *Lung Cancer*. 59, 355–363.
50. Infante, M., Sestini, S., Galeone, C., Marchianò, A., Lutman, F.R., Angeli, E., Calareso, G., Pelosi, G., Sozzi, G., Silva, M., Sverzellati, N., Cavuto, S., La Vecchia, C., Santoro, A., Alloisio, M., Pastorino, U. (2017): Lung cancer screening with low–dose spiral computed tomography: Evidence from a pooled analysis of two italian randomized trials. *Eur. J. Cancer. Prev.* 26, 324–329.
51. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. (2020): Bericht Nr. 982. Lungenkrebscreening mittels Niedrigdosis–Computertomografie.
52. International Association for the Study of Lung Cancer. (2018): Press release. NELSON study shows CT screening for nodule volume management reduces lung cancer mortality by 26 percent in men.
53. International Business Machines (IBM) Corporation, New York, Vereinigte Staaten von Amerika (2020).
54. Jayamani, J., Osman, N.D., Tajuddin, A.A., Salehi, Z., Ali, M.H., Aziz, M.Z.A. (2019): Determination of computed tomography number of high–density materials in 12–bit, 12–bit extended and 16–bit depth for dosimetric calculation in treatment planning system. *J. Radiother. Pract.* 18, 285–294.
55. Kauczor, H.–U., Baird, A.–M., Blum, T.G., Bonomo, L., Bostantzoglou, C., Burghuber, O., Cepická, B., Comanescu, A., Couraud, S., Devaraj, A., Jespersen, V., Morozov, S., Nardi Agmon, I., Peled, N., Powell, P., Prosch, H., Ravara, S., Rawlinson, J., Revel, M.–P., Silva, M., Snoeckx, A., Ginneken van, B., Meerbeeck, J.P., Vardavas, C., Stackelberg von, O., Gaga, M. (2020): ESR/ERS statement paper on lung cancer screening. *Eur. Respir. J.* 55, 1–18.



56. Kim, J.H., Yoon, H.J., Lee, E., Kim, I., Cha, Y.K., Bak, S.H. (2021): Validation of Deep-Learning Image Reconstruction for Low-Dose Chest Computed Tomography Scan: Emphasis on Image Quality and Noise. *Korean J. Radiol.* 22, 131–138.
57. König, B., Guberina, N., Köhl, H., Zylka, W. (2020): Eignet sich die statistische modellbasierte iterative Rekonstruktion mit low-dose Protokollen der Computertomographie für Lungenkrebs-Screening? Erste Ergebnisse einer Phantomstudie. *Rofo.* 192, 110.
58. König, B., Guberina, N., Köhl, H., Zylka, W. (2019): Design and first results of a phantom study on the suitability of iterative reconstruction for lung-cancer screening with low-dose computer tomography. *Curr. Dir. Biomed. Eng.* 5, 593–596.
59. König, B., Guberina, N., Köhl, H., Zylka, W. (2020): Validation of iterative CT reconstruction by inter and intra observer performance assessment of artificial lung foci. *Curr. Dir. Biomed. Eng.* 6, 534–537.
60. König, B., Köhl, H., Zylka, W. (2021): CT lung-cancer screening: does iterative reconstruction matter? *Shanghai Chest.*
61. Koning, H.D., Aalst, C.V.D., Haaf, K.T., Oudkerk, M. (2018): Effects of volume CT lung cancer screening: Mortality results of the NELSON randomised-controlled population based trial. *J. Thorac. Oncol.* 13, 185.
62. Köhl, H. (2018): Private Kommunikation.
63. Kurfürst, M.C.R. (2018): Einfluss verschiedener iterativer Rekonstruktionstechniken auf die Bildqualität und Detailerkennbarkeit in der CT Pulmonalisangiographie mit einem Niedrigdosisprotokoll und reduziertem Kontrastmittelvolumen im Vergleich zu der herkömmlichen gefilterten Rückprojektion. Dissertation. UH Hamburg: Medizinische Fakultät.
64. Kutschmann, M. (2020): Private Kommunikation.
65. Kwicien, R., Kopp-Schneider, A., Blettner, M. (2011): Concordance analysis. *Dtsch. Arztebl.* 108, 515–521.
66. Landis, J.R., Koch, G.G. (1977): The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.
67. Lee, K.B., Goo, H.W. (2018): Quantitative image quality and histogram-based evaluations of an iterative reconstruction algorithm at low-to-ultralow radiation dose levels: A phantom study in chest CT. *Korean J. Radiol.* 19, 119–129.
68. Leitz, W., Axelsson, B., Szendrö, G. (1995): Computed tomography dose assessment—A practical approach. *Radiat. Prot. Dosim.* 57, 377–380.
69. Li, J., Udayasankar, U.K., Toth, T.L., Seamans, J., Small, W.C., Kalra, M.K. (2007): Automatic patient centering for MDCT: Effect on radiation dose. *Am. J. Roentgenol.* 188, 547–552.
70. Macri, F., Greffier, J., Pereira, F., Mandoul, C., Khasanova, E., Gualdi, G., Beregi, J. (2016): Ultra-low-dose chest CT with iterative reconstruction does not alter anatomical image quality. *Diagn. Interv. Imag.* 97, 1131–1140.
71. Mascalchi, M., Sali, L. (2017): Lung cancer screening with low dose CT and radiation harm—from prediction models to cancer incidence data. *Ann. Transl. Med.* 5, 1–6.

72. McHugh, M.L. (2012): Interrater reliability: The kappa statistic. *Biochem. Med.* 22, 276–282.
73. Medixant, Poznań, Polen (2016).
74. Microsoft Corporation, Redmond, Vereinigte Staaten von Amerika (2016).
75. Miller, A.R., Jackson, D., Hui, C., Deshpande, S., Kuo, E., Hamilton, G.S., Lau, K.K. (2019): Lung nodules are reliably detectable on ultra-low-dose CT utilising model-based iterative reconstruction with radiation equivalent to plain radiography. *Clin. Radiol.* 74, e17–e22.
76. Möller, T.B., Reif, E. (2019): Taschenatlas Einstelltechnik: Röntgendiagnostik, Angiografie, CT, MRT. 6. Aufl. Stuttgart: Georg Thieme Verlag.
77. Nagel, H.D., Galanski, M., Hidajat, N., Maier, W., Schmidt, Th. (2002): Dosimetrische Grundlagen, Dosiswerte bei CT-Untersuchungen, Dosisbeeinflussende Faktoren, Niedrigdosis-CT im klinischen Alltag. In: Nagel, H.D. (Hrsg.): *Strahlenexposition in der Computertomographie. Grundlagen, Einflussfaktoren, Dosisermittlung, Optimierung, Zahlenwerte, Begriffe*. 3. Aufl. S. 5–53. Hamburg: CTB Publications.
78. Nanavaty, P., Alvarez, M.S., Alberts, W.M. (2014): Lung cancer screening: advantages, controversies, and applications. *Cancer Control.* 21, 9–14.
79. National Institutes of Health, Bethesda, Vereinigte Staaten von Amerika (2019).
80. Nguyen, H.T. (2019): Implementierung einer statistischen Analyse zur Bewertung der subjektiven Qualität von iterativ rekonstruierten CT Bildern. Masterthesis. Westfälische Hochschule Gelsenkirchen Bocholt Recklinghausen: Fachbereich Elektrotechnik und angewandte Naturwissenschaften.
81. O’Dowed, E.L., Baldwin, D.R. (2018): Lung cancer screening—low dose CT for lung cancer screening: recent trial results and next steps. *Br. J. Radiol.* 91, 1–6.
82. Paci, E., Puliti, D., Pegna, A.L., Carrozzi, L., Picozzi, G., Falaschi, F., Pistelli, F., Aquilini, F., Ocello, C., Zappa, M., Carozzi, F.M., Mascalchi, M. (2017): Mortality, survival and incidence rates in the ITALUNG randomised lung cancer screening trial. *Thorax.* 72, 825–831.
83. Papula, L. (2009): *Mathematische Formelsammlung für Ingenieure und Naturwissenschaftler: mit zahlreichen Rechenbeispielen und einer ausführlichen Integraltafel*. 9. Aufl. Wiesbaden: Vieweg und Teubner Verlag.
84. Papula, L. (2016): *Vektoranalysis, Wahrscheinlichkeitsrechnung, Mathematische Statistik, Fehler- und Ausgleichsrechnung: mit 550 Abbildungen, zahlreichen Beispielen aus Naturwissenschaft und Technik sowie 295 Übungsaufgaben mit ausführlichen Lösungen*. 7. Aufl. Wiesbaden: Springer Vieweg Verlag.
85. Pastorino, U., Rossi, M., Rosato, V., Marchianò, A., Sverzellati, N., Morosi, C., Fabbri, A., Galeone, C., Negri, E., Sozzi, G., Pelosi, G., La Vecchia, C. (2012): Annual orbiennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. *Eur. J. Cancer. Prev.* 21, 308–315.

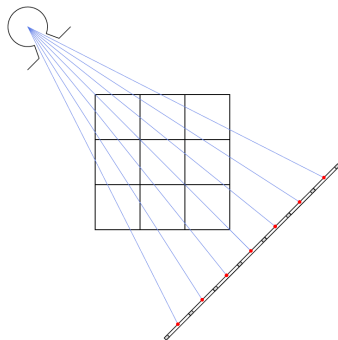
86. Philips Healthcare. (2021): Innovations–O–MAR metal artifact reduction algorithm. Automatic metal artifact reduction algorithm for more confident contouring. Abgerufen am 29.06.2021 von <https://www.philips.com.au/healthcare/education-resources/publications/hotspot/omar-metal-artifact-reduction>.
87. Prayer, F., Röhrich, S., Pan, J., Hofmanninger, J., Langs, G., Prosch, H. (2020): Künstliche Intelligenz in der Bildgebung der Lunge. *Radiologe*. **60**, 42–47.
88. Prokop, M. (2001): Dosisoptimierung in der thorakalen Computertomographie. *Radiologe*. **41**, 269–278.
89. Prosch, H., Ebner, L. (2019): Lungenkarzinom–Screening. *Radiologe*. **59**, 19–22.
90. Radon, J. (1917): Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Ber. Verh. Sächs. Akad. Wiss. Leipzig*. **69**, 262–277.
91. Ramirez–Giraldo, J.C., Grant, K.L., Raupach, R. (2017): Advanced modeled iterative reconstruction. *Siemens Healthineers*. **1**, 1–16.
92. Rampado, O., Depaoli, A., Marchisio, F., Gatti, M., Racine, D., Ruggeri, V., Ruggirello, I., Darvizeh, F., Fonio, P., Ropolo, R. (2021): Effects of different levels of CT iterative reconstruction on low–contrast detectability and radiation dose in patients of different sizes: an anthropomorphic phantom study. *Radiol. Med.* **126**, 55–62.
93. Rasch, B., Friese M., Hofmann, W., Naumann, E. (2008): Voraussetzungen für die Anwendung eines t–Tests, die Konstruktion eines t–Tests. In: Rasch, B. (Hrsg.): *Quantitative Methoden: Einführung in die Statistik*. 2. Aufl. S. 59–90. Heidelberg: Springer Medizin Verlag.
94. Saghir, Z., Dirksen, A., Ashraf, H., Bach, K.S., Brodersen, J., Clementsen, P.F., Døssing, M., Hansen, H., Kofoed, K.F., Larsen, K.R., Mortensen, J., Rasmussen, J.F., Seersholm, N., Skov, B.G., Thorsen, H., Tønnesen, P., Pedersen, J.H. (2012): CT screening for lung cancer brings forward early disease. The randomised danish lung cancer screening trial: Status after five annual screening rounds with low–dose CT. *Thorax*. **67**, 296–301.
95. Siegmund–Schultze, N. (2019): Onkologie: Wendepunkt Beim Lungenkrebs–Screening für Risikogruppen. *Dtsch. Arztebl.* **116**, 686–689.
96. Siemens Healthineers. (2020): Ansätze zur Dosisreduzierung in der Computertomographie. Abgerufen am 09.04.2021 von <https://www.siemens-healthineers.com/de-ch/medical-imaging/low-dose/low-dose-information-by-modality/low-dose-in-computed-tomography/computed-tomography-approaches-to-dose-reduction>.
97. Siemens Healthineers. (2021): iMAR. Abgerufen am 14.05.2021 von <https://www.siemenshealthineers.com/de/computed-tomography/options-upgrades/clinical-applications/imar>.
98. Siemens Healthineers. (2021): SOMATOM Force. Zwei Schritte voraus mit Dual–Source–CT. Abgerufen am 29.06.2021 von <https://www.siemens-healthineers.com/de/computed-tomography/dual-source-ct/somatom-force>.

99. Spekowius, G., Wendler, T. (2006): *Advances in Healthcare Technology: Shaping the Future of Medical Care*. Dordrecht: Springer Verlag.
100. Stiller, W. (2011): Grundlagen der Mehrzeilendetektor–Computertomographie: Teil 1: Technischer Aufbau und physikalisch–technische Grundlagen. *Radiologe*. 51, 625–640.
101. The MathWorks, Natick, Vereinigte Staaten von Amerika (2018).
102. Tsalafoutas, I.A., Kharita, M.H., Al–Naemi, H., Kalra, M.K. (2020): Radiation dose monitoring in computed tomography: Status, options and limitations. *Phys. Med.* 79, 1–15.
103. Turbell, H. (2001): *Cone–beam reconstruction using filtered backprojection*. Dissertation. LiU Linköping: Fachbereich Elektroingenieurwesen.
104. Tuy, H.K. (1983): An inversion formula for cone–beam reconstruction. *SIAM J. Appl. Math.* 43, 546–552.
105. U.S. Preventive Services Task Force. (2021): Final recommendation statement. Lung cancer: Screening. Abgerufen am 03.09.2021 von <https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/lung-cancer-screening>.
106. Verdun, F., Racine, D., Ott, J., Tapiovaara, M., Toroi, P., Bochud, F., Veldkamp, W., Schegerer, A., Bouwman, R., Giron, I.H., Marshall, N., Edyvean, S. (2015): Image quality in CT: From physical measurements to model observers. *Phys. Med.* 31, 823–843.
107. Viry, A., Aberle, C., Racine, D., Knebel, J.–F., Schindera, S.T., Schmidt, S., Becce, F., Verdun, F.R. (2018): Effects of various generations of iterative CT reconstruction algorithms on low–contrast detectability as a function of the effective abdominal diameter: A quantitative task–based phantom study. *Phys. Med.* 48, 111–118.
108. Weiß, C. (2013): *Basiswissen Medizinische Statistik: mit 20 Tabellen*. 6. Aufl. Berlin, Heidelberg: Springer Medizin Verlag.
109. Wielpütz, M.O., Wroblewski, J., Lederlin, M., Dinkel, J., Eichinger, M., Koenigkam–Santos, M., Biederer, J., Kauczor, H.–U., Puderbach, M.U., Jobst, B. J. (2015): Computer–aided detection of artificial pulmonary nodules using an ex vivo lung phantom: Influence of exposure parameters and iterative reconstruction. *Eur. J. Radiol.* 84, 1005–1011.
110. Willeminck, M.J., Noël, P.B. (2019): The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence. *Eur. Radiol.* 29, 2185–2195.
111. Wirtz, M., Kutschmann, M. (2007): Analyse der Beurteilerübereinstimmung für kategoriale Daten mittels Cohens Kappa und alternativer Maße. *Rehabilitation*. 46, 370–377.
112. Ye, K., Chen, M., Zhu, Q., Lu, Y., Yuan, H. (2021): Effect of adaptive statistical iterative reconstruction–V (ASiR–V) levels on ultra–low–dose CT radiomics quantification in pulmonary nodules. *Quant. Imag. Med. Surg.* 11, 2344–2353.

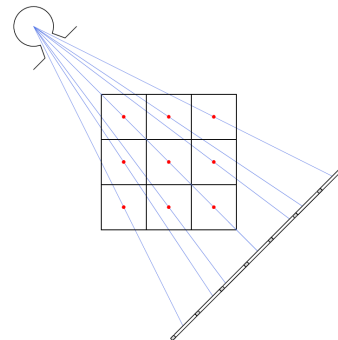
113. Yousaf–Khan, U., van der Aalst, C., de Jong, P.A., Heuvelmans, M., Scholten, E., Lammers, J.–W., van Ooijen, P., Nackaerts, K., Weenink, C., Groen, H., Vliegenthart, R., ten Haaf, K., Oudkerk, M., de Koning, H. (2017): Final screening round of the NELSON lung cancer screening trial: The effect of a 2.5–year screening interval. *Thorax*. 72, 48–56.
114. Zeng, W., Zeng, L.–M., Xu, X., Hu, S.–X., Liu, K.–L., Zhang, J.–G., Peng, W.–L., Xia, C.–C., Li, Z.–L. (2021): Noise reduction effect of deep–learning–based image reconstruction algorithms in thin–section chest ct. *Sichuan Da Xue Xue Bao Yi Xue Ban*. 52, 286–292.
115. Zhang, J., Teng, J., Bai, Y. (2019): Improving sparse compressed sensing medical CT image reconstruction. *Aut. Control. Comp. Sci.* 53, 281–289.
116. Zhang, Y., Jiang, B., Zhang, L., Greuter, M.J.W., de Bock, G.H, Zhang, H., Xie, X. (2021): Lung nodule detectability of artificial intelligence–assisted ct image reading in lung cancer screening. *Curr. Med. Imaging*. 17.

# A Anhang

## A.1 Projektoren



(a) Strahlbasierter Projektor



(b) Voxelbasierter Projektor

Abb. A.1: Prinzip eines (a) strahlbasierten und eines (b) voxelbasierten Projektors für eine Detektorzeile. Dabei betrachtet der strahlbasierte Projektor ausgehend von der Quelle Geraden, deren Endpunkt in der Mitte eines Dixel (rote Punkte, links) liegen. Die Durchstrahlungslängen durch die einzelnen Voxel werden dabei als Strahl definiert. Berechnungsmethoden und Varianten eines strahlbasierten Projektors vergleicht (Turbell, 2001, S. 141 ff.). Der voxelbasierte Projektor definiert den Strahl durch Geraden, welche die einzelnen Voxel im Mittelpunkt (rote Punkte, rechts) schneiden. Die Endpunkte der Gerade auf dem Detektor werden dabei an beliebigen Stellen getroffen. Projektor und Berechnungsverfahren sind einsehbar in (De Man & Basu, 2003).

## A.2 Datenblätter und Abnahmeprotokoll






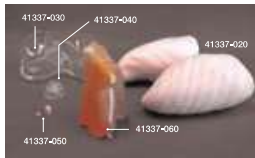




Multipurpose Chest Phantom N1 "LUNGMAN"   X-ray Phantom_4	
<h3 style="margin: 0;">Comparison</h3> <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>X-Ray Image <span style="color: red;">○</span> Simulated tumors (HU# 100)</p>  <p>CT Image <span style="color: red;">○</span> Simulated tumors (HU# 100)</p>  </div> <div style="width: 50%;"> <p>Comparison between Plain X-ray and CT, as well as between these images and the direct observation of the phantom, helps trainees to have three dimensional understanding and to improve X-ray interpretation skills.</p>  <p>Improve interpretation skills</p> </div> </div>	
<h3 style="margin: 0;">Multipurpose Chest Phantom N1 "LUNGMAN" PH-1</h3> <div style="display: flex; justify-content: space-between;"> <div style="width: 30%;">  </div> <div style="width: 65%;"> <p><b>Set Includes:</b></p> <ul style="list-style-type: none"> <li>1 male chest torso</li> <li>main body: synthetic bones are embedded</li> <li>internal parts: separates into four parts</li> <li>mediastinum: heart, trachea</li> <li>pulmonary vessels (right and left)</li> <li>abdomen (diaphragm) block: no internal structure</li> <li>30 simulated tumors</li> <li>3 varieties of Hounsfield number: approx -800, -630, +100</li> <li>5 sizes for each type: diameters 3, 5, 8, 10, 12 mm</li> </ul> <p><b>Materials:</b></p> <ul style="list-style-type: none"> <li>soft tissue: polyurethane (gravity 1.06)</li> <li>synthetic bones: epoxy resin</li> <li>phantom size: 43 x 40 x 48H cm, chest girth 94 cm</li> <li>weight: approx. 18 kg</li> <li>packing size: 30 x 70 x 50 cm, 27 kg</li> </ul> <p style="text-align: right; font-size: small;">Specifications are subject to change.</p> </div> </div>	
<div style="display: flex; justify-content: space-between;"> <div style="width: 30%;"> <p><b>Optional parts</b></p>  <p>41337-010 Chest Plates</p> </div> <div style="width: 65%;">  <p>41337-030 41337-040 41337-020 41337-050 41337-060</p> <p><b>Components for Radioisotope</b></p> <ul style="list-style-type: none"> <li>41337-020 Lungs of urethane</li> <li>41337-030 Liver RI container</li> <li>41337-040 Gallbladder RI container</li> <li>41337-050 Pulmonary nodule RI container</li> <li>41337-060 Mediastinum with left myocardium RI container</li> <li>41337-070 Simulated tumors</li> </ul> <div style="display: flex; justify-content: space-around; font-size: x-small;"> <div style="text-align: center;">  <p>CT</p> </div> <div style="text-align: center;">  <p>PET</p> </div> <div style="text-align: center;">  <p>Fusion</p> </div> </div> </div> </div>	
 <p>www.kyotokagaku.com rw-kyoto@kyotokagaku.co.jp</p>	
Head Office: 15 Kitaneokoya-cho, Fushimi-ku, Kyoto, 612-8388, JAPAN Tel: +81-75-605-2510 Fax: +81-75-605-2519 USA Office: 3109 Lomita Boulevard, Torrance, CA 90505-5108, USA Tel: 1-310-325-8860 Fax: 1-310-325-8867	

Abb. A.2: Produktblatt des Lungman-N1-Phantoms. Die Aufnahmen (oben, links) zeigen Beispiele für konventionelle Röntgen- bzw. CT-Bilder mit dem Phantom und (oben, rechts) Trainingsbilder am Alternator. Aufgeführt werden (mittig) die Materialangaben, Gewicht, Modellkomponenten und -abmaße sowie die zugehörigen auf Gewebe kalibrierten HU-Werte (bei 120 kV Röhrenspannung). Weitere Angaben erfolgen bezüglich (unten) Zusatzkomponenten und Modifikationen für das Phantom.

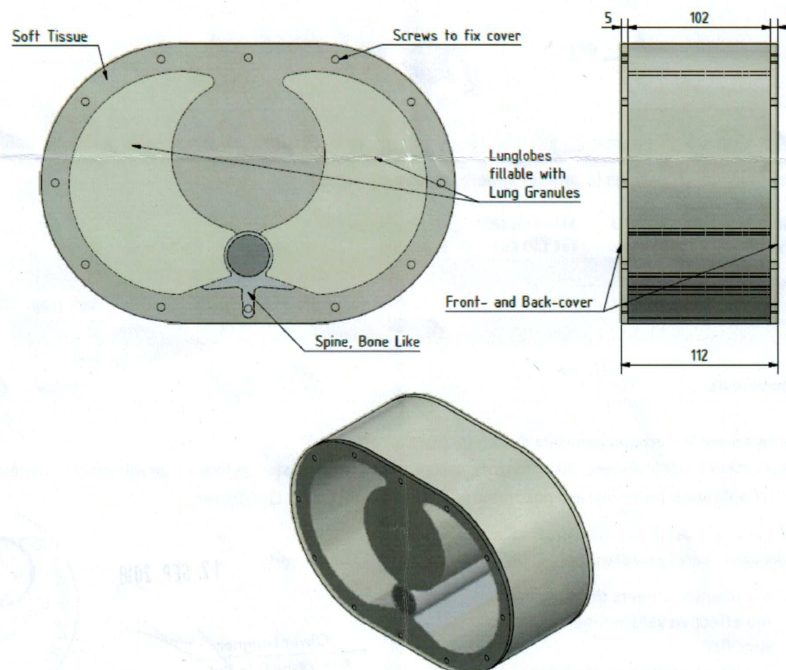
<b>Phantom:</b>	QRM-Lung-Nodule-Phantom and Lung Nodules + Extension Rings
<b>Serial:</b>	QSA-708 (Thorax), QSA-709 (Spherical Nodules), QSA-710 (Spiculated Nodules), QSA-708/M (Ring M, Soft Tissue), QSA-708/L (Ring L, Adipose Tissue)
<b>Project:</b>	B18061, B18062, B18063

**Lung-Nodule-Phantom:**

The Lung-Nodule-Phantom offers two hollow lung lobes to fill with lung granules and place lung nodules of different densities inside. The Lung Material can be replaced to receive different densities of lung tissue.

**PLEASE NOTE:**

Handle phantom with care during lung material replacement and lung nodule positioning. Screw in the screws of the top cover very carefully.



Lung Nodule Phantom, general dimensions – all measures in mm

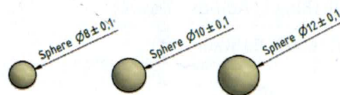
Abb. A.3: Erste Seite des Abnahmeprotokolls des QRM-Thorax-Phantoms, der Lungenrundherdmodelle und der Extensionsringe. Aufgeführt sind die Modellteile mit Serien- und Projektnummern (oben), die Drauf-, Seiten- und Isoansicht des konstruierten QRM-Bodys mit Bemaßungen und Beschriftung (mittig, unten).



### Nodule specification and Measurements

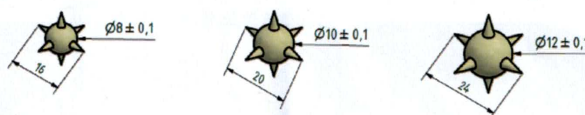
CT measurements are performed on a Siemens SOMATOM Definition Flash CT using a dedicated CT scan protocol. All given CT-values (HU) are based on that CT scanner.

*\*Effective values measured on the CT system at 120 kVp at a 20mm diameter cylinder (of the material) in air. Tolerances within  $\pm 10$  HU.*



### Spherical Nodules

Nodule Type #	Nominal HU (at 120 kV)	Effective HU* (at 120 kV)	Ø / Quantity			Color
			8 mm	10 mm	12 mm	
1	-50 HU	-49 HU	1	1	1	dark gray
3	-690 HU	-682 HU	1	1	1	beige



### Spherical-Spiculated Nodules (diameter sphere/sphere+spiculae)

Nodule Type #	Nominal HU (at 120 kV)	Effective HU* (at 120 kV)	Ø / Quantity			Color
			8/16 mm	10/20 mm	12/24 mm	
4	100 HU	90 HU	1	1	1	opaque
5	-50 HU	-53 HU	1	1	1	dark gray
6	-690 HU	-682 HU	1	1	1	beige

### Acceptance tests

The phantoms are tested mechanically during production process.

All measurement tools (callipers, micrometres, scales, etc.) are calibrated in fixed time intervals to national standards. Calibration protocols are documented within our ISO 9001 QM-System.

The undersigned hereby certifies that

- the phantom meets the specification
- the effective values meets the specification
- the phantom passed acceptance test

Möhrendorf, 12. SEP. 2018

Oliver Langner

QRM / Baidersdorfer Str. 22  
91096 Möhrendorf



**Important Note:**

The above-appointed QRM phantom (test device) is **NOT** a Medical Device and is not intended to diagnose, treat, cure or prevent any disease! QRM phantoms are designed and manufactured for the use in **research studies and experimental setups**.

Abb. A.4: Zweite Seite des Abnahmeprotokolls des QR M–Thorax–Phantoms, der Lungenrundherdnmodelle und der Extensionsringe. Aufgeführt sind die Spezifikationen für die Messungen und Lungenrundherdmodelle (Modalität, Kalibration und Toleranzen) der Abnahmeprüfung (oben) sowie die Modelleigenschaften der Kugeln und Spikulae (mittig) und die unterschriebene Zertifizierung des Herstellers (unten).



Quality Assurance in  
Radiology and Medicine

QRM-Extensionrings

## Extension / Obese Rings

To simulate obese patients these Extension Rings can be used with our standard Thorax, Cardio, Abdomen and Liver Phantoms.

Our extension rings are a good choice if you want to test the influences on image quality in CT and other 3D imaging techniques of higher attenuation due to an increased abdominal or chest girth.

The extension rings fits properly on our standard phantoms as there are the QRM-Thorax, QRM-Cardio-Phantom, QRM-Abdomen or the QRM-Liver and the QRM-Oval-Body.

The basic material is available as tissue equivalent, fat or water equivalent. Other materials (specified CT-values HU) are available upon request.

### Specifications

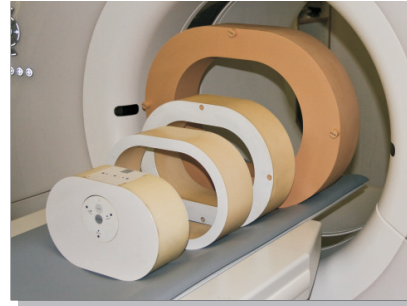
Extension Ring **M** outside ..... 350 x 250 mm  
inside ..... 300 x 200 mm  
height ..... 100 mm

Extension Ring **L** outside ..... 400 x 300 mm  
inside ..... 300 x 200 mm  
height ..... 100 mm

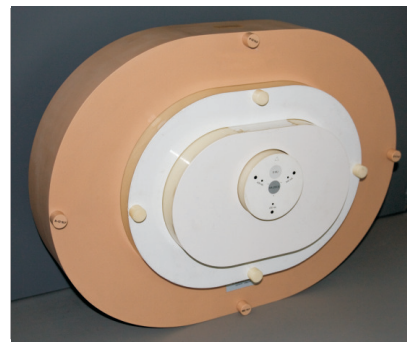
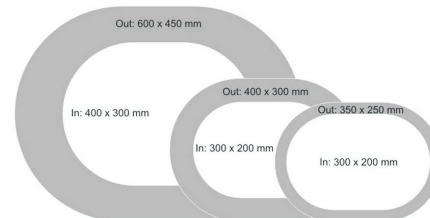
Extension Ring **XL** outside ..... 600 x 450 mm  
inside ..... 400 x 300 mm  
height ..... 100 mm

*If not specified otherwise, we will deliver the ring based on our standard material (tissue equivalent resin, valid for 120 kV).*

*Extension rings are also available with drilled holes for dosimetry purposes.*



QRM-Extensionrings



QRM-Extensionrings

WZ 21.121.04

QRM GmbH Dorfstrasse 4  
91096 Moehrendorf Germany

www.qrm.de  
info@qrm.de

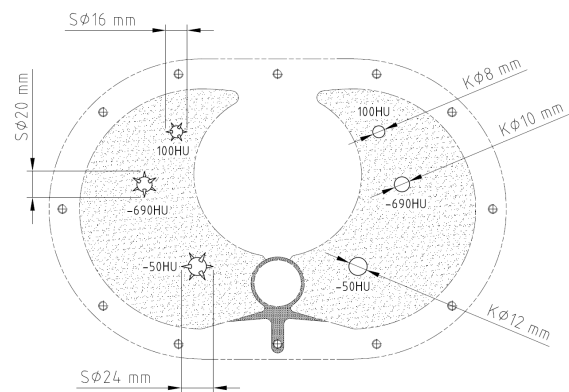
Tel: +49-(0)9133-6031-76  
Fax: +49-(0)9133-6031-78

Abb. A.5: Datenblatt der Extensionsringe. Die Modellspezifikationen (links) werden angegeben mit Materialangaben, Innen-, Außen- bzw. Höhenmaßen und Adaptionsempfehlungen. Des Weiteren ist das Material auf Gewebe zugeordnete HU-Werte bei 120 kV Röhrenspannung kalibriert worden. Die Ringansichten (rechts) bilden die Modelle innerhalb der CT-Gantry (oben), schematisch mit Maßangaben (mittig) und adaptiert (unten) ab.

### A.3 QRM–Thorax–Phantom mit Granulatfüllung



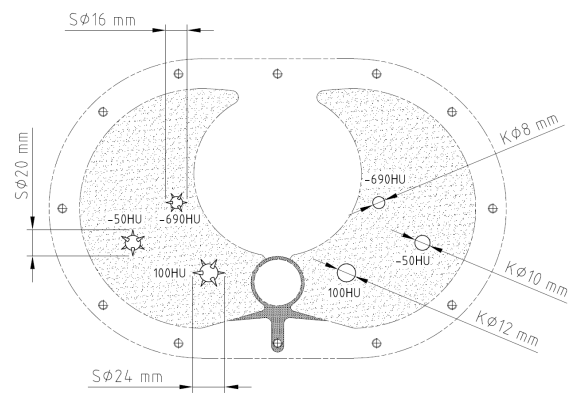
(a) QRM–Body, Draufsicht, Füllstand von 25 mm



(b) Modellkonstruktion, Draufsicht, Füllstand von 25 mm



(c) QRM–Body, Draufsicht, Füllstand von 50 mm



(d) Modellkonstruktion, Draufsicht, Füllstand von 50 mm

Abb. A.6: Draufsicht und maßstabsgerechte Konstruktion des QRM–Body mit Granulatfüllung. Abgebildet sind die Positionen der Targets bei einem Füllstand von 25 mm (a) und 50 mm (b) Granulat innerhalb des Phantoms. Die Konstruktionszeichnungen (b) und (d) zeigen die absoluten Durchmessermaße der Kugeln, K bzw. Spikulae, S sowie die HU–Werte. Ferner befinden sich die Kugeln im rechten und die Spikulae im linken Modelllungenlappen.

## A.4 Technische Spezifikationen des CT–Scanners

Tab. A.1: Technische Spezifikationen und Produktdetails des Dual–Source–CT Somatom Force (Siemens Healthineers, Forchheim, Deutschland). Die kumulierten Spezifikationen orientieren sich an (Healthineers, 2021).

Technische Spezifikationen	Produktdetails
Röhre	2x Vectron <sup>TM</sup> –Röntgenröhre
Detektoren	2x Stellar <sup>Infinity</sup> –Detektor mit 3D–Antistreustrahlenkollimator
Anzahl der akquirierten Schichten	384
Rotationsdauer	bis zu 0,25 s
Zeitliche Auflösung	bis zu 66 ms
Generatorleistung	240 kW
kV–Werte	70 bis 150 kV
Räumliche Auflösung	0,24 mm
Maximale Scangeschwindigkeit	737 mm/s mit Turbo Flash
Tischbelastung	bis zu 307 kg
Gantryöffnung	78 cm

## A.5 Prüfbild 212 mit ROI–Messregionen für die Konstanzprüfung

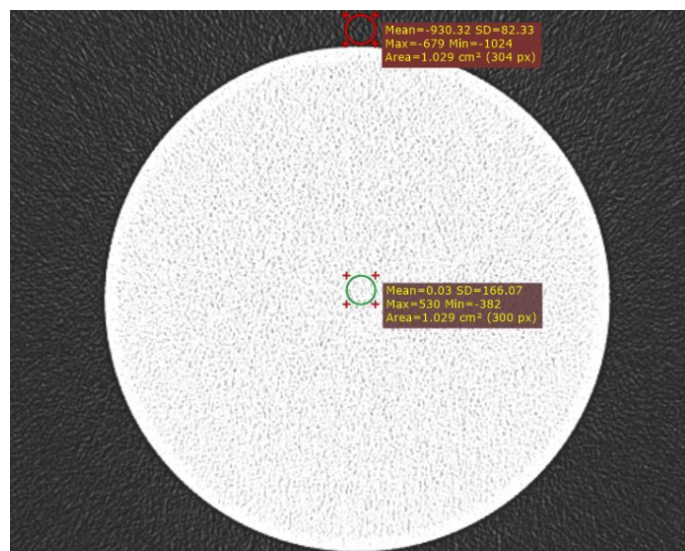


Abb. A.7: Axiales Wasserphantom\_comp\_MxT Modeltype 10742380 Prüfbild 212, CTD/ Kern/ Rekonstruktion/ Schichtdicke = LD/ Bl57 FBP/ 1 mm, zur Messung objektiver Bildqualitätsmaße. Es wurde in den Messregionen im Zentrum des Wasserphantoms (ROI<sub>W</sub>, grün) und in Luft (ROI<sub>L</sub>, rot) der mittlere HU–Wert (Mean, gelb), die zugehörige Standardabweichung (SD, gelb) sowie der Pixelwert (Area (px), gelb) gemessen.

## A.6 HU–Profile der 100 HU– und –50 HU–Kugel

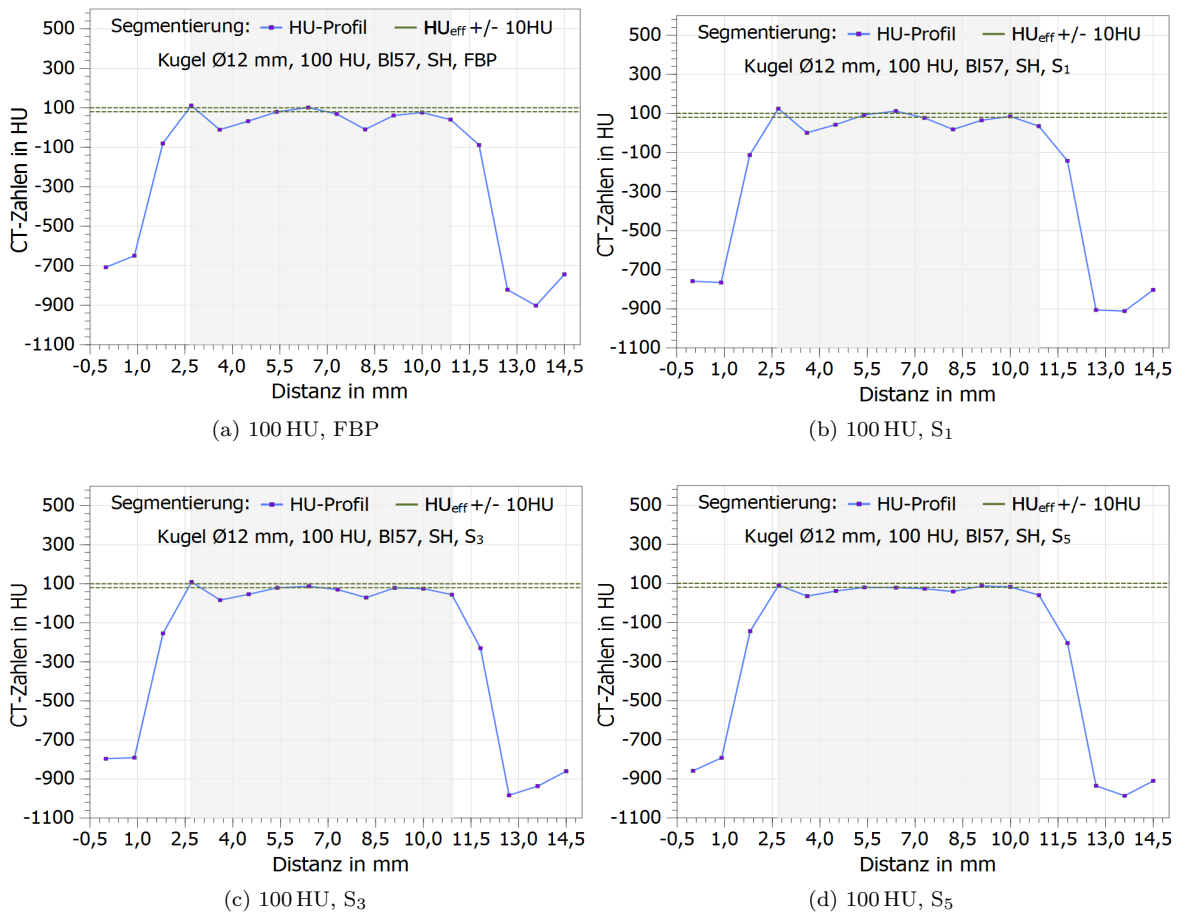


Abb. A.8: Segmentierte HU–Profile aus den Lungman–N1–Kugelsetup–CT–Bildern für die 100 HU–Kugel mit 12 mm Durchmesser. Die Darstellung erfolgt für das CT–Dosisprotokoll SH, den Standard–Lungen–BI57–Kern und die Rekonstruktionsalgorithmen FBP (a) sowie ADMIRE in den IR–Stärken S<sub>1</sub>, S<sub>3</sub>, S<sub>5</sub> (b) bis (d). Die Profile werden über die CT–Zahlen in HU und über die Segmentierungslinie in mm abgebildet. Die grau hinterlegten HU–Profilwerte liegen innerhalb der Kugel. In Bezug auf die Bildpixelgröße der verwendeten Rekonstruktionen von 0,824 mm<sup>2</sup> ergab sich eine Auflösung von 1,101 Pixel/mm.



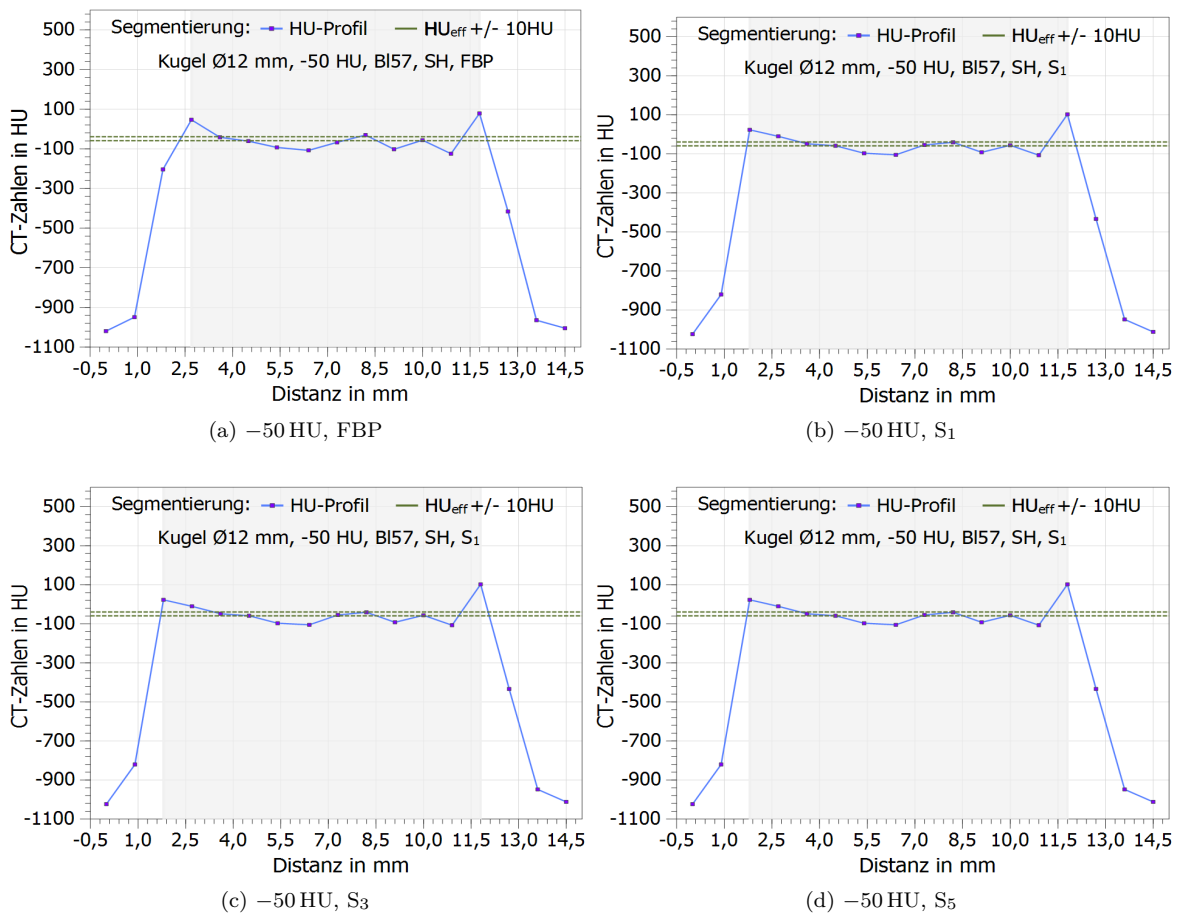


Abb. A.9: Segmentierte HU-Profile aus den Lungman-N1-Kugelsetup-CT-Bildern für die -50 HU-Kugel mit 12 mm Durchmesser. Die Darstellung erfolgt für das CT-Dosisprotokoll SH, den Standard-Lungen-BI57-Kern und die Rekonstruktionsalgorithmen FBP (a) sowie ADMIRE in den IR-Stärken S<sub>1</sub>, S<sub>3</sub>, S<sub>5</sub> (b) bis (d). Die Profile werden über die CT-Zahlen in HU und über die Segmentierungslinie in mm abgebildet. Die grau hinterlegten HU-Profilwerte liegen innerhalb der Kugel. In Bezug auf die Bildpixelgröße der verwendeten Rekonstruktionen von 0,824 mm<sup>2</sup> ergab sich eine Auflösung von 1,101 Pixel/mm.

## A.7 Rangordnungstestergebnisse für variable Phantomdurchmesser

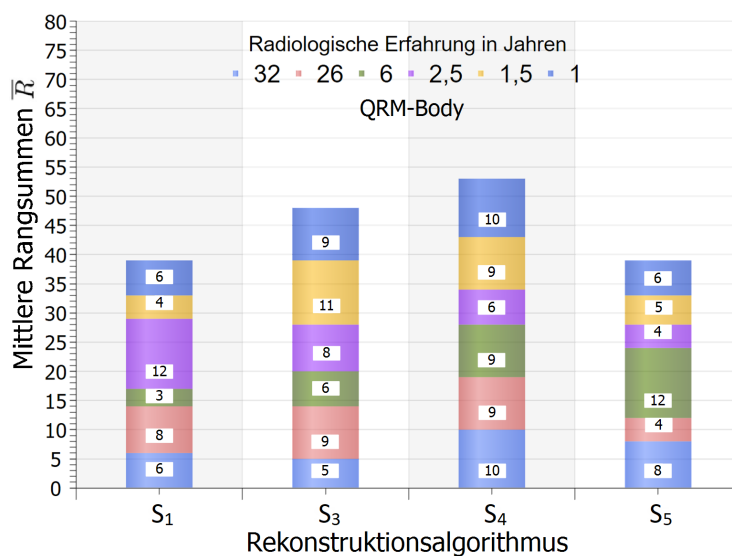


Abb. A.10: Kumulative und individuelle Ergebnisse des Rangordnungstests mit den CT-Bildern des QRM-Body aus Probe 2 mit dem Standard-Lungen-BI57-Kern. Die Darstellung erfolgt über den Rekonstruktionsalgorithmus ADMIRE in den IR-Stärken S<sub>1</sub>, S<sub>3</sub> bis S<sub>5</sub>. Die mittleren Rangsummen  $\bar{R}$  stellen die subjektiv empfundene Eignung der Prüfbilder zur Lungenrundherddetektion dar. Die radiologische Erfahrung wird den Farben zugeordnet: 1 Jahr (dunkelblau), 1,5 Jahre (gelb), 2,5 Jahre (lila), 6 Jahre (grün), 26 Jahre (rot) und 32 Jahre (hellblau).

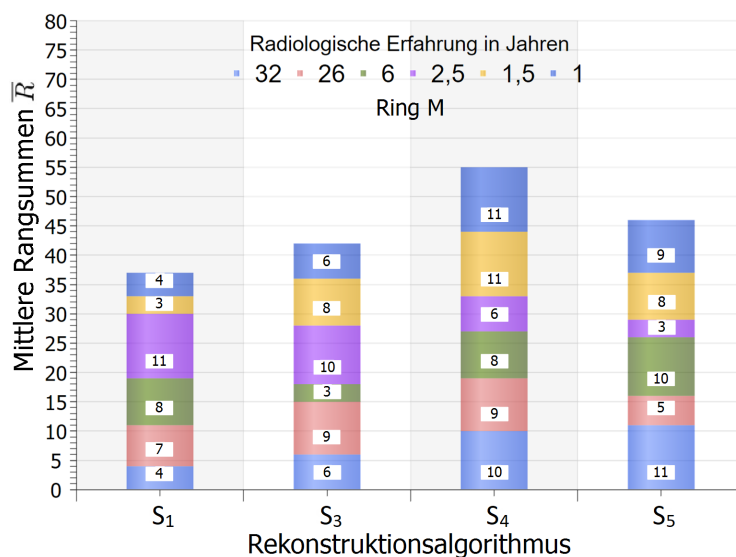


Abb. A.11: Kumulative und individuelle Ergebnisse des Rangordnungstests mit den CT-Bildern des QRM-Body mit adaptiertem Ring M aus Probe 2 mit dem Standard-Lungen-BI57-Kern. Die Darstellung erfolgt über den Rekonstruktionsalgorithmus ADMIRE in den IR-Stärken S<sub>1</sub>, S<sub>3</sub> bis S<sub>5</sub>. Die mittleren Rangsummen  $\bar{R}$  stellen die subjektiv empfundene Eignung der Prüfbilder zur Lungenrundherddetektion dar. Die radiologische Erfahrung wird den Farben zugeordnet: 1 Jahr (dunkelblau), 1,5 Jahre (gelb), 2,5 Jahre (lila), 6 Jahre (grün), 26 Jahre (rot) und 32 Jahre (hellblau).

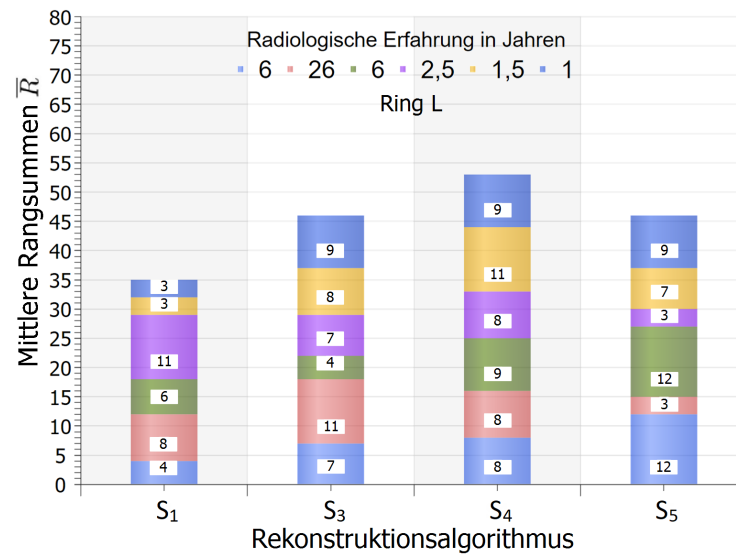


Abb. A.12: Kumulative und individuelle Ergebnisse des Rangordnungstests mit den CT-Bildern des QRM-Body mit adaptiertem Ring L aus Probe 2 mit dem Standard-Lungen-BI57-Kern. Die Darstellung erfolgt über den Rekonstruktionsalgorithmus ADMIRE in den IR-Stärken S<sub>1</sub>, S<sub>3</sub> bis S<sub>5</sub>. Die mittleren Rangsummen  $\bar{R}$  stellen die subjektiv empfundene Eignung der Prüfbilder zur Lungenrundherddetektion dar. Die radiologische Erfahrung wird den Farben zugeordnet: 1 Jahr (dunkelblau), 1,5 Jahre (gelb), 2,5 Jahre (lila), 6 Jahre (grün), 26 Jahre (rot) und 32 Jahre (hellblau).



## A.8 Tabellenverzeichnis

Tab. 1.1	Europäische und US-amerikanische randomisierte kontrollierte Studien zum Lungenkrebs-Screening mit Niedrigdosis-CT . . . . .	9
Tab. 1.2	Vorstellung der IR-Implementationen führender Hersteller . . . . .	11
Tab. 2.1	Rangdatenschema für den Friedman-Test zur Rangvarianzanalyse . . . . .	33
Tab. 2.2	Kontingenztabelle von Cohen's $\kappa$ . . . . .	36
Tab. 2.3	Interpretation von Cohen's $\kappa$ . . . . .	37
Tab. 2.4	Häufigkeitstabelle von $\kappa_{\text{Fleiss}}$ . . . . .	38
Tab. 2.5	Interpretation von $\kappa_{\text{Fleiss}}$ . . . . .	39
Tab. 2.6	Modelleigenschaften der artifiziellen Lungenrundherde für beide geometrischen Ausführungen . . . . .	41
Tab. 2.7	Scanparameter für die CT-Experimente mit beiden anthropomorphen Lungenphantomen . . . . .	45
Tab. 2.8	Eingesetzte Hardware mit Spezifikationen und Produktdetails . . . . .	48
Tab. 3.1	Mittelwert $\overline{\text{HU}}$ und Standardabweichung der Stichprobe $s$ aus den HU-Profilwerten innerhalb der Kugelmodelle . . . . .	53
Tab. 3.2	Subjektiv wahrgenommene Erkennbarkeit der Lungenrundherde anhand der Anzahl der individuellen Höchstbewertungen . . . . .	57
Tab. 3.3	Differenzen der Minimal- und Maximalwerte des $\overline{\text{SNR}}$ . . . . .	59
Tab. 3.4	Differenzen der Minimal- und Maximalwerte des $\overline{\text{CNR}}$ . . . . .	62
Tab. 3.5	Die objektiven Bildqualitätsparameter $\overline{\text{SNR}}$ und $\overline{\text{CNR}}$ aus den CT-Bildern mit Spikulae des Lungman-N1-CT-Experiments . . . . .	64
Tab. 3.6	Individuelle Friedman-Werte $\chi_r^2$ der einzelnen Rangordnungstestungen mit Probe 1 . . . . .	66
Tab. 3.7	Ergebnisse der Interrater-Reliabilitätsanalyse mit Probe 1 durch Fleiss' $\kappa$ . . . . .	67
Tab. 3.8	Individuelle Ergebnisse der Intrarater-Reliabilitätsanalyse mit Probe 1 durch Cohen's $\kappa$ und seiner Modifikation $\kappa_{\text{ney}}$ . . . . .	68
Tab. 3.9	Individuelle Friedman-Werte $\chi_r^2$ der einzelnen Rangordnungstestungen mit Probe 2 . . . . .	70
Tab. 3.10	Ergebnisse der Interrater-Reliabilitätsanalyse mit Probe 2 durch Fleiss' $\kappa$ . . . . .	71

---

Tab. 3.11	Individuelle Ergebnisse der Intrarater–Reliabilitätsanalyse mit Probe 2 durch Cohen’s $\kappa$ und seiner Modifikation $\kappa_{\text{ney}}$ . . . . .	73
Tab. 3.12	Scannerintegrierte CT–spezifische Dosisgrößen des Dosisreports und die abgeschätzte effektive Dosis . . . . .	74
Tab. A.1	Technische Spezifikationen und Produktdetails des Dual–Source–CT Somatom Force . . . . .	104

## A.9 Abbildungsverzeichnis

Abb. 2.1	Veranschaulichung der dritten CT-Generation . . . . .	15
Abb. 2.2	Zweidimensionale räumliche Abtastung mit einer Parallelstrahlgeometrie in der CT . . . . .	17
Abb. 2.3	Bandbegrenzter Ram-Lak-, Shepp-Logan- und Hamming-Filter im Fourierraum mit zugehörigem Kern im Radonraum . . . . .	20
Abb. 2.4	Zweidimensionale Veranschaulichung der Modellierung der Elemente $a_{ij}$ von der Systemmatrix $\mathbf{A}$ . . . . .	23
Abb. 2.5	Schema der ADMIRE-Implementation anhand ihrer Projektions- und Rauschkorrekturschleife . . . . .	25
Abb. 2.6	Lungenrundherdmodelle mit runder und spikulierter Geometrie . . . . .	40
Abb. 2.7	Anthropomorphes Lungman-N1-Phantom PH-1 R16511 . . . . .	41
Abb. 2.8	Ventrale und dorsale Ansicht des Lungman-N1-Phantoms mit Lungenrundherdmodellen und zugehöriger Konstruktionszeichnung . . . . .	42
Abb. 2.9	Anthropomorphes Lungenrundherd-QRM-Thorax-Phantom mit Extensionsringen . . . . .	43
Abb. 2.10	Draufsicht des QRM-Body mit 75 mm Granulatfüllung, Lungenrundherdmodellen und zugehöriger Konstruktionszeichnung . . . . .	44
Abb. 2.11	Waterphantom_comp_MxT Modeltype 10742380 innerhalb der CT-Gantry und zugehöriges Prüfbild . . . . .	45
Abb. 2.12	Lagerung des Lungman-N1-Phantoms und des QRM-Body innerhalb der CT-Gantry mit zugehörigen Prüfbildern . . . . .	46
Abb. 2.13	Lungman-N1-Prüfbild mit der Landmarke RL und LL der rechten und linken Lunge zur Messung der objektiven Bildqualitätsmaße . . . . .	50
Abb. 2.14	Vierfeld-Benutzerschnittstelle für den Rangordnungstest mit Probe 1 und Probe 2 . . . . .	51
Abb. 3.1	HU-Profile aus den Lungman-N1-Kugelsetup-CT-Bildern für den Modelltyp -690 HU . . . . .	53
Abb. 3.2	Individuelle Erkennbarkeit der artifiziellen Lungenrundherde anhand der sechsstufigen Likert-Skala . . . . .	55
Abb. 3.3	Mittleres $\overline{\text{SNR}}$ mit zugehöriger Standardabweichung des Mittelwerts der objektiven Bildqualitätsanalyse . . . . .	58

Abb. 3.4	Mittleres $\overline{\text{CNR}}$ mit zugehöriger Standardabweichung des Mittelwerts der objektiven Bildqualitätsanalyse . . . . .	61
Abb. 3.5	Die SNR- und CNR-Verteilungen der Lungman-N1-CT-Bilder mit dem Bl57-Kern, FBP- und modellbasierter IR-Rekonstruktion der Stärke $S_3$ .	63
Abb. 3.6	Kumulative und individuell wahrgenommene Eignung der CT-Bilder des Rangordnungstests mit Probe 1 zur Lungenrundherddetektion . . . . .	65
Abb. 3.7	Vergleich der Eignung der CT-Bilder des Rangordnungstests mit Probe 1 zur Lungenrundherddetektion . . . . .	66
Abb. 3.8	Ergebnisse der Intrarater-Reliabilitätsanalyse mit dem modifizierten $\kappa_{\text{ney}}$ -Test für den Rangordnungstest mit Probe 1 . . . . .	67
Abb. 3.9	Kumulative und individuell wahrgenommene Eignung der CT-Bilder des Rangordnungstests mit Probe 2 zur Lungenrundherddetektion . . . . .	69
Abb. 3.10	Vergleich der Eignung der CT-Bilder des Rangordnungstests mit Probe 2 zur Lungenrundherddetektion . . . . .	71
Abb. 3.11	Ergebnisse der Intrarater-Reliabilitätsanalyse mit dem modifizierten $\kappa_{\text{ney}}$ -Test für den Rangordnungstest mit Probe 1 . . . . .	72
Abb. A.1	Prinzip eines strahlbasierten und eines voxelbasierten Projektors . . . . .	98
Abb. A.2	Produktblatt des Lungman-N1-Phantoms . . . . .	99
Abb. A.3	Erste Seite des Abnahmeprotokolls des QRM-Thorax-Phantoms einschließlich der Lungenrundherdmodelle und der Extensionsringe . . . . .	100
Abb. A.4	Zweite Seite des Abnahmeprotokolls des QRM-Thorax-Phantoms einschließlich der Lungenrundherdmodelle und der Extensionsringe . . . . .	101
Abb. A.5	Datenblatt der Extensionsringe . . . . .	102
Abb. A.6	Draufsicht des QRM-Body mit 50 mm bzw. 25 mm Granulatfüllung, Lungenrundherdmodellen und zugehöriger Konstruktionszeichnung . . . . .	103
Abb. A.7	Axiales Prüfbild 212 des Waterphantom_comp_MxT mit ROI-Messregionen für die Konstanzprüfung . . . . .	104
Abb. A.8	HU-Profile aus den Lungman-N1-Kugelsetup-CT-Bildern für den Modelltyp 100 HU . . . . .	105
Abb. A.9	HU-Profile aus den Lungman-N1-Kugelsetup-CT-Bildern für den Modelltyp -50 HU . . . . .	106
Abb. A.10	Kumulative und individuelle Ergebnisse des Rangordnungstests mit Probe 2 und dem QRM-Body . . . . .	107

Abb. A.11 Kumulative und individuelle Ergebnisse des Rangordnungstests mit  
Probe 2 und dem adaptierten Ring M . . . . . 107

Abb. A.12 Kumulative und individuelle Ergebnisse des Rangordnungstests mit  
Probe 2 und dem adaptierten Ring L . . . . . 108

## A.10 Abkürzungsverzeichnis

<b>ADMIRE</b>	Advanced Modeled Iterative Reconstruction
<b>AIDR 3D</b>	Adaptive Iterative Dose Reduction 3D
<b>ASiR</b>	Adaptive Statistical Iterative Reconstruction
<b>ASiR-V</b>	Adaptive Statistical Iterative Reconstruction-V
<b>AUC</b>	Area Under the Curve (deutsch: Fläche unterhalb der Kurve)
<b>BfS</b>	Bundesamt für Strahlenschutz
<b>Bl57</b>	Body Lung 57 (deutsch: Körper Lunge 57)
<b>BMI</b>	Body-Mass-Index
<b>Br32</b>	Body Regular 32 (deutsch: Körper Regulär 32)
<b>Br69</b>	Body Regular 69 (deutsch: Körper Regulär 69)
<b>CAD</b>	Computer-Aided Detection (deutsch: Computergestützte Detektion)
<b>CARE Dose4D</b>	Produktname (Röhrenstrom-Modulationstechnik)
<b>CARE kV</b>	Produktname (Röhrenspannung-Modulationstechnik)
<b>CNN</b>	Convolutional Neuronal Network (deutsch: Neuronales Faltungsnetzwerk)
<b>CNR</b>	Contrast-to-Noise Ratio (deutsch: Kontrast-Rausch-Verhältnis)
<b>COSMOS</b>	Continuous Observation of Smoking Subjects
<b>CS</b>	Compressed Sensing (deutsch: Komprimiertes Abtasten)
<b>CT</b>	Computertomographie
<b>CTD</b>	Computertomographie-Dosisprotokoll
<b>CTDI</b>	Computertomographie-Dosisindex
<b>CTDI<sub>vol</sub></b>	Volumen-Computertomographie-Dosisindex
<b>CTDI<sub>w</sub></b>	gewichteter Computertomographie-Dosisindex
<b>DANTE</b>	Detection and screening of early lung cancer by novel imaging technology and molecular assays Trial
<b>DICOM</b>	Digital Imaging and Communications in Medicine (deutsch: Digitale Bildgebung und Kommunikation in der Medizin)

---

<b>DL</b>	Deep-Learning
<b>DLCST</b>	Danish Lung Cancer Screening Trial
<b>DLP</b>	Dosis-Längen-Produkt
<b>DMS</b>	Dose Monitoring System (deutsch: Dosisüberwachungssystem)
<b>DNN</b>	Deep Neuronal Network (deutsch: Tiefes Neuronales Netzwerk)
<b>EM</b>	Expectation-Maximaziation
<b>FBP</b>	Filtered Back Projection (deutsch: Gefilterte Rückprojektion)
<b>FDA</b>	Food and Drug Administration
<b>FIRST</b>	Forward Projected Model-Based Iterative Reconstruction Solution
<b>FFT</b>	Fast-Fourier-Transformation (deutsch: Schneller Fouriertransformationsalgorithmus)
<b>G-BA</b>	Gemeinsamer Bundesausschuss der großen Selbstverwaltungsorganisationen im Gesundheitswesen
<b>GUI</b>	Graphical User Interface (deutsch: Grafische Benutzerschnittstelle)
<b>HU</b>	Hounsfield Units (deutsch: Hounsfield-Einheiten)
<b>I-Mod</b>	Röhrenstrommodulation
<b>iDose</b>	Produktname (Iterative Rekonstruktionstechnik)
<b>iMAR</b>	Iterative Metal Artifact Reduction
<b>IMR</b>	Iterative Model Reconstruction
<b>IR</b>	Iterative Rekonstruktion
<b>ITALUNG</b>	Italian Lung Cancer Screening Trial
<b>IQWiG</b>	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
<b>KI</b>	Künstliche Intelligenz
<b>LD</b>	Low-Dose (deutsch: Niedrigdosis)
<b>LS</b>	Least Squares (deutsch: Kleinste-Quadrate-Schätzung)
<b>LUSI</b>	Lung Cancer Screening Intervention Trial
<b>NELSON</b>	Nederlands-Leuvens Longkankaer Screenings Onderzoek
<b>NLST</b>	National Lung Screening Trail
<b>MAP</b>	Maximum-a-Posteriori

---

<b>MRF</b>	Markoffsches Zufallsfeld
<b>ML</b>	Maschine Learning (deutsch: Maschinelles Lernen)
<b>ML–M</b>	Maximum–Likelihood Methode
<b>MILD</b>	Multicentric Italian Lung Detection Trial
<b>MO</b>	Model Observer (deutsch: Modellbeobachter)
<b>O–MAR</b>	Metal Artifact Reduction for Orthopedic Implants
<b>ROC</b>	Receiver Operating Characteristic–Curve (deutsch: Grenzwertoptimierungskurve)
<b>ROI</b>	Region of Interest (deutsch: Region von Interesse)
<b>S<sub>x</sub></b>	IR–Stärke auf dem Niveau x
<b>SAFIRE</b>	Sinogram Affirmed Iterative Reconstruction
<b>SH</b>	Standard–Hochkontrast
<b>SNR</b>	Signal–to–Noise Ratio (deutsch: Signal–Rausch–Verhältnis)
<b>UKLS</b>	United Kingdom Lung Cancer Screening Trial
<b>ULD</b>	Ultra–Low–Dose (deutsch: Ultra–Niedrigdosis)
<b>Veo</b>	Produktname (Modellbasierte iterative Rekonstruktionstechnik)
<b>FOV</b>	Field of View (deutsch: Sichtfeld)
<b>X–Care</b>	Produktname (Organbasierte Dosismodulationstechnik)



## B Danksagung

Mein Dank gilt Herrn Priv.–Doz. Dr. med. Hilmar Kühl und Herrn Prof. Dr. rer. nat. Waldemar Zylka für die freundliche und kompetente Begleitung meiner Dissertation. Ich bedanke mich vielmals für das zur Verfügung gestellte Promotionsthema, die wissenschaftliche Anleitung und das Vertrauen in meine Fähigkeiten.

Meinen Dank spreche ich Herrn Prof. Dr. rer. nat. Klaus Ewen aus, für die Motivation und die Empfehlung meiner Person zur Promotion. Ich danke Herrn Prof. Dr. med. Heinz Otto, Frau Priv.–Doz. Dr. med. Nika Guberina, Herrn Priv.–Doz. Dr. med. Uwe Keske, Herrn Dr. med. Sebastian Gayk, Frau Dr. med. Saskia Thüring, Herrn Dr. med. Lukasz Wszol sowie Herrn Dr. med. Christian Gernhardt für das Mitwirken an den subjektiven Bildqualitätsprüfungen.

Mein Dank gilt Herrn Dipl.–Phys. Jörg Wassenberg, Herrn Dr. rer. nat. Martin Altebockwinkel und Frau Sabine Altebockwinkel für das Korrekturlesen meiner Promotionsarbeit. Überdies bedanke ich mich bei Frau Hoai Thu Nguyen M.Sc. und Herrn Sebastian Blex für das Mitwirken an Implementierungsarbeiten und den CT–Experimenten.

Ein besonderer Dank gilt meinem Ehemann Ronny und meiner Mutter Elisabeth für die unterstützende Betreuung unserer Kinder, die gegebene Zuversicht mein Ziel konsequent weiterzuverfolgen und die Begleitung der vielen organisatorischen Wege.

Ich widme diese Arbeit meinen Kindern, die mir immer viel Glück gewünscht und sich bei jedem Teilerfolg für mich gefreut haben.

*Laurenz David,*

*Hanna Louisa,*

*Jonas Robin.*

## **C Lebenslauf**

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.