

Aus dem Nationalen Zentrum für Strahlenforschung in der Onkologie – OncoRay
Direktorin: Prof. Dr. Mechthild Krause

Evaluierung eines Detektionssystems
für prompte Gammastrahlung
zur Behandlungskontrolle bei klinischen
Protonentherapiebestrahlungen

D i s s e r t a t i o n s s c h r i f t
zur Erlangung des akademischen Grades
Doctor rerum medicinalium
(Dr. rer. medic.)
vorgelegt
der Medizinischen Fakultät Carl Gustav Carus
der Technischen Universität Dresden

von
M. Sc. Jonathan Berthold
aus Dresden / Sachsen

Dresden 2023

1. Gutachter: Prof. Dr. Wolfgang Enhardt
2. Gutachter: Prof. Dr. Dietmar Georg

Tag der mündlichen Prüfung: 23.08.2023

Prof. Dr. Christian Richter
Vorsitzender der Prüfungskommission

Kurzfassung

Die Protonentherapie zeichnet sich durch eine konformale und fokussierte Tumorbestrahlung aus, die es ermöglicht, gesundes Gewebe besser zu schonen als bei der konventionellen Strahlentherapie. Dieses Potential wird jedoch durch Unsicherheiten bei der Vorhersage der Protonenreichweite im Gewebe oder durch anatomische Veränderungen über den Verlauf der Therapie eingeschränkt. In der vorliegenden Arbeit wurde daher der klinische Nutzen eines Reichweiteverifikationssystems auf Grundlage von Prompt-Gamma-Imaging (PGI) zur Behandlungskontrolle untersucht. Dafür wurden Messungen mit einem PGI-System während Prostata- und Kopf-Hals-Tumor-Bestrahlungen durchgeführt und retrospektiv ausgewertet. Einerseits konnte dabei mittels PGI die Genauigkeit verschiedener Methoden zur Reichweitevorhersage überprüft werden. Es zeigte sich, dass die 2019 klinisch eingeführte Methode zur Reichweitevorhersage (DirectSPR) nicht von der mit PGI gemessenen Protonenreichweite in Prostata-Tumor-Bestrahlungen abweicht, wodurch die Reduktion der auf DirectSPR basierenden Reichweiteunsicherheiten unabhängig bestätigt werden konnte. Andererseits konnte die Detektionsfähigkeit von PGI bei der Erkennung relevanter und nicht relevanter anatomischer Veränderungen in applizierten Bestrahlungsfeldern nachgewiesen werden. Insbesondere wurde für die feldweise Klassifizierung der Prostata-Bestrahlungen eine Sensitivität und Spezifität von 74% bzw. 79% festgestellt. Damit konnte in dieser Dissertation erstmals systematisch das klinische Anwendungspotential eines Systems zur PGI-Reichweiteverifikation gezeigt werden. Als zusätzliche Untersuchung wurde in einer Kollaboration mit dem Massachusetts General Hospital zum ersten Mal ein Vergleich zwischen zwei verschiedenen, auf prompter Gammastrahlung basierenden Systemen zur Reichweiteverifikation durchgeführt. Dazu wurde ein standardisiertes Studienprotokoll etabliert, welches die Vergleichbarkeit und die klinische Implementierung von Reichweiteverifikationssystemen generell unterstützen könnte.

Abstract

Proton therapy is a conformal and focused irradiation of the tumor, which allows for a better sparing of healthy tissue than with conventional radiotherapy. However, this potential is limited by uncertainties from the proton range prediction in the patient or anatomical changes over the course of the treatment. Therefore, in this work, the clinical benefit of a range verification system based on the prompt-gamma-imaging (PGI) method for treatment verification was investigated. For this purpose, measurements were carried out with a PGI system during prostate and head and neck cancer irradiations and evaluated retrospectively. On the one hand, PGI was used to review the accuracy of several range prediction methods. The results showed that a specific method for range prediction (DirectSPR), which was clinically introduced in 2019, does not deviate from the PGI-measured proton range in prostate cancer irradiations. This means that the reduction of the range uncertainties with DirectSPR could be independently confirmed. On the other hand, the detection capability of PGI in identifying relevant and non-relevant anatomical changes in delivered treatment fields was demonstrated. In particular, for the fieldwise classification of prostate irradiations a sensitivity and specificity of 74% and 79% was determined, respectively. Thus, the clinical potential of a PGI range verification system was for the first time systematically demonstrated in this thesis. Furthermore, in a collaboration with the Massachusetts General Hospital a first-time comparison of two different range verification systems based on prompt gamma radiation was conducted. Therefore, a standardized study protocol was established, which could generally foster the comparability and clinical implementation of range verification systems.

Inhaltsverzeichnis

Abkürzungsverzeichnis	vii
Abbildungsverzeichnis	viii
Tabellenverzeichnis	x
1 Einleitung	1
2 Strahlentherapie mit Protonen	5
2.1 Physikalische Grundlagen der Protonentherapie	5
2.2 Behandlungsablauf in der Protonentherapie	8
2.2.1 Bildgebung zur Therapieplanung	8
2.2.2 Bestrahlungsplanung	10
2.2.3 Strahlapplikation	12
2.3 Genauigkeit in der Protonentherapie	15
2.3.1 Ursachen für Behandlungs- und Reichweiteunsicherheiten	15
2.3.2 Aktueller Stand der Behandlungs- und Reichweiteverifikation	17
3 Methodik der Reichweiteverifikation mittels Prompt-Gamma-Bildgebung (PGI)	23
3.1 Funktionsprinzip der PGI-Schlitzkamera	23
3.2 Datenaufnahme und -verarbeitung	24
3.2.1 Detektoraufbau und Signalaufnahme	24
3.2.2 PGI-Simulation und Bestimmung der Reichweiteabweichung	25
3.3 Charakterisierung des PGI-Prototyps	28
3.3.1 Kalibrierung des Systems	29
3.3.2 Positionierungspräzision	34
3.4 Überblick zur PRIMA-Studie	35
3.5 Experimentelle Studien zur PGI-Simulationsgenauigkeit	37
3.5.1 Abhängigkeit vom PGI-Sichtfeld und der Protonenenergie	37
3.5.2 Validierung der erweiterten Simulationssoftware	41
3.5.3 Abhängigkeit von der Tumorentität	43
3.5.4 Schlussfolgerungen	47

4	Validierung der CT-basierten Reichweitevorhersage mittels PGI	51
4.1	Konzept der Validierung	52
4.2	Gesamtabschätzung der Validierungsunsicherheit	53
4.3	Ergebnisse der Validierung	56
4.4	Diskussion	58
5	Detektionsfähigkeit anatomischer Veränderungen mittels PGI	61
5.1	Prinzipieller Aufbau der Studie	61
5.2	Grundwahrheit auf Basis von CT- und Dosisinformationen	62
5.2.1	Manuelle Klassifizierung	64
5.2.2	Klassifizierung auf Grundlage von integrierten Tiefendosisprofilen	65
5.2.3	Ergebnis der Etablierung einer CT-basierten Grundwahrheit	69
5.3	Etablierung einer Klassifikation auf Basis von PGI-Daten	70
5.3.1	Verarbeitung der PGI-Daten mittels Cluster-Algorithmus	71
5.3.2	Definition von spot- oder clusterbasierten Klassifikationsmodellen	72
5.4	Ergebnisse der PGI-Detektionsfähigkeit	74
5.4.1	Auswertung für Patienten mit Prostata-Tumor	74
5.4.2	Auswertung für Patienten mit Tumoren im Kopf-Hals-Bereich	78
5.5	Diskussion	81
6	Genauigkeit zweier Reichweiteverifikationsmethoden – bizentrischer Vergleich	89
6.1	Material und Methoden	89
6.1.1	Bildgebung	90
6.1.2	Bestrahlungsplanung	91
6.1.3	Durchführung und Auswertung	92
6.2	Ergebnisse	94
6.3	Diskussion	98
7	Zusammenfassung	101
8	Summary	105

Abkürzungsverzeichnis

ACC	Korrektklassifikationsrate, engl. <i>accuracy</i>
AUC	Fläche unter der Kurve, engl. <i>area under the curve</i>
CBCT	Kegelstrahl-Computertomographie, engl. <i>cone-beam CT</i>
CSDA	Näherung mittels kontinuierlicher Abbremsung, engl. <i>continuous slowing down approximation</i>
CT	Computertomographie
CTV	Klinisches Zielvolumen, engl. <i>clinical target volume</i>
DECT	Mittels zweier Röntgenspektren aufgenommener CT-Scan, engl. <i>dual-energy CT</i>
DirectSPR	Verfahren zur direkten Bestimmung von SPR-Werten mittels DECT
FOV	Sichtfeld, engl. <i>field of view</i>
GTV	Makroskopisches Tumolvolumen, engl. <i>gross tumor volume</i>
GUM	Engl. <i>Guide to the expression of Uncertainty in Measurement</i>
DVH	Dosis-Volumen-Histogramm
HLUT	Hounsfield Lookup-Tabelle
HU	Hounsfield Einheit, engl. <i>Hounsfield unit</i>
ITD	Integrierte Tiefendosis
kCT	Kontroll-CT
LYSO	Lutetium-Yttrium-Oxyorthosilikat
MGH	Massachusetts General Hospital
monoCT	Monoenergetische CT
MRT	Magnetresonanztomographie
MKK	Matthews Korrelationskoeffizient
MTRA	Medizinisch-technischer Radiologieassistent
OAPT	Online-adaptive Protonentherapie
OAR	Risikoorgan, engl. <i>organ at risk</i>
PBS	Scanning-Verfahren mit Bleistiftstrahl, engl. <i>pencil-beam scanning</i>
pCT	Planungs-CT
PET	Positronenemissionstomographie
PMMA	Polymethylmethacrylat
Protonen-CT	Protonen-Computertomographie
PG	Prompte Gammastrahlung

PGI	Prompt-Gamma-Bildgebung, engl. <i>Prompt-Gamma-Imaging</i>
PGPI	Prompt-Gamma-Peak-Integral
PGS	Prompt-Gamma-Spektroskopie
PGT	Prompt-Gamma-Timing
PRIMA	Patientenstudie mittels <i>PRompt-gamma-IMAGING</i>
PTV	Planungszielvolumen, engl. <i>planning target volume</i>
RBW	Relative biologische Wirksamkeit
ROC	Operationscharakteristik eines Beobachters, engl. <i>receiver operating characteristic</i>
SECT	Mittels eines Röntgenspektrums aufgenommener CT-Scan, engl. <i>single-energy CT</i>
SiPM	Silizium-basierter Photoelektronenvervielfacher, engl. <i>silicon photomultiplier</i>
SOBP	Ausgedehnter Bragg-Peak, engl. <i>spread-out Bragg-peak</i>
SPR	Bremsvermögen relativ zu Wasser, engl. <i>stopping-power ratio</i>
UKD	Universitätsklinikum Carl Gustav Carus Dresden
UPTD	Universitäts Protonen Therapie Dresden
WDF	Wahrscheinlichkeitsdichtefunktion

Abbildungsverzeichnis

2.1	Dosisvergleich Protonen und Photonen	7
2.2	Überblick zur SPR-Konversion	10
2.3	Dosisverteilung Prostata-Karzinom	12
2.4	Passive und dynamische Strahlapplikation	14
2.5	Emission prompter Gammastrahlung über Eindringtiefe	19
3.1	Skizze PGI-Prinzip	24
3.2	Detektoraufbau PGI-Kamera	25
3.3	PGI-Signalprozessierung	27
3.4	Generationen des PGI-Systems	28
3.5	Koordinatensystem zur Kalibrierung des PGI-Systems	30
3.6	Charakterisierung des PGI-Systems	33
3.7	Untersuchung des FOV unter 270° Gantrywinkel	38
3.8	Untersuchung des FOV unter 90° Gantrywinkel	39
3.9	Untersuchung der Invertierung des Detektorarrays	40
3.10	Untersuchung der lateralen Spotpositionen im FOV	41
3.11	Phantom-Halterung zur Simulationsvalidierung	42
3.12	Validierung der erweiterten Simulation	43
3.13	Prostata-Tumor-Patienten: Anwendung der FOV-Korrektur	44
3.14	Prostata-Tumor-Patienten: Begründung für Pelvis-Korrektur	45
3.15	Kopf-Hals-Tumor-Patienten: Verlauf der PGI-Reichweiteabweichung im FOV	47
3.16	Prostata-Tumor-Patienten: Heuristische Korrekturen	50
4.1	Validierung Protonenreichweite: Histogramm der Reichweiteabweichungen	57
4.2	Validierung Protonenreichweite: Histogramm der CT-Zahl auf Spotachse	57
5.1	Studiendesign: Detektion anatomischer Veränderungen	63
5.2	Beispiele anatomischer Veränderungen: Prostata	66
5.3	Beispiele anatomischer Veränderungen: Kopf-Hals	67
5.4	Grundwahrheit anatomischer Veränderungen: Prostata	70
5.5	Grundwahrheit anatomischer Veränderungen: Kopf-Hals	71
5.6	Visualisierung Cluster-Algorithmus	72
5.7	Korrelation von PGI- und ITD-Reichweiteabweichungen: Prostata	75

5.8	Ergebnisse für die Optimierung der PGI-Klassifikationsmodelle: Prostata . . .	76
5.9	Korrelation von PGI- und ITD- Reichweiteabweichungen: Kopf-Hals	78
5.10	Ergebnisse für die Optimierung der PGI-Klassifikationsmodelle: Kopf-Hals . .	80
5.11	Beispiel einer anatomischen Veränderung des Weichgewebes: Prostata . . .	84
6.1	Design der bizentrischen Vergleichsstudie zweier PG-Systeme	90
6.2	Experimenteller Aufbau des PGI-Systems für Vergleichsstudie	92
6.3	Anthropomorphes Kopfphantom: DirectSPR vs. Grundwahrheit	93
6.4	Abweichung der PG-basierten Systeme von der Absolutreichweite	96
6.5	Spotweise Präzision und detektierte Ereignisanzahl für PGI und PGS	97
6.6	Detektierbarkeit von relativen Reichweiteabweichungen	98

Tabellenverzeichnis

2.1	Spektrallinien prompter Gammastrahlung	18
3.1	Globale Korrekturen bei Kalibrierung des PGI-Systems	34
3.2	Überblick Prostata-Tumor-Patienten in dieser Studie	35
3.3	Überblick Kopf-Hals-Tumor-Patienten in dieser Studie	36
3.4	Überblick heuristische Korrekturen für die PGI-Auswertung	48
4.1	Unsicherheitsbeiträge der Reichweitevalidierung	56
5.1	Zusammenfassung Spotfilter und Clustering: Prostata	75
5.2	AUC, Sensitivität und Spezifität der PGI-Klassifikationsmodelle: Prostata	76
5.3	Aufteilung der Prostata-Tumor-Patienten zur Optimierung des PGI-Modells	77
5.4	Zusammenfassung Spotfilter und Clustering: Kopf-Hals	79
5.5	AUC, Sensitivität und Spezifität der PGI-Klassifikationsmodelle: Kopf-Hals	80
5.6	Aufteilung der Kopf-Hals-Tumor-Patienten zur Optimierung des PGI-Modells	81
6.1	Überblick zur Detektierbarkeit relativer Reichweiteabweichungen	98

1 Einleitung

Durch die Entdeckung der Röntgenstrahlung im Jahre 1895 durch Wilhelm C. Röntgen wurde der Grundpfeiler für eine neue Art der Krebstherapie gelegt: Die Strahlentherapie ermöglicht es, Tumore lokal und nichtinvasiv zu behandeln. Sie stellt neben dem operativen Eingriff und der Chemotherapie eine der wichtigsten Säulen der Krebstherapie dar, deren Erfolg nicht zuletzt aus dem Zusammenwirken der verschiedenen Behandlungsmethoden resultiert (Baumann et al., 2016). Eine Krebserkrankung ist weltweit eine der häufigsten Todesursachen mit mehr als 19,3 Millionen Neuerkrankungen und 10 Millionen Todesfällen im Jahr 2020 (Bray et al., 2021; Sung et al., 2021). Diese Entwicklung wird sich aufgrund einer stetig wachsenden und älter werdenden Weltbevölkerung fortsetzen, was der Krebsprävention und -forschung einen wichtigen Stellenwert zukommen lässt (White et al., 2014). Die 2014 am Universitätsklinikum Carl Gustav Carus Dresden (UKD) in Betrieb genommene Universitäts Protonen Therapie Dresden (UPTD) stellt einen wichtigen Forschungsstandort für die Strahlentherapie dar und bildet die Grundlage für die hier vorliegende Arbeit.

Bereits 1946 schlug der Physiker Robert R. Wilson vor, beschleunigte Protonen für die Krebstherapie einzusetzen (Wilson, 1946). Aufgrund der dabei auftretenden technischen Herausforderungen entstanden entsprechende Therapiezentren in Krankenhäusern jedoch erst gegen Ende des 20. Jahrhunderts. Ein Vorteil dieser Ionenstrahlung im Vergleich zur Photonenstrahlung liegt in der endlichen Reichweite der Ionen im Gewebe und der charakteristischen Energieabgabe entlang der Wegstrecke im Gewebe (Bragg-Kurve), wodurch potentiell eine bessere Schonung des umliegenden gesunden Gewebes erreicht werden kann (vgl. Kapitel 2). Die endliche Reichweite stellt jedoch gleichzeitig eine Herausforderung für die Bestrahlungsplanung und -applikation dar, insbesondere gegenüber anatomischen Veränderungen und Unsicherheiten in der Vorhersage der Protonenreichweite. Letztere beruht standardmäßig auf Patientenaufnahmen mittels Computertomographie (CT) und wird mit Unsicherheiten von etwa 3,5% der Protonenreichweite berücksichtigt (Paganetti, 2012). Die Reduktion von Unsicherheiten hat das Potential, wesentlich zur Schonung von gesundem Gewebe beizutragen und ist somit eine der aktuellen Herausforderungen für die moderne Protonentherapie.

Die Vision von einer online-adaptiven Protonentherapie (Albertini et al., 2020; Paganetti et al., 2021b), d. h. der Anpassung der Bestrahlung in Echtzeit, treibt die Entwicklung neuer Technologien und Methoden in der Protonentherapie entscheidend an und führt zu neuen Kooperationen zwischen Klinik, Forschung und Industrie (vgl. Wiegand, 2022). Die Zielstellung,

den Bestrahlungsplan täglich oder sogar in Echtzeit während der Bestrahlung anzupassen setzt einen Kreislauf bestehend aus 3D-Bildgebung des Patienten, Software zur schnellen Planadaptation und der Qualitätssicherung der Bestrahlung voraus. Aufgrund der angestrebten Reduktion von Sicherheitssäumen mit einem solchen Arbeitsablauf, rücken Messsysteme für die Kontrolle der Protonenbehandlungen in den Fokus, welche die Dosisdeposition bzw. Reichweite der Protonen im Gewebe detektieren können. Im Hinblick auf eine Echtzeit-Verifikation der Behandlung, d. h. noch während der Patient bestrahlt wird, hat sich insbesondere die Detektion von prompt emittierter Gammastrahlung aus Anregungen der Atomkerne im Gewebe als vielversprechend erwiesen. Obwohl diese Idee bereits 2003 vorgestellt wurde (Jongen & Stichelbaut, 2003), hat es bis 2015 gedauert, dass ein erstes Prototypsystem am Patienten zum Einsatz kam (Richter et al., 2016) und später auch in weiteren klinischen Bestrahlungen im dynamischen Scanning-Verfahren angewendet wurde (Xie et al., 2017, 2020). Das Detektionssystem, welches auch in dieser Arbeit zum Einsatz gekommen ist, beruht auf dem Prinzip der sogenannten Prompt-Gamma-Bildgebung, engl. *Prompt-Gamma-Imaging* (PGI). Bei dieser wird die Intensitätsverteilung der im Patienten emittierten Gammastrahlung bestimmt, wodurch auf die Reichweite bzw. Reichweiteabweichung der Protonen im Gewebe geschlossen werden kann.

Nachdem das Potential von PGI bereits in oben genannten Einzelfallstudien gezeigt werden konnte, bedarf es jedoch noch systematischer Studien mit einer Vielzahl an Patienten, um den klinischen Nutzen einer routinemäßigen Anwendung von PGI zu demonstrieren und zu quantifizieren. Folglich ist es die Hauptaufgabe dieser Arbeit, das klinische Anwendungspotential eines PGI-Systems der zweiten Generation, welches am OncoRay – Nationales Zentrum für Strahlenforschung in der Onkologie seit 2018 zur Verfügung stand, systematisch zu untersuchen. Eine maßgebliche Voraussetzung hierfür war die Kalibrierung und Charakterisierung des Systems (Kapitel 3). Im Rahmen einer klinischen Studie wurden Protonenbestrahlungen von insgesamt 15 Prostata-Tumor-Patienten und 4 Kopf-Hals-Tumor-Patienten mittels PGI in mehreren Bestrahlungsfractionen monitoriert und retrospektiv ausgewertet. Die gemessenen Daten zeichnen sich dadurch aus, dass vom Tag der Bestrahlung sowohl PGI-Aufnahmen als auch aktuelle CT-Scans in Behandlungsposition vorliegen, die mit einem im Behandlungsraum befindlichen CT-Scanner auf Schienen aufgenommen wurden. In einer ersten Untersuchung wurde mittels PGI die Genauigkeit von CT-basierter Reichweitevorschau im Patienten überprüft. Anlass dafür war nicht zuletzt ein neues Verfahren (DirectSPR) mit verbesserter Vorhersagegenauigkeit, welches auf einem mittels zweier Röntgenspektren aufgenommenen CT-Scan basiert und 2019 erstmals für die Bestrahlungsplanung in der UPTD eingeführt worden ist (Peters et al., 2022). Die Genauigkeit wurde systematisch für das bisherige Standardverfahren zur Reichweitevorschau sowie das DirectSPR- und ein darauf basierendes Verfahren anhand von fünf monitorierten Prostata-Patienten untersucht (Kapitel 4). Grundvoraussetzung für diese Validierung war außerdem die umfassende Abschätzung der Genauigkeit des PGI-Systems. In einer zweiten Untersuchung wurde die

Detektionsfähigkeit des PGI-Systems im Bezug auf anatomische Veränderungen unter der Bestrahlung ermittelt. Im Fokus stand hierbei die Erkennung von Reichweiteveränderungen der Protonen zwischen dem Tag der Bestrahlungsplanung und der eigentlichen Bestrahlung, die anhand der verfügbaren CT-Daten ermittelt werden konnten. Für die abgestrahlten Bestrahlungsfelder wurde pro Bestrahlungsfeld eine Grundwahrheit bzgl. relevanter und nicht relevanter Veränderungen definiert. Mit dieser wurde ein PGI-basiertes Klassifikationsmodell verglichen, um das Detektionsvermögen von PGI zu quantifizieren (Kapitel 5).

Als zusätzliche, dritte Untersuchung wurde in einer bizentrischen Studie das im Patienten untersuchte PGI-System mit dem am Massachusetts General Hospital (MGH) entwickelten klinischen Prototypen für Reichweiteverifikation mittels prompter Gammastrahlung verglichen. Dazu wurde im selben anthropomorphen Kopfphantom ein realistisch eingezeichnetes Zielvolumen im Gehirn unter klinischen Bedingungen bestrahlt. Dadurch konnten die Genauigkeit, Präzision und Robustheit der Messsysteme gegenüber der Absolutreichweite der Protonen oder einer relativen Reichweiteverschiebung untersucht werden (Kapitel 6). Primäres Ziel der Studie war die Etablierung standardisierter Experimente, um die objektive Vergleichbarkeit verschiedener Reichweiteverifikationssysteme und damit deren klinische Implementierung zu fördern.

2 Strahlentherapie mit Protonen

Die Krebsbehandlung mit beschleunigten Protonen ist eine spezielle Art der Partikeltherapie und stellt eine stetig an Bedeutung zunehmende Säule der externen Strahlentherapie dar, was sich anhand der wachsenden Anzahl von Patienten pro Jahr zeigt (Particle Therapy Co-Operative Group, 2022). Im Vergleich zur konventionellen Strahlentherapie mit Photonen erreicht man in der Partikeltherapie eine ebenso konformale Dosisabdeckung des Tumors bei gleichzeitig besserer Schonung des gesunden Gewebes (Normalgewebe), wie Abbildung 2.1 A beispielhaft zeigt. Die Ursache dafür liegt in den unterschiedlichen physikalischen Wechselwirkungen der Strahlungsarten und wird in diesem Kapitel genauer erläutert. Zahlreiche Studien zeigen bereits das Potential der Protonentherapie aufgrund der verbesserten Normalgewebsschonung bei gleicher Tumorkontrolle (Water et al., 2011; Yuan et al., 2019). Gegenstand aktueller Forschung ist es, den positiven Behandlungseffekt der Protonentherapie gegenüber anderer Therapieformen und Bestrahlungsmodalitäten, wie bspw. der konventionellen Photonentherapie, systematisch nachzuweisen (Langendijk et al., 2018; Mishra et al., 2017; Weber et al., 2020). Dabei stellt die Erhöhung der Genauigkeit und Präzision bei der Strahlenanwendung einen wichtigen Aspekt zur Verbesserung des Behandlungserfolgs dar (Knopf & Lomax, 2013). In diesem Kapitel sollen die Ursachen für Unsicherheiten in der Protonentherapie identifiziert werden. Ein wichtiger Teilaspekt hierbei ist die Verifikation der Dosisapplikation im Patienten, für welche im weiteren Verlauf die aktuellen Lösungsansätze vorgestellt werden.

2.1 Physikalische Grundlagen der Protonentherapie

Im Gegensatz zu Photonen wechselwirken Protonen aufgrund ihrer Ladung kontinuierlich über die gesamte Weglänge im Patienten mit dem umliegenden Gewebe. Dabei vollziehen sie elastische und inelastische Stöße mit den dort befindlichen Atomkernen und Hüllenelektronen. Die Energieübertragung für die in der Protonentherapie relevanten kinetischen Energien bis etwa 250 MeV erfolgt maßgeblich über elastische Stöße mit den Hüllenelektronen (vgl. Krieger, 2019). Die differentielle Abgabe der Energie E pro Weglänge x , auch genannt Bremsvermögen $S \equiv -(dE/dx)$, wird hierbei durch die genäherte Bethe-Bloch-Formel

$$S \approx k_0 \frac{z^2}{\beta^2} n \left[\ln \left(2m_e c^2 \frac{\beta^2}{1 - \beta^2} \right) - \beta^2 - \ln I \right] \quad (2.1)$$

beschrieben (Berger et al., 1993; Bethe, 1930), mittels welcher sich die Eindringtiefe und der typische Verlauf der Dosis im Gewebe entlang der Wegstrecke ableiten lassen. In Gleichung 2.1 bezeichnen z die Ladung des Projektils ($z = 1$ für Protonen), $\beta = v/c$ das Verhältnis zwischen Projektilgeschwindigkeit und Lichtgeschwindigkeit, n und I die Elektronendichte bzw. das mittlere Ionisationspotential des zu durchdringenden Materials, $m_e c^2$ die Ruheenergie des Elektrons und $k_0 = 5,1 \cdot 10^{-25} \text{ MeV cm}^2$ eine vom Material und Projektil unabhängige Konstante. In der Protonentherapie wird das Bremsvermögen relativ zu Wasser, engl. *stopping-power ratio* (SPR),

$$\frac{S}{S_{\text{Wasser}}} = \frac{n}{n_{\text{Wasser}}} \frac{\ln \left(2m_e c^2 \frac{\beta^2}{1-\beta^2} \right) - \beta^2 - \ln I}{\ln \left(2m_e c^2 \frac{\beta^2}{1-\beta^2} \right) - \beta^2 - \ln I_{\text{Wasser}}} = \hat{n} \hat{L} \quad (2.2)$$

als dimensionslose und näherungsweise von der Protonenenergie unabhängige Materialeigenschaft eingeführt. Man definiert die relative Elektronendichte $\hat{n} = n/n_{\text{Wasser}}$ und die relative Bremszahl \hat{L} . Es gibt verschiedene Verfahren für die Ermittlung der SPR-Verteilung im Patienten, welche im Unterkapitel 2.2 näher erläutert werden. Die Reichweite R im Patienten für Protonen mit der kinetischen Energie E_0 lässt sich durch die Näherung mittels kontinuierlicher Abbremsung, engl. *continuous slowing down approximation* (CSDA), bestimmen:

$$R(E_0) = \int_{E_0}^0 \frac{1}{S} dE. \quad (2.3)$$

Da die zugrunde liegenden Wechselwirkungen der Abbremsung statistisch variieren und damit auch die individuellen Streuwinkel der einzelnen Wechselwirkungen, kann die tatsächliche Reichweite eines Protons jedoch abweichen (*Reichweitenstraggling*). Außerdem wird durch die Vielzahl an Streuprozessen ein sonst paralleler Protonenstrahl beim Durchdringen von Materie senkrecht zur Strahlachse aufgeweitet (*Winkelstraggling*).

Neben der Ionisation der Atome wird ein kleiner Teil der Energie eines Protonenstrahls auch durch die selteneren inelastische Stöße mit den Atomkernen abgegeben. Solche Reaktionen führen zur Fragmentierung oder Anregung dieser Nuklide, was durch Methoden zur Überprüfung der Dosisverteilung oder Reichweite der Protonen im Gewebe ausgenutzt wird (vgl. Unterkapitel 2.3.2). Die Dosis D ist hierbei ein Maß für die absorbierte Energie pro Masse mit der Maßeinheit Gray ($1 \text{ Gy} = 1 \text{ J/kg}$). Die von den Protonen deponierte Dosis entlang ihrer Eindringtiefe ist durch ein ausgeprägtes Dosismaximum mit abruptem Abfall der Dosis geprägt (Abbildung 2.1 B). Dieser sogenannte Tiefendosisverlauf ergibt sich aus der Bethe-Bloch-Formel (Gleichung 2.1) aufgrund der Proportionalität $S \propto \beta^{-2}$ für die klinisch relevanten Protonenenergien und wird als Bragg-Kurve mit Bragg-Peak bezeichnet. Die Lage des Bragg-Peaks entlang der Eindringtiefe hängt dabei von der Energie der eingehenden Protonen ab. Aus dem Verlauf der Tiefendosis wird jedoch auch deutlich, dass eine

Verschiebung des Bragg-Peaks z. B. durch anatomische Veränderungen (Abbildung 2.1 C) oder Reichweiteunsicherheiten schwerwiegende dosimetrische Konsequenzen hat, wie in Unterkapitel 2.3.1 diskutiert.

Das Besondere der Protonenstrahlung ist somit, dass diese im Vergleich zur Photonenstrahlung eine endliche Reichweite im Gewebe besitzt. Durchdringen hingegen Photonen das Gewebe, so nimmt die Intensität der primären Strahlung exponentiell entsprechend des Absorptionskoeffizienten ab. Dies hat zur Folge, dass für tiefer liegende Tumore die Dosisdeposition geringer ist als für das davorliegende gesunde Gewebe (Abbildung 2.1 B). Nur durch eine auf den Tumor fokussierte Überlagerung von verschiedenen Einstrahlrichtungen kann erreicht werden, dass im Tumor mehr Dosis als im Normalgewebe appliziert wird. Im Vergleich dazu reicht in der Protonentherapie bereits eine Einstrahlrichtung aus, um mit einem ausgedehnten Bragg-Peak, engl. *spread-out Bragg-peak* (SOBP), durch Überlagerung verschiedener Bragg-Kurven eine höhere Dosis im Tumor als im gesunden Gewebe zu erreichen. In der Praxis wird jedoch auch in der Protonentherapie mit

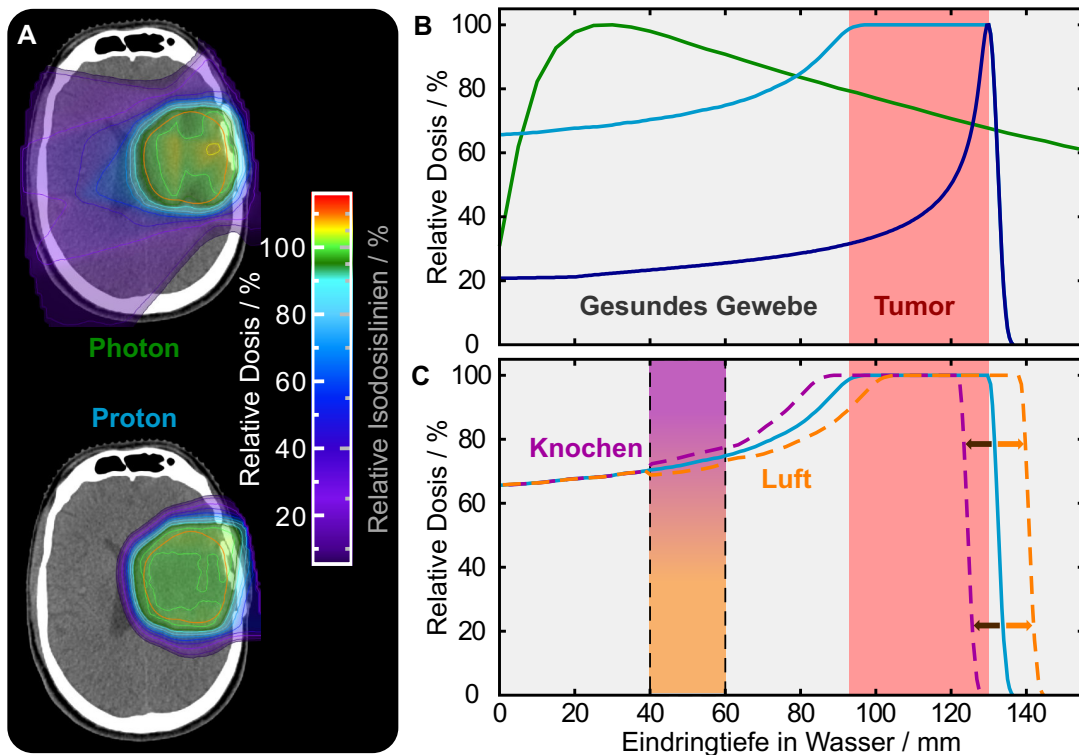


Abbildung 2.1: (A) Dosisverteilung in einer exemplarischen transversalen CT-Schicht eines Hirn-Tumor-Patienten für einen Photonen- und Protonen-basierten Bestrahlungsplan. (B) Tiefendosisverteilung von 15 MV Photonen (grün), 135 MeV Protonen (dunkelblau) und eines an das Tumervolumen (rot) angepassten ausgedehnten Bragg-Peaks (hellblau). (C) Variation des proximal zum Tumor liegenden Gewebes (orange/violett), welche zur Verringerung (höhere Dichte, z. B. durch Knochen) oder Erhöhung (niedrigere Dichte, z. B. durch Luft) der Protonenreichweite und damit zur kritischen Dosisverringern (-erhöhung) im Tumor (gesunden Gewebe) führt. (Quelle: In Anlehnung an Wohlfahrt, 2018)

mehreren Einstrahlrichtungen geplant, um Unsicherheiten zu berücksichtigen und Risikoorgane, engl. *organs at risk* (OARs), optimal zu schützen (Knopf & Lomax, 2013).

2.2 Behandlungsablauf in der Protonentherapie

2.2.1 Bildgebung zur Therapieplanung

Nachdem ein Patient für die Behandlung mit Protonen ausgewählt wurde, beginnt die Planung der Bestrahlung, wofür dreidimensionale Bildgebung des Patienten zum Einsatz kommt (vgl. Jäkel, 2018a). Die Bildgebung dient dabei zum einen der Lokalisation des Tumors und eventueller Metastasen sowie zur Abgrenzung der umliegenden OARs. Zum anderen wird aus der Bildgebung der für die Dosisberechnung relevante Gewebeparameter, der SPR-Wert, lokal ermittelt. Neben der Magnetresonanztomographie (MRT), die u. a. einen hohen Weichgewebekонтраст zur Gewebeabgrenzung liefert, und der Positronenemissionstomographie (PET), welche für funktionelle Bildgebung genutzt werden kann, ist die Computertomographie (CT) der aktuelle Maßstab für die quantitative Bildgebung des Patienten zur Bestrahlungsplanung. Bereits einige Tage vor der ersten Bestrahlung wird von jedem Patienten ein initialer CT-Scan, der sogenannte Planungs-CT-(pCT-)Scan, aufgenommen. Der pCT-Scan dient neben der Gewebeabgrenzung auch zur Bestimmung der SPR-Verteilung im Patienten, auf welcher die Dosisapplikation im Patienten geplant wird. Demgegenüber steht die Bestimmung der SPR-Verteilung im Patient mittels Protonen-Computertomographie-(Protonen-CT-)Scan, welche von technischer Seite her für die klinische Anwendung jedoch noch nicht ausgereift ist (Johnson, 2017). Zum einen sind die aktuellen klinischen Beschleunigeranlagen nicht darauf ausgelegt, Protonenenergien zu liefern, die den Patienten in jedem Fall durchdringen und zum anderen ist die räumliche Auflösung der Methode durch die Coulomb-Streuung limitiert. Unter Berücksichtigung neuerer Entwicklungen zur CT-basierten SPR-Vorhersage (siehe unten) ist noch nicht abzusehen, dass die Protonen-CT-Bildgebung überhaupt Einzug in die klinischen Abläufe der Protonentherapie erhalten wird (Johnson, 2017), wenngleich unter Laborbedingungen bereits eine ähnlich gute SPR-Genauigkeit in Phantomen erzielt werden konnte (Dedes et al., 2019).

Grundlagen der CT-basierten Reichweitevorschau

Die CT-Bildgebung ist ein etablierter Maßstab für die Bestrahlungsplanung in der konventionellen Strahlentherapie. Für den Einsatz von CT-Bildgebung für die Bestrahlungsplanung in der Protonentherapie gilt es Unsicherheiten abzuschätzen, nicht zuletzt auch aufgrund der Unterschiede in den Wechselwirkungen von Photonen und Protonen mit Gewebe. Weithin verbreitet für die Vorhersage der Protonenreichweite ist die Verwendung klassischer CT-Scans, d. h. ein mittels eines Röntgenspektrums aufgenommener CT-Scan, engl. *single-energy CT* (SECT). Vor einigen Jahren wurde ein weiteres CT-Verfahren zur klinischen

Anwendung gebracht: Ein mittels zweier Röntgenspektren aufgenommener CT-Scan, engl. *dual-energy CT* (DECT), ermöglicht es, eine verbesserte Vorhersage der SPR-Verteilung und damit der Protonenreichweite abzuleiten (Wohlfahrt & Richter, 2020). Im Rahmen dieser Arbeit wurden verschiedene Ansätze zur CT-basierten Reichweitevorschau im Patienten gegenübergestellt (Abbildung 2.2), die im Folgenden kurz näher erklärt werden.

Heuristischer Ansatz Die beim SECT-Scan gemessene Gewebeeigenschaft pro Volumenelement, den sogenannten Voxeln, ist der vom Röntgenspektrum abhängige Absorptionskoeffizient μ . Gemessen wird dieser auf Basis der Hounsfield Einheit, engl. *Hounsfield unit* (HU), welche eine lineare Transformation des Absorptionskoeffizienten zur CT-Zahl H mittels

$$H = 1000 \cdot \frac{\mu - \mu_{\text{Wasser}}}{\mu_{\text{Wasser}} - \mu_{\text{Luft}}} \quad (2.4)$$

darstellt. Dabei repräsentieren μ_{Wasser} und μ_{Luft} den mit dem jeweiligen CT-Scanner gemessenen Absorptionskoeffizienten von Wasser bzw. Luft unter Standardbedingungen (Fani et al., 2014). Aus dieser Definition ergibt sich, dass die CT-Zahl von Wasser $H_{\text{Wasser}} = 0$ HU und die von Luft $H_{\text{Luft}} = -1000$ HU ist. Die CT-Zahl für DECT-Scans kann aus der gewichteten Überlagerung der beiden Scans berechnet werden. In dieser Arbeit wurde die Anwendung syngo.CT DE Monoenergetic Plus (syngo.MonoPlus) als Teil der Bildnachverarbeitungssoftware syngo.via (Siemens Healthineers, Forchheim, Deutschland) zur Erstellung von 79 keV pseudo-monoenergetischen CT (monoCT) Bilddaten (aus Scans mit 80/140 kVp Röhrenspannung) eingesetzt. Für diese Überlagerung ist das Bildrauschen minimal und der Einfluss der Strahlauhfärtung im CT reduziert (Wohlfahrt et al., 2017a).

Die voxelweise Umrechnung der CT-Zahl in einen SPR-Wert erfolgt standardmäßig über eine Hounsfield Lookup-Tabelle (HLUT) und kann sowohl für SECT- als auch DECT-Scans angewendet werden (Wohlfahrt et al., 2017a). Die Zuordnung von CT-Zahl zu SPR ist jedoch nicht eindeutig, d. h. Materialien können die gleiche CT-Zahl, aber unterschiedliche SPR-Werte besitzen. Daher hat unter anderem die Auswahl der Materialien zur Kalibrierung einen Einfluss auf die Genauigkeit der SPR-Vorhersage. Diese kann sich zwischen Instituten teilweise erheblich unterscheiden, wie eine kürzlich erschienene Studie zeigt (Peters et al., 2021). Zur Kalibrierung der HLUT gibt es verschiedene Möglichkeiten. In dieser Arbeit wurde eine empirische Kalibrierung (Jäkel et al., 2001) für die 79 keV monoCT-Scans verwendet, welche im Folgenden als *Standard-HLUT* bezeichnet wird (Wohlfahrt et al., 2017a).

Eine unmittelbare Kalibrierung der HLUT anhand von Gewebeproben, sei es vom Menschen oder von Tieren, ist aufgrund des schnellen Gewebeerfalls der Proben nur schwer experimentell umzusetzen. Hingegen können Gewebeproben unter anderem zur Validierung und Anpassung einer bestehenden HLUT verwendet werden (Meijers et al., 2020). Eine weitere Möglichkeit zur Validierung und Anpassung einer HLUT beruht auf dem Verfahren zur direkten Bestimmung von SPR-Werten mittels DECT (DirectSPR), wie unten beschrieben. Dabei

wird in DECT-Scans von Patienten retrospektiv der Zusammenhang zwischen der CT-Zahl im monoCT-Scan und den mit DirectSPR bestimmten SPR-Werten pro Voxel ermittelt, welcher als Grundlage für die Anpassung der HLUT dient. In der vorliegenden Arbeit wurde eine solche *adaptierte HLUT* genutzt, die auf der Anpassung der HLUT anhand der jeweiligen SPR-Mediane pro CT-Zahl von insgesamt 50 Patienten-Scans (25 Hirn-Tumor- und 25 Prostata-Tumor-Patienten) basiert (Wohlfahrt et al., 2020).

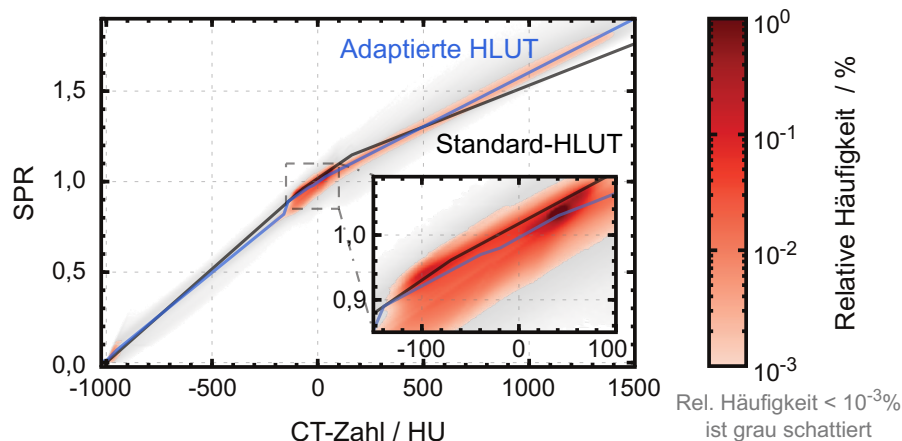


Abbildung 2.2: Überblick über die verwendeten Methoden zur CT-basierten SPR-Übersetzung. Als Grundlage der CT-Zahl dient ein mittels DECT bestimmter 79 keV monoCT-Scan, auf welchem eine Standard-HLUT empirisch kalibriert wurde (schwarz). Durch die Anwendung von DirectSPR konnte die Häufigkeitsverteilung der SPR-Werte gegenüber der CT-Zahl für insgesamt 50 Patienten bestimmt werden (rote Farbskala). Zuletzt wurde auf Grundlage dieser Häufigkeitsverteilung eine adaptierte HLUT definiert (blau). (Quelle: In Anlehnung an Wohlfahrt et al., 2019)

DirectSPR Bei der DirectSPR-Methode wird jedem Voxel ein individueller SPR-Wert zugeordnet, der sich wie folgt aus einem DECT-Scan ermitteln lässt. Aus der Definition des relativen Bremsvermögens (Gleichung 2.2) zeigt sich dessen Abhängigkeit von der relativen Elektronendichte \hat{n} und der relativen Bremszahl \hat{L} . Diese beiden Parameter können für jedes Voxel nach umfangreicher Kalibrierung aus einem DECT-Scan extrahiert werden. Für weitere Informationen zur Kalibrierung und Implementierung des in dieser Arbeit verwendeten DirectSPR-Verfahrens, welches gemeinsam mit Siemens Healthineers (Forchheim, Deutschland) entwickelt wurde, wird auf die Arbeit von Peters et al. (2022) verwiesen.

2.2.2 Bestrahlungsplanung

Für die Planung oder Anpassung der Dosisverteilung ist die Abgrenzung des Tumervolumens im pCT-Scan von Bedeutung, die ggf. unter Einbeziehung der in Unterkapitel 2.2.1 genannten Bildgebungsmodalitäten erfolgt (Troost et al., 2015). Man unterscheidet das sichtbare makroskopische Tumervolumen, engl. *gross tumor volume* (GTV), vom klinischen

Zielvolumen, engl. *clinical target volume* (CTV), welches die vermutete mikroskopische Proliferation von Tumorzellen angrenzend zum GTV mit einbezieht (DeLuca, 2007; Landberg et al., 1993, 1999). Bei der softwarebasierten inversen Bestrahlungsplanung wird, unter Berücksichtigung der jeweils anlagenspezifischen Strahleigenschaften, die Verteilung der verschriebenen Dosis für das CTV im Patienten optimiert und berechnet. Dafür werden anatomische und physikalische Unsicherheiten der Bestrahlung mit einbezogen (Unterkapitel 2.3.1). Das aus der konventionellen Strahlentherapie stammende Konzept mittels Planungszielvolumen, engl. *planning target volume* (PTV), als Sicherheitsaum um das CTV, stößt in der Protonentherapie jedoch auf einige Grenzen, wodurch die Methode der robusten Optimierung entwickelt wurde (Unkelbach et al., 2018). Bei der robusten Optimierung wird die Dosisverteilung so geplant, dass unter definierten Unsicherheitsquellen – etwa in der Patientenpositionierung oder der Vorhersage der Protonenreichweite – weiterhin die vom Arzt verschriebene Dosis im CTV appliziert wird, sodass die explizite Definition eines PTV nicht mehr notwendig ist (Unkelbach & Paganetti, 2018).

Bei der Bestrahlungsplanung wird zudem auch die Dosisdeposition in OARs nach klinikinternen Richtlinien berücksichtigt, um mögliche Normalgewebskomplikationen zu reduzieren (DeLuca, 2007). Solche Richtlinien können sich z. B. auf die maximale oder mittlere Dosis in einem bestimmten Organ oder Volumenanteil beziehen. Dargestellt wird der Zusammenhang zwischen Dosis und Volumen in einem Dosis-Volumen-Histogramm (DVH), wie beispielhaft für einen Prostata-Patienten in Abbildung 2.3 gezeigt. In dieser Arbeit kam die Bestrahlungsplanungssoftware RayStation (RaySearch Laboratories AB, Stockholm, Schweden) mit robuster Optimierung in den Versionen 6 und 8B für die klinische Planung zum Einsatz. Für den Zweck dieser Arbeit, wurde eine Monte-Carlo-basierte Dosisberechnung noch einmal einheitlich in der RayStation Forschungsversion 9B mit einer Voxelgröße von $1 \times 1 \times 1 \text{ mm}^3$ und der klinisch gängigen Unsicherheit von 0,3% durchgeführt.

Um Komplikationen im Normalgewebe durch die Bestrahlung zu vermeiden, wird die Dosis sukzessive über einen längeren Zeitraum appliziert, was in der Strahlentherapie als Fraktionierung bezeichnet wird. Die Grundlage für die Fraktionierung bildet dabei das gegenüber dem Tumorgewebe meist ausgeprägtere Reperaturvermögen des Normalgewebes in Bestrahlungspausen. Dadurch lassen sich mit der fraktionierten Bestrahlung bei gleicher Tumorkontrolle vor allem Spätreaktionen des Normalgewebes verhindern, d. h. also, dass sich das therapeutische Fenster der Behandlung vergrößert (vgl. Karger, 2018). Von Standardfraktionierung spricht man bei einer täglichen Einzeldosis von 1,8–2 Gy mit fünf Fraktionen pro Woche und Gesamtdosen von typischerweise 40–70 Gy, was je nach Tumorart einem Bestrahlungszeitraum von 3 bis 7 Wochen entspricht (vgl. Karger, 2018). Je nach Tumorart kann es aber auch gerechtfertigt sein, eine geringere oder höhere Einzeldosis zu verwenden. Bei gleichbleibender Gesamtdosis für das Tumolvolumen bedeutet das, die Anzahl der Fraktionen entsprechend zu erhöhen oder zu verringern (Hyper- oder Hypofraktionierung).

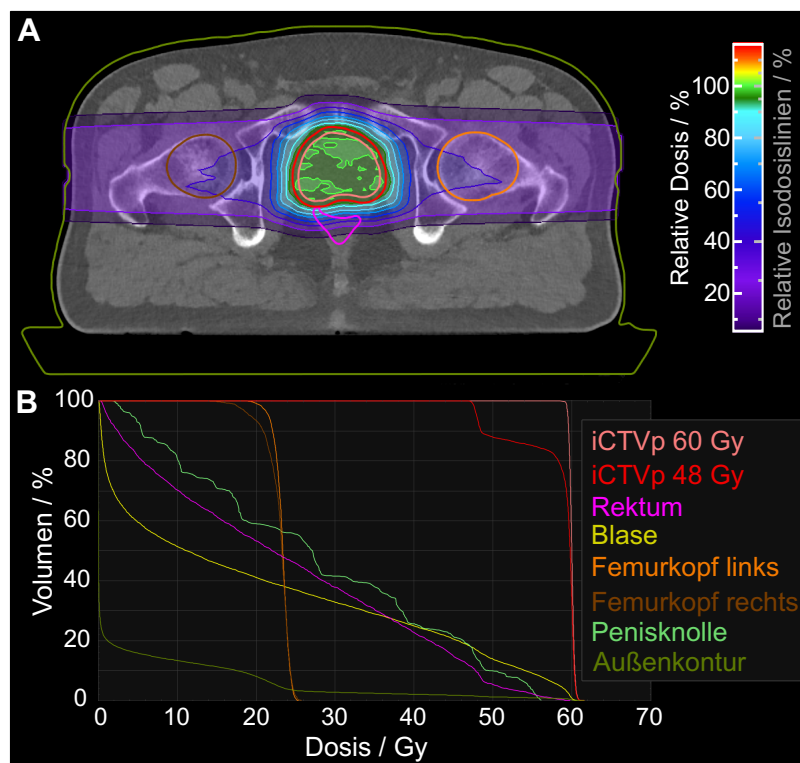


Abbildung 2.3: (A) Beispiel für eine bei Prostata-Tumor-Patienten typische Dosisverteilung mit zwei lateralen Feldern im transversalen Schnitt. (B) Resultierendes Dosis-Volumen-Histogramm für die primären internen klinischen Zielvolumina (iCTVp) und Risikoorgane bei der Prostatabestrahlung.

2.2.3 Strahlanwendung

Protonen für die Protonentherapie werden durch Ionisation von Wasserstoffatomen erzeugt und durch eine Beschleunigeranlage auf die für die Therapie benötigten Energien beschleunigt. Die beiden wesentlichen Beschleunigertypen, die hierbei zum Einsatz kommen, sind relativistische Zyklotrone und Synchrotrone (vgl. Schippers, 2012). Als Synchrotron bezeichnet man einen Ringbeschleuniger, bei dem die Protonen durch Ablenkmagneten auf einer geschlossenen Bahn gehalten werden und durch im Strahlengang befindliche Beschleunigungskavitäten ein oder mehrere Bündel von Protonen gleichzeitig beschleunigt werden können. Die Hochfrequenz zur Beschleunigung der Protonen sowie die Feldstärke der Ablenkmagnete werden hierbei mit der Protonengeschwindigkeit synchronisiert. Zyklotrone sind kompakte Ringbeschleuniger, in denen die Protonen durch ein ständiges Magnetfeld auf eine Kreisbahn gezwungen werden. Die Beschleunigung erfolgt typischerweise zwischen zwei oder vier Metallelektroden, den sogenannten Duanten, an denen eine Hochfrequenz anliegt. Da die Protonen durch die Beschleunigung eine Spiralbahn durchlaufen, ist das Zyklotron jedoch nur für die Extraktionen einer immer gleichbleibenden Protonenenergie konzipiert, sodass die Energiemodulation für die Therapie über ein separates Energieabstufungs- und Energieauswahlsystem zu erfolgen hat. Für die klinisch relevante Extraktionsenergie der Protonen (etwa 230-250 MeV) ist weiterhin die relativistische Massenzunahme zu berücksichtigen.

sichtigen. Beschleuniger, die diese kompensieren, sind zum einen das Synchrozyklotron, bei dem die Beschleunigungsfrequenz mit der zunehmenden Umlaufdauer des Protonenbündels synchronisiert wird. Zum anderen werden Isochronzyklotrone genutzt, bei denen das Magnetfeld in radialer Richtung zunimmt und für eine gleichbleibende Umlaufdauer der Protonen sorgt, wodurch – wie beim klassischen Zyklotron – eine quasi kontinuierliche Extraktion von Protonen ermöglicht wird. Die in dieser Arbeit ausgewerteten Bestrahlungen an einem Isochronzyklotron wurden mit einem Proteus235 System (Ion Beam Applications S.A., Louvain-La-Neuve, Belgien) am UPTD durchgeführt.

Über eine evakuierte Strahlführung mit Führungs- und Fokussierungsmagneten werden die Protonen vom Beschleuniger in den Behandlungsraum der Patienten geleitet. Mit einer 360°-rotierbaren Gantry kann dort im Vergleich zu einer stationären Strahlführung das maximale Potential der Protonentherapie genutzt werden, denn die Gantry ermöglicht in Kombination mit der Patientenliege die Bestrahlung aus beliebigen Einstrahlrichtungen und damit die individuell optimale Schonung des gesunden Gewebes. Abhängig vom verwendeten Bestrahlungsverfahren wird der Strahl unmittelbar vor oder am Austritt aus der Strahlführung konfiguriert. Es werden grundlegend zwei Verfahren unterschieden: das passive Streuverfahren und das dynamische Scanning-Verfahren (DeLuca, 2007). Im Folgenden sollen exemplarisch dafür das passive Doppelstreuverfahren bzw. das diskrete Scanning-Verfahren mit Bleistiftstrahl, engl. *pencil-beam scanning* (PBS), vorgestellt werden – beide Verfahren werden standardmäßig an der UPTD eingesetzt.

Doppelstreuverfahren

In diesem Verfahren wird der meist nur einige Millimeter ausgedehnte Protonenstrahl durch zwei dedizierte Streuelemente aufgeweitet (Abbildung 2.4 A). Das erste Streuelement dient dazu, den Strahl zu verbreitern, indem ein Material mit hoher Kernladungszahl Z zum Einsatz kommt. Für die Bestrahlung ist hingegen ein flaches laterales, d. h. senkrecht zur Strahlrichtung, Strahlprofil von Vorteil, um den Tumor homogen mit Dosis abdecken zu können. Das zweite Streuelement besteht daher aus einer Kombination eines Hoch- und Niedrig- Z -Materials mit unterschiedlichen Aufstreueigenschaften, sodass das gaußförmige Strahlprofil geglättet wird, ohne dabei jedoch eine unterschiedliche Energieverteilung der Protonen in lateraler Richtung zu erzeugen. Die Eindringtiefe des Protonenstrahls kann mit einem sich drehenden Rad unterschiedlich dicker Bereiche moduliert werden, aber auch eine stationäre Lösung mittels sogenanntem Ridgefilter ist möglich. Der lateral und in der Eindringtiefe aufgestreute Protonenstrahl durchläuft kurz vor dem Patienten noch lateral formende Kollimatoren und einen patientenspezifischen Reichweitekompensator, welcher den Strahl konformal an die in Strahlrichtung distale Form des Tumors anpasst. Ein Nachteil dieses Verfahrens ist daher, dass die in Strahlrichtung proximale Form des Tumors nicht berücksichtigt werden kann und somit vergleichsweise viel Dosis in Teilen des proximalen gesunden Gewebes appliziert wird (Abbildung 2.4 A). Ein Vorteil ist hingegen, dass das komplette Tumolvolumen

quasi gleichzeitig und in kurzer Zeit bestrahlt werden kann, was den Einfluss von Bewegung auf die Bestrahlung (z. B. bei Lungentumoren) deutlich reduziert.

Pencil-Beam Scanning

Beim dynamischen Scanning wird der bleistiftdicke Protonenstrahl mithilfe von zwei senkrecht zum Strahl angeordneten Ablenkmagneten in lateraler Richtung über den Tumor bewegt. Das Tumolvolumen wird hierbei in Schichten gleicher Protonenenergie aufgeteilt, den sogenannten *Layern*, und im Falle des diskreten PBS pro Layer weiter in Bestrahlungspunkte, den sogenannten *Spots* (Abbildung 2.4 B). Dadurch kann die Form des Tumors optimal berücksichtigt werden, sodass das proximale Gewebe im Vergleich zum passiven Doppelstreuverfahren besser geschont werden kann. Bei der passiven Strahlapplikation sind hingegen steile laterale Gradienten des Bestrahlungsfeldes hervorzuheben, die im PBS durch den Querschnitt bzw. die Penumbra der Spots beschränkt sind. Weiterhin erfolgt im PBS die Modulation der Eindringtiefe und der lateralen Ausdehnung des Bestrahlungsfeldes zeitlich klar getrennt, wodurch sich die Dauer der Bestrahlung verlängert. Die Zeitabstände zwischen den Spots liegen nur in der Größenordnung von Millisekunden, doch insbesondere durch die mechanische Bewegung von Bauteilen in der Strahlführung zum Anpassen der Protonenenergie benötigt der Wechsel der Layer meist mehr als eine Sekunde. Dadurch kann sich die Bewegung des Patienten (z. B. durch Atmung, Herzschlag, Peristaltik) nachteilig auf die

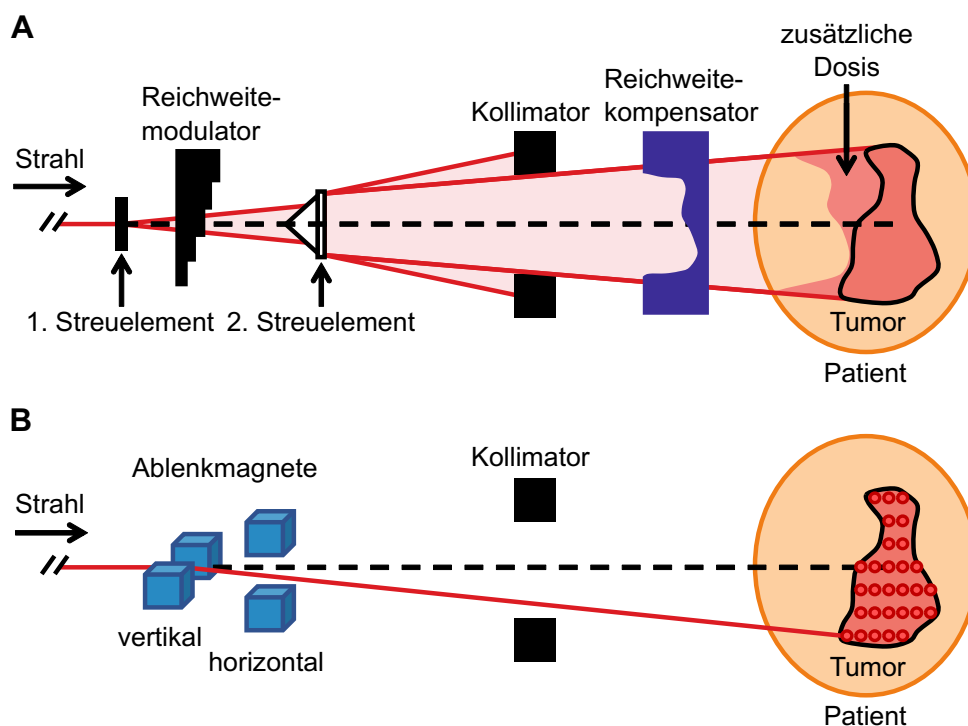


Abbildung 2.4: Schematische Darstellung der Funktionsweise des Doppelstreuverfahrens (A) und des PBS-Verfahrens (B). (Quelle: In Anlehnung an Schlegel et al., 2018)

applizierte Dosis auswirken. Es gibt jedoch einige Ansätze, die das Auftreten von Bereichen mit extrem hoher und niedriger Dosis, welche durch die Überlagerung der Bestrahlung mit regelmäßigen Patientenbewegungen entstehen, kompensieren. Beispiele hierfür sind die wiederholte Abstrahlung des Plans mit jeweils reduzierter Dosis (engl. *repainting*) oder die zeitliche Synchronisation der Bestrahlung mit der Bewegung (engl. *gating*) (vgl. Schlegel et al., 2018). Im Gegensatz dazu liegen weitere Vorteile des PBS darin, dass keine patientenspezifischen Bauteile (Apertur und Reichweitekompensator) gefertigt werden müssen, dass die Dosisbelastung durch Streustrahlung (z. B. Neutronen) reduziert wird und dass die relative Protonenausbeute für die Therapie größer ist.

2.3 Genauigkeit in der Protonentherapie

Der Zusammenhang zwischen Reichweite und Dosisdeposition der Protonen in der Protonentherapie verdeutlicht, dass eine Reichweiteabweichung schwerwiegende dosimetrische Konsequenzen nach sich ziehen kann (vgl. Abbildung 2.1 C). Der Behandlungserfolg hängt also maßgeblich von der Berücksichtigung und Begrenzung aller Unsicherheiten entlang des im vorherigen Abschnitt beschriebenen Behandlungsablaufs ab. In diesem Abschnitt sollen die wichtigsten Quellen für Unsicherheiten und der praktische Umgang mit ihnen vorgestellt werden. Hieran anknüpfend wird der aktuelle Stand von Methoden zur *in vivo* Reichweiteverifikation vorgestellt.

2.3.1 Ursachen für Behandlungs- und Reichweiteunsicherheiten

Anatomische Veränderungen des Patienten

Über den Verlauf der Therapie muss die Variabilität der Patientenanatomie berücksichtigt werden. Man unterscheidet anatomischen Veränderungen, die innerhalb einer Bestrahlungsfraktion (intrafraktionell) auftreten und solche zwischen verschiedenen Bestrahlungsfraktionen (interfraktionell). Beispiele für intrafraktionelle Veränderungen sind Änderungen aufgrund der Atmung oder des Herzschlags des Patienten, aber auch aperiodische Prozesse bspw. im Magen-Darm-Trakt. Diese werden nach Möglichkeit bereits in der Bestrahlungsplanung oder der Strahlapplikation berücksichtigt (Albertini et al., 2020; De Ruyscher et al., 2015). Auch interfraktionelle Veränderungen, wie z. B. die Füllung von Kavitäten und Hohlräumen oder auch Tumorschrumpfung und Gewichtsveränderung des Patienten, können schwerwiegende Folgen auf die Dosisverteilung im Patienten und somit die Tumorkontrolle und die Wahrscheinlichkeit für Normalgewebekomplikationen haben (Paganetti et al., 2021b). Kleinere Änderungen sind im Rahmen von Sicherheitssäumen abgedeckt, darüber hinausgehende relevante klinische Veränderungen führen dazu, dass der Bestrahlungsplan angepasst werden muss (Planadaptation). Die Herausforderung ist es, schwer vorhersehbare

anatomische Veränderungen rechtzeitig zu erkennen, um darauf reagieren zu können. Hierfür könnten zukünftig neben klassischen Bildgebungsmodalitäten wie regelmäßigen CT- oder MRT-Scans auch Systeme zur *in vivo* Dosis- und Reichweiteverifikation eine wesentliche Rolle spielen. Unter adaptiver Protonentherapie fasst man daran anknüpfend die Anpassung des Bestrahlungsplans an die aktuelle Patientenanatomie – optimalerweise in Echtzeit (online) – zusammen (Albertini et al., 2020; Paganetti et al., 2021b). Das Ziel der online-adaptiven Protonentherapie (OAPT) ist die Entwicklung eines Kreislaufs aus Bildgebung, Planadaptation und Verifikation mit geeigneten Methoden zur Qualitätssicherung der Bestrahlung.

Positionierung und Strahlanwendung

Durch die fraktionsweise Bestrahlung ist es notwendig, dass der Patient zu jeder Fraktion exakt so im Isozentrum des Bestrahlungsgeräts (virtueller Punkt im Raum, in dem sich die Drehachsen von Gantry, Kollimator und Tisch kreuzen) positioniert wird, wie durch die Bestrahlungsplanung auf dem pCT-Scan vorgegeben. Der Patient wird hierbei meist anhand von Lasern und planarer Röntgenbildgebung, einer Kegelstrahl-Computertomographie, engl. *cone-beam CT* (CBCT), oder mittels eines CT-Scanners auf Schienen, engl. *CT on rails*, im Isozentrum positioniert (vgl. Jäkel, 2018a). Zur Erhöhung der Präzision in der Positionierung werden die Patienten oftmals z. B. mithilfe von Vakuumkissen oder Kopfmaske immobilisiert (vgl. Echner, 2018). In Abhängigkeit der jeweils verwendeten Methoden und der Tumorentität können Unsicherheiten in der Positionierung abgeschätzt und bereits in der Bestrahlungsplanung berücksichtigt werden. Weiterhin wird die Einhaltung von vorher festgelegten Toleranzen für die Strahlanwendung, d. h. Lage des Isozentrums, Strahlenergie und Strahlprofil, in der Regel täglich bis turnusmäßig überprüft (vgl. Jäkel, 2018b).

CT-basierte Reichweitevorschau

Die wichtigste Grundlage für die Berechnung der Protonenreichweite ist die CT-basierte Vorhersage der SPR-Verteilung (Unterkapitel 2.2.1), welche mit gewissen Unsicherheiten behaftet ist. Für das Standard-HLUT Verfahren wird typischerweise ein Sicherheitsraum mit einem relativen Anteil von 2,5–3,5% der Protonenreichweite und einem absoluten Anteil von 1–3 mm angenommen (Paganetti, 2012). Bei der Verwendung von DirectSPR in der Bestrahlungsplanung konnten die Unsicherheiten an der UPTD erstmals von 3,5%+2 mm auf 2,0%+2 mm für Prostatatumore und 1,7%+2 mm für Kopftumore reduziert werden (Peters et al., 2022).

Weitere Ursachen

Zu den weiteren Ursachen für Unsicherheiten der Behandlung gehört einerseits die Abgrenzung des Zielvolumens und der OARs (Persson et al., 2012). Andererseits gibt es Unsicherheiten in der Berechnung der physikalischen und biologischen Strahlendosis im Patienten.

Die Geschwindigkeit und kommerzielle Verfügbarkeit von Monte-Carlo-Algorithmen für die Dosisberechnung hat über die letzten Jahr deutlich zugenommen, weshalb bereits an vielen Kliniken – wie auch in dieser Studie an der UPTD – Monte-Carlo-basierte Dosisberechnungen standardmäßig zum Einsatz kommen (Fracchiolla et al., 2021). Gegenüber analytischen Modellen sind Monte-Carlo-basierte Modelle genauer, da sie insbesondere Streuprozesse besser simulieren (Yepes et al., 2018). Die erhöhte biologische Wirksamkeit (im Bezug auf Zellüberleben) der durch Protonen eingetragenen Dosis im Vergleich zu Photonen wird aktuell durch einen konstanten Wichtungsfaktor von 1,1, der sogenannten relativen biologischen Wirksamkeit (RBW), berücksichtigt (DeLuca, 2007). Aus *in vitro* Experimenten ist bekannt, dass es am Ende der Protonenreichweite zu einer höheren RBW von bis zu 1,7 kommen kann (Paganetti, 2014). Für die Einführung von Modellen mit variablem RBW in die Bestrahlungsplanung fehlt momentan die Datengrundlage, jedoch gibt es erste klinische Indizien für einen möglichen Zusammenhang zwischen Normalgewebekomplikationen und variablem RBW (Eulitz et al., 2023; Sørensen et al., 2021).

2.3.2 Aktueller Stand der Behandlungs- und Reichweiteverifikation

Ein wesentlicher Aspekt im Umgang mit Unsicherheiten ist die Verifikation der Behandlung im Patienten. Methoden zur Behandlungs- und Reichweiteverifikation könnten dazu beitragen, das Normalgewebe besser zu schonen (Knopf & Lomax, 2013). In dieser Arbeit wird ein solches System zur *in vivo* Reichweiteverifikation untersucht.

Es lässt sich grundsätzlich zwischen direkten und indirekten Methoden unterscheiden, um die Reichweite oder Reichweiteabweichungen der Protonen im Patienten zu bestimmen (Knopf & Lomax, 2013). Während bei direkten Methoden die Reichweite auf die unmittelbare Detektion von Protonen zurückgeführt werden kann, wird die Reichweite bei indirekten Methoden hingegen anhand von damit korrelierten, sekundären Signaturen bestimmt. Für das eigentliche Bestrahlungsfeld können die Protonen nicht zu Verifikationsmessungen außerhalb des Patienten herangezogen werden, da sie ihre komplette Energie bereits im Gewebe deponieren. Bei direkten Verfahren muss daher eine Messsonde in den Patienten eingebracht werden, was nur für bestimmte Tumorentitäten möglich ist und zudem einen invasiven Eingriff bedeutet (Knopf & Lomax, 2013). Genau wie die Detektion von langsam auftretenden Langzeit-Gewebeveränderungen im MRT (Knopf & Lomax, 2013) spielen solche Verfahren für die Echtzeit-Verifikation eine eher untergeordnete Rolle und sollen hier nicht näher betrachtet werden. Die indirekten Methoden fokussieren sich hauptsächlich auf drei unterschiedliche Arten von Signaturen: die Emission von Gammastrahlung aus induzierten Kernanregungen, die Emission von Annihilationsstrahlung von erzeugten Positronenemittern und die Aussendung thermoakustischer Signale entlang des Protonenstrahls. Im Folgenden werden darauf basierende Verifikationsmethoden vorgestellt. Außerdem wird auf die Methode der Protonenradiographie als eine weitere Möglichkeit zur direkten *in vivo*

Reichweiteverifikation eingegangen, welche jedoch keine unmittelbare Überprüfung des Behandlungsstrahls darstellt.

Prompte Gammastrahlung

Durch die Protonenbestrahlung werden angeregte Kerne induziert, welche sich mit einer Halbwertszeit von ≤ 1 ns unter Aussendung sogenannter *prompter Gammastrahlung (PG)* abregen können. Diese hochenergetische Strahlung ist im Wesentlichen gekennzeichnet durch ein breites Energiespektrum bis 7 MeV mit einigen charakteristischen Spektrallinien (Smeets et al., 2012). Ein Großteil der PG-Emission im Patienten stammt aus Kernabregungen der Isotope ^{12}C und ^{16}O (Tabelle 2.1).

Aufgrund der geringen Absorptionswahrscheinlichkeit von PG im Patienten kann diese für die Reichweiteverifikation genutzt werden, wie erstmals durch Jongen und Stichelbaut (2003) vorgeschlagen. Herkömmliche Gammastrahlungsdetektoren aus der Nuklearmedizin sind insbesondere durch die höhere Gammaenergie und Detektorlast, u. a. verursacht durch den Neutronenuntergrund, nicht für die Anforderungen in der Protonentherapie konzipiert (Smeets et al., 2012). Dies führte zur Entwicklung neuer Detektoren und zur Entdeckung der drei wesentlichen PG-basierten Detektionsprinzipien für den Zweck der Reichweiteverifikation in der Protonentherapie: Prompt-Gamma-Bildgebung, engl. *Prompt-Gamma-Imaging (PGI)*, Prompt-Gamma-Spektroskopie (PGS) sowie Prompt-Gamma-Timing (PGT) (vgl. Krimmer et al., 2018). Grundlage für all diese Methoden ist, dass die PG-Emission

Tabelle 2.1: Überblick wichtiger Spektrallinien von prompter Gammastrahlung aus nuklearen Wechselwirkungen mit ^{16}O und ^{12}C . Einige Spektrallinien mit ähnlicher Energie wurden zusammengefasst und deren Energie gemittelt. (Quelle: Verburg und Seco, 2014)

Targetkern ^{16}O		Targetkern ^{12}C	
Energie / MeV	Reaktion	Energie / MeV	Reaktion
1,89	$^{16}\text{O}(p, pp\gamma)^{15}\text{N}$	2,00	$^{12}\text{C}(p, x\gamma)^{11}\text{C}$
2,0	$^{16}\text{O}(p, x\gamma)^{15}\text{O}$	2,1	$^{12}\text{C}(p, pp\gamma)^{11}\text{B}$
	$^{16}\text{O}(p, x\gamma)^{11}\text{C}$		$^{12}\text{C}(p, x\gamma)^{10}\text{C}$
2,31	$^{16}\text{O}(p, x\gamma)^{14}\text{N}$	2,8	$^{12}\text{C}(p, x\gamma)^{11}\text{C}$
2,8	$^{16}\text{O}(p, p'\gamma)^{16}\text{O}$		$^{12}\text{C}(p, x\gamma)^{10}\text{B}$
	$^{16}\text{O}(p, x\gamma)^{14}\text{N}$	4,44	$^{12}\text{C}(p, p'\gamma)^{12}\text{C}$
	$^{16}\text{O}(p, x\gamma)^{11}\text{C}$	4,80	$^{12}\text{C}(p, x\gamma)^{11}\text{C}$
	$^{16}\text{O}(p, x\gamma)^{10}\text{B}$		
3,68	$^{16}\text{O}(p, x\gamma)^{13}\text{C}$		
4,44	$^{16}\text{O}(p, x\gamma)^{12}\text{C}$		
5,2	$^{16}\text{O}(p, x\gamma)^{15}\text{O}$		
	$^{16}\text{O}(p, pp\gamma)^{15}\text{N}$		
	$^{16}\text{O}(p, x\gamma)^{15}\text{O}$		
	$^{16}\text{O}(p, pp\gamma)^{15}\text{N}$		
6,32	$^{16}\text{O}(p, x\gamma)^{15}\text{N}$		
7,0	$^{16}\text{O}(p, p'\gamma)^{16}\text{O}$		
	$^{16}\text{O}(p, p'\gamma)^{16}\text{O}$		

deutlich mit der Protonenreichweite korreliert (Min et al., 2006), wobei etwa jedes zehnte Proton an einer nuklearen Wechselwirkung beteiligt ist, welche PG induziert (Smeets et al., 2012). Der Wirkungsquerschnitt dieser Reaktionen hängt von der Protonenenergie ab und führt zu einem ähnlichen Verlauf der PG-Emission über der Eindringtiefe (Abbildung 2.5) wie die Tiefendosisverteilung der Protonen.

Ändert sich also die Reichweite der Protonen, so hat dies auch direkt Auswirkungen auf den Intensitätsverlauf der PG entlang des Protonenstrahls. Ein wesentlicher Unterschied besteht allerdings darin, dass die Energieschwelle für nukleare Wechselwirkungen deutlich höher liegt als für Wechselwirkungen mit den Hüllenelektronen. Daher induzieren Protonen kurz vor ihrer vollständigen Abbremsung keine PG mehr. Dies stellt aber nur bei lokal großen Dichteunterschieden im Gewebe eine Einschränkung für die Korrelation der PG-Emission mit der Protonenreichweite dar.

Prompt-Gamma-Imaging Die PGI-Methode basiert auf der räumlichen Auflösung der PG-Emission. Für die Bestimmung der Protonenreichweite wird sich die örtliche Korrelation zwischen Dosis und PG-Emission zunutze gemacht. Es gibt zwei Möglichkeiten eine entsprechende räumliche Auflösung zu erreichen: mittels mechanischer oder elektronischer Kollimation (vgl. Krimmer et al., 2018). In dieser Arbeit wurde ein System mit einem *knife-edge*,

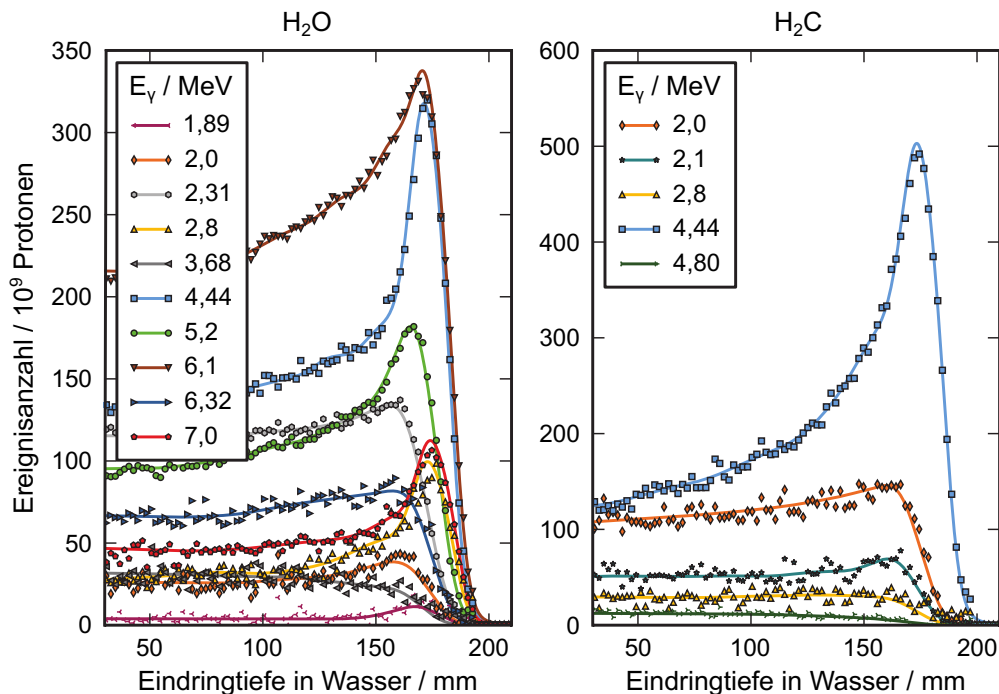


Abbildung 2.5: Gemessene (Datenpunkte) und simulierte (Linien) Emission spezifischer Spektrallinien prompter Gammastrahlung (vgl. Tabelle 2.1) über der Eindringtiefe eines 165 MeV Protonenstrahls (CSDA-Reichweite in Wasser: 18,6 cm) für ein Target aus Wasser (H₂O) und Polyethylen (H₂C). (Quelle: In Anlehnung an Verburg und Seco, 2014)

aus dem Engl. für keilförmig zulaufend, Schlitzkollimator (Bom et al., 2012; Smeets et al., 2012) während klinischer Protonenbestrahlungen verwendet, welches detaillierter in Kapitel 3 beschrieben wird. Weitere Beispiele für die Realisierung mechanischer Kollimation sind Lochblenden (Kim et al., 2009), sowie Mehrfach-Schlitz-Systeme (Min et al., 2012; Park et al., 2019). Elektronische Kollimation der PG lässt sich durch das Prinzip einer Compton-Kamera erreichen (Everett et al., 1977; Kabuki et al., 2009). Hierbei wird aufgrund der Compton-Streuung die Einfallsrichtung der detektierten Photonen rekonstruiert.

Prompt-Gamma-Spektroskopie Das Verhältnis der Wirkungsquerschnitte für die Anregung verschiedener Kernniveaus ändert sich entlang der Eindringtiefe der Protonen (Abbildung 2.5). Dadurch lässt sich in einer gewissen Eindringtiefe aus der Häufigkeit von detektierter PG zu unterschiedlichen Energien etwas über die Zusammensetzung des Gewebes und die restliche Reichweite der Protonen ableiten. Auf diesem Prinzip beruht die PGS-Methode, bei welcher ein schmaler Kollimator einen bestimmten Punkt entlang des Protonenstrahls im Patienten fokussiert, um an dieser Stelle die Energie und damit die verbleibende Reichweite der Protonen zu ermitteln (Verburg & Seco, 2014). Mithilfe eines ersten Prototyps (Hueso-González et al., 2018) wurde diese Methode bereits erfolgreich im Patienten angewandt (Verburg et al., 2020). In dieser Arbeit wird ein Vergleich zwischen diesem Prototypen und dem untersuchten PGI-System unter klinischen Bedingungen beschrieben (Kapitel 6).

Prompt-Gamma-Timing Die Methoden PGT und Prompt-Gamma-Peak-Integral (PGPI) beruhen auf der Messung der Flugzeit der Protonen im Gewebe, welche mit der Eindringtiefe korreliert ist (Golnik et al., 2014). Wie beschrieben wird PG entlang der Wegstrecke der Protonen emittiert und entsprechend kann mit einem unkollimierten System mit geeigneter Zeitauflösung die Ankunftszeit von PG entlang des Protonenstrahls detektiert werden. Um die Flugzeit der Protonen zu ermitteln wird eine Referenzzeit benötigt (häufig einfach die Zyklotronfrequenz), die anschaulich z. B. dem Eintritt der Protonen in den Patienten entspricht. Die Verteilung der Flugzeiten verbreitert sich, wenn die Protonen eine größere Flugstrecke im Patienten zurücklegen, also eine größere Reichweite besitzen. Entsprechend lässt sich aus Veränderungen der Breite oder des Mittelwerts der Verteilung eine Reichweiteabweichung ermitteln (Golnik et al., 2014). Zusätzlich zur Breite der Verteilung ändert sich auch das Integral der Verteilung aufgrund der abweichenden Winkelverteilung von PG am Detektor. Dadurch können prinzipiell auch Reichweiteabweichungen bestimmt werden, was unter dem Begriff PGPI bekannt ist (Krimmer et al., 2017). Mithilfe von mehreren Detektoren lassen sich darüber potentiell auch Abweichungen in der Patientenpositionierung und Dosisapplikation ableiten (Krimmer et al., 2017). Die PGT-Methode befindet sich momentan noch in der Translation zur klinischen Anwendung (Hueso-González et al., 2015; Werner et al., 2019).

Positronenemissionstomographie

Der indirekte Zusammenhang zwischen Dosisdeposition und dem Entstehen von Positronemittern durch Protonenstrahlung stellt eine Möglichkeit dar, die Protonenbestrahlung mit PET zu verifizieren (Zhu & Fakhri, 2013). Die Isotope ^{11}C , ^{13}N und ^{15}O haben aufgrund der Zusammensetzung menschlichen Gewebes den größten Einfluss auf das PET-Signal. Diese Betastrahler zerfallen mit einer charakteristischen Halbwertszeit von 2–20 min unter Aussendung eines Positrons. Dieses annihiliert mit einem Elektron im Gewebe zu zwei entgegengesetzt gerichteten Photonen mit einer Energie von je 511 keV, die mit PET detektiert werden können. Jedoch bewirkt der menschliche Metabolismus durch Perfusion und Diffusion im Gewebe, dass die gemessene Aktivität nicht mehr unmittelbar mit dem Entstehungsort, also dem Ort der Dosisdeposition korreliert ist. Diese biologische Verwaschung, engl. *biological washout*, und die freie Weglänge des Positrons im Gewebe limitieren die Genauigkeit der Methode (Zhu & Fakhri, 2013). Durch eine sofortige Messung mit einem PET-Scanner am Bestrahlungsgerät, engl. *in-beam* PET, kann die Genauigkeit der Methode verbessert werden (Shakirin et al., 2011). Erstmals wurde ein solches in-beam PET am GSI Helmholtzzentrum für Schwerionenforschung (Darmstadt, Deutschland) klinisch implementiert (Enghardt et al., 2004). Mit einem der aktuellen, klinisch erprobten Prototypen (Ferrero et al., 2018) konnten morphologische Veränderungen im Patienten den jeweils betreffenden PBS-Spots (Fiorina et al., 2021) und sogar voxelweise zugeordnet werden (Kraan et al., 2022).

Ionoakustik

Als Ionoakustik bezeichnet man eine Methode zur Reichweiteverifikation anhand von akustischen Signalen, die aufgrund der lokalen Erwärmung und der daraus resultierenden Druckerhöhung im Gewebe entlang eines Ionenstrahls entstehen (Assmann & Parodi, 2020; Hayakawa et al., 1995). Um Reichweiteabweichungen zu detektieren wird die Laufzeit der tatsächlich während der Bestrahlung gemessenen ionoakustischen Signale mit einer theoretischen Erwartung abgeglichen. Die Messung der ionoakustischen Signale erfolgt nach Propagation der Druckwelle durch den Körper an einem geeigneten Sensor für Ultraschallwellen. Aus den Signalen kann die Lage des Bragg-Peaks berechnet und bei Kombination mit einem live Ultraschallbild in diesem markiert werden (Kellnberger et al., 2016; Patch et al., 2016). Bislang konnte in homogenen Wasserphantomen eine Auflösung der Reichweite im Submillimeterbereich nachgewiesen werden (Assmann et al., 2015; Lehrack et al., 2017). Herausforderungen für die Translation der Methode in die Klinik sind unter anderem die Modellierung des Signals in heterogenem Gewebe und das geringe Signal-Rausch-Verhältnis bei typischer Dosisapplikation (Jones et al., 2018).

Protonenradiographie

Bei der Protonenradiographie wird die Abbremsung von Protonen im Gewebe entlang einer Einstrahlrichtung gemessen. Um eine Radiographie des Patienten mit Protonen zu erstellen, benötigen die Protonen daher ausreichend Energie, um den Patienten zu durchdringen. Diese und andere technische und physikalische Herausforderungen in der Detektion eines Protonenstrahls mit geringer Intensität erschweren die klinische Implementierung von Systemen zur Protonenradiographie (Johnson, 2017). Es gibt jedoch bereits einen kommerziellen Prototypen für die klinische Anwendung (Sarosiek et al., 2021). Ein weiteres hieran angelehntes Konzept basiert auf dem Testen der Reichweite, engl. *range probing*, anhand von ausgewählten PBS-Testspots (Mumot et al., 2010). Dieses Konzept wurde sogar schon während einzelner Patientenbehandlungen angewendet (Meijers et al., 2021). Beide Konzepte eignen sich prinzipiell zur Verifizierung der Patientenpositionierung und zur Feststellung anatomischer Veränderungen und von Abweichungen der SPR-Vorhersage.

3 Methodik der Reichweiteverifikation mittels Prompt-Gamma-Bildgebung

Wie in Kapitel 2 dargelegt, können Methoden zur Reichweite- und damit Behandlungsverifikation in der Protonentherapie dazu beitragen, die Therapie so effektiv und schonend wie möglich durchzuführen. In dieser Arbeit wird ein Prototyp-System für die PGI-Methode im Hinblick auf seine klinische Anwendung und die daraus resultierenden Vorteile für den Patienten evaluiert. Das folgende Kapitel dient zur Einführung in das Detektionsprinzip sowie der konkreten Charakterisierung des untersuchten Systems für eine Patientenstudie am OncoRay – Nationales Zentrum für Strahlenforschung in der Onkologie.

3.1 Funktionsprinzip der PGI-Schlitzkamera

Eine PGI-Schlitzkamera besteht aus einer Schlitzblende zwischen zwei keilförmig zulaufenden Kollimatorelementen und einem Detektor. Das Funktionsprinzip basiert auf der eindimensionalen Projektion der entlang der Strahlachse der Protonen emittierten PG auf den Detektor (Abbildung 3.1). Bei dem in dieser Studie verwendeten Prototypen kommt ein 4 cm dicker Wolfram-Kollimator zum Einsatz (Smeets et al., 2012), der die hochenergetische Gammastrahlung weitestgehend absorbiert (bis auf ca. 4% bei 4 MeV, vgl. Hubbell und Seltzer, 2004). Die durch die Schlitzblende verlaufende PG hingegen trifft auf ein orthogonal zur Schlitzblende – also in Strahlrichtung – segmentiertes Detektorarray, durch welches ein inverses Intensitätsprofil der PG, im Folgenden *PGI-Signal* genannt, aufgenommen wird. Durch die absolute Lage des PGI-Systems im Behandlungsraum ist das PGI-Signal mit der Reichweite der Protonen im Patienten verknüpft. Im Speziellen lassen sich Reichweiteabweichungen aus dem Vergleich von gemessenem PGI-Signal mit einer simulierten Erwartung des PGI-Signals ermitteln. Für jeden PBS-Spot kann so, nach Rückprojektion der PGI-Signale auf die spotspezifische Strahlachse, eine Reichweiteabweichung berechnet werden. Der Abbildungsmaßstab für die Rückprojektion leitet sich hierbei aus dem Verhältnis der Abstände zwischen Kollimator- und Detektorebene d_{proj} sowie zwischen Kollimatorebene und dazu – im besten Fall – paralleler Strahlachse d_{real} ab. Dementsprechend ergibt sich für das PGI-System ein mit dem Abstand der Strahlachse zum Kollimator keilförmig verlaufendes Sichtfeld, engl. *field of view* (FOV), der Länge l_{real} entlang der Strahlachse, welches durch

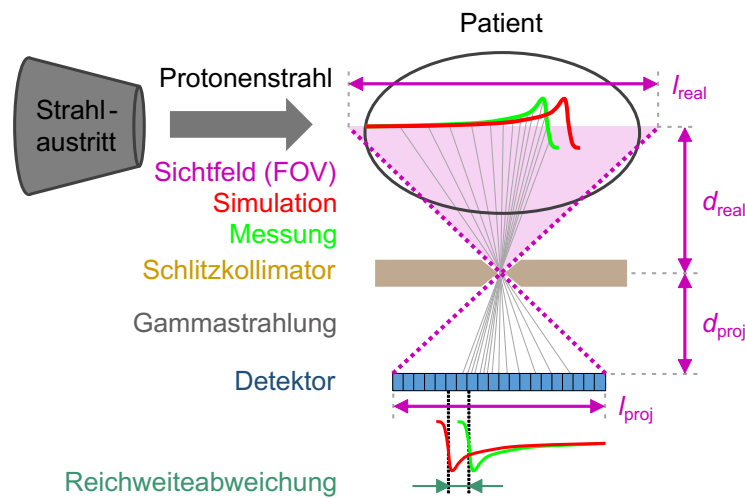


Abbildung 3.1: Skizze zum Detektionsprinzip von Reichweiteabweichungen mittels PGI.

die Ausmaße des Detektorarrays l_{proj} sowie den Abstand zwischen Kollimator- und Detektorebene definiert ist:

$$l_{real} = \frac{d_{real}}{d_{proj}} \cdot l_{proj} \quad (3.1)$$

3.2 Datenaufnahme und -verarbeitung

In diesem Unterkapitel soll der Ablauf zur Detektion von Reichweiteabweichungen anhand des durch die Firma IBA (Ion Beam Applications S.A. Louvain-La-Neuve, Belgien) entwickelten und in dieser Arbeit zum Einsatz gekommenen Systems dargestellt werden (Perali et al., 2014; Smeets et al., 2012).

3.2.1 Detektoraufbau und Signalaufnahme

Der Aufbau des Detektors ist in Abbildung 3.2 dargestellt. Der Detektor ist aus je 20 Segmenten in zwei unabhängigen Reihen aufgebaut. Jedes Segment besteht aus einem Lutetium-Yttrium-Oxyorthosilikat (LYSO) Szintillationskristall (Maße: $100 \times 4 \times 31,5$ mm), der an einen Silizium-basierten Photoelektronenvervielfacher, engl. *silicon photomultiplier* (SiPM), gekoppelt ist. Anhand einer zentralen elektronischen Steuereinheit können die SiPM betrieben und ausgelesen werden (Perali et al., 2014). Aufgrund der Temperaturabhängigkeit der SiPM im Bezug auf die Durchschlagsspannung ist es wichtig, dass die Kamera vor dem Einsatz aufgewärmt wird (Priegnitz et al., 2016a). Da bei einer Messung die Temperatur im SiPM zunimmt, werden hierzu zyklisch Messungen im 60-Sekunden-Takt mit ca. einer Stunde Vorlauf gestartet. Dadurch geht die Temperatur an den SiPM in eine Sättigung über, so dass direkt vor der PGI-Messung ein Temperaturequilibrium herrscht. Während der Signalaufnahme wird die Anzahl der Ereignisse im Energiebereich 3–6 MeV (Smeets et al., 2012)

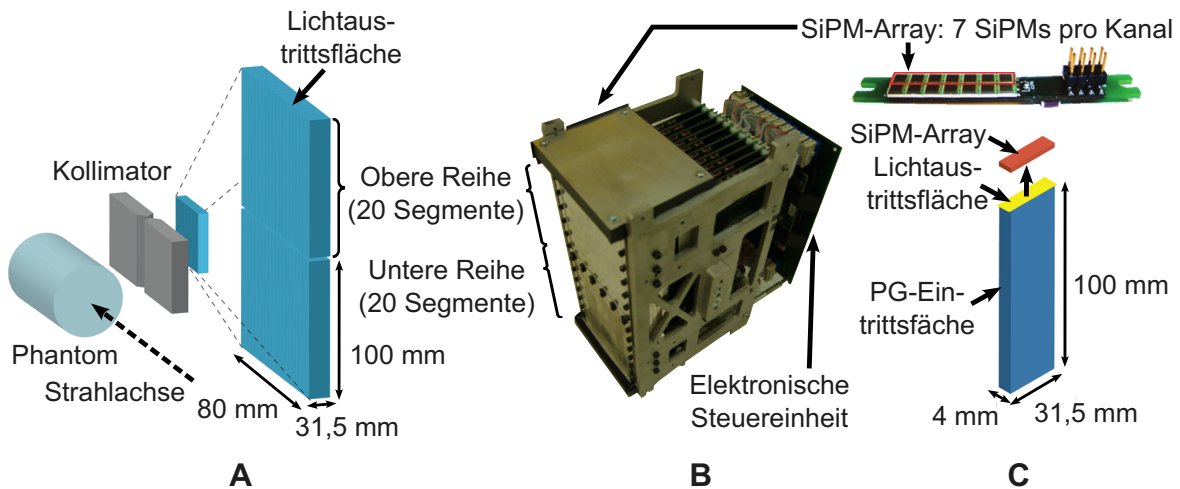


Abbildung 3.2: Aufbau des verwendeten PGI-Detektors. (A) Anordnung der Segmente aus Szintillationskristallen in zwei unabhängigen Reihen. (B) Zusammengesetzter Prototyp in einem Aluminiumgehäuse und mit elektronischer Steuereinheit. (C) Koppelung eines Szintillationskristalls an ein SiPM-Array. (Quellen: In Anlehnung an Perali et al., 2014, und Priegnitz et al., 2015)

in Zeitabständen von 0,5 ms integriert. Diese Zeiteinteilung hat sich als verlässlich bei der Zuordnung der Daten zu den einzelnen Spots gezeigt. Diese Zuordnung erfolgt retrospektiv über das Zählratehistogramm der Messung und auf Grundlage der aufgezeichneten Log-Dateien der Bestrahlungsmaschine.

3.2.2 PGI-Simulation und Bestimmung der Reichweiteabweichung

Die analytische Simulation der PGI-Signale wurde von IBA entwickelt und basiert im Kern auf Monte-Carlo-Simulationen (Sterpin et al., 2015). Im Wesentlichen erfolgt die Berechnung der PGI-Signale in zwei Schritten: (1) die Simulation der Emission von PG im Patienten und (2) die Übertragung der PG zum Detektor. In Vorbereitung für die Simulation ist es notwendig, Parameter der Bestrahlungsanlage, wie z. B. den energieabhängigen Strahlquerschnitt und die Lage der PBS-Ablenkmagnete, einzulesen.

Im ersten Teil der Simulation wird dann anhand des Patienten-CT-Scans und über exemplarische Pfade innerhalb jedes PBS-Spots (engl. *raytracing*) der Weg der Protonen im Patienten bis zur vollständigen Energieabgabe nachverfolgt. Entlang der Protonenpfade wird die momentane Energie der Protonen entsprechend der CT-basierten SPR-Vorhersage und des CSDA-Ansatzes ermittelt. In der Simulation wird nur die Emission von PG über die Nuklide ^{12}C , ^{14}N , ^{16}O , ^{31}P und ^{40}Ca anhand der hinterlegten Material- und Massendichteübersetzung berücksichtigt. Das lokale PG-Emissionsspektrum entlang der Protonenpfade wird im Bereich von 3–8 MeV mittels der lokalen Zusammensetzung des Gewebes, der momentanen Protonenenergie und der mittels Monte-Carlo-Algorithmen vorab bestimmten Wahrscheinlichkeit der PG-Emission für jedes der berücksichtigten Nuklide berechnet. Bei allen

Simulationen, die im Rahmen dieser Arbeit durchgeführt wurden, basiert die für jede Körperregion spezifische Materialübersetzung auf den tabellierten Geweben von Woodard und White (1986). Hingegen ist die Übersetzung in eine Massendichte ein Effekt zweiter Ordnung, weshalb hier vereinfachend ein generischer, über alle Körperregionen gemittelter Ansatz verwendet wurde, der ebenso auf den Daten von Woodard und White (1986) basiert. Die Konversionen beruhen auf einer dedizierten Kalibrierung für CT-Scanner und Scanprotokoll (Wohlfahrt et al., 2018) und im Fall von DirectSPR auf dem MATA-Ansatz (engl. *MATerial Assignment*) (Permatasari et al., 2020).

Der zweite Teil der Simulation umfasst die Übertragung der PG zum Detektor, welche über ein Modell zur Impulsantwort des Systems, d. h. ein Modell des PGI-Signals einer Punktquelle, approximiert ist. Für unterschiedliche geometrische Konfigurationen des Detektors, das energieabhängige Detektoransprechvermögen und die Abschwächung der PG durch den Patienten wurden die angepassten Parameter dieser Impulsantwort tabelliert. Darüber kann die simulierte PG-Emission durch Faltung mit der entsprechenden Impulsantwort in ein PGI-Signal umgewandelt werden.

Beim Einlesen der Messung werden eine Reihe von im Nachfolgenden beschriebenen Korrekturen durchgeführt und schlussendlich die Reichweiteabweichung für jeden Spot bestimmt (Abbildung 3.3). Zunächst erfolgt eine geometrische Korrektur der Messdaten, welche auf Simulationen in einem Wasserphantom basiert (Petzoldt et al., 2018). Die geometrische Korrektur berücksichtigt die nicht gleichbleibende Impulsantwort im FOV der Kamera durch die Detektorgeometrie, den Fluss von Neutronen oder neutroneninduzierten Teilchen und die relative Lage der Detektorsegmente bei der Streuung zwischen den Segmenten. Weiterhin erfolgt bei der Prozessierung der gemessenen PGI-Signale eine Streckung oder Stauchung aller Signaleinträge mit einem globalen Skalierungsfaktor, welcher die Ungenauigkeit in der absoluten Normierung der gemessenen Signale gegenüber der simulierten Signale kompensieren soll. Dieser Skalierungsfaktor wird als Median der Verhältnisse von gemessenen zu simulierten Amplituden der PGI-Signale bestimmt. Der verbleibende vertikale Versatz der Signale aufgrund eines großen, fast konstanten Untergrundes (Sterpin et al., 2015) wird durch die vertikale Ausrichtung der Signale auf ihre halbe Amplitude korrigiert. Zur Verringerung der statistischen Unsicherheit in den gemessenen PGI-Signalen wird außerdem pro PBS-Layer eine zweidimensionale Gaußsche Glättung ($\sigma = 7,8$ mm, vgl. Nenoff et al., 2017) durchgeführt, die im Folgenden *Spotaggregation* genannt wird. Zum Schluss wird für jedes der aggregierten und nochmals geglätteten PGI-Signale eine Reichweiteabweichung berechnet, indem die Methode der kleinsten Quadrate zum Einsatz kommt (Janssens et al., 2016). Durch die Faltung mit einer Gauß-Funktion ($\sigma = 50$ mm) soll die Bestimmung der Reichweiteabweichung auf die abfallende Flanke des PGI-Signals fokussiert und damit Unterschiede am Rand der PGI-Signale unterdrückt werden.

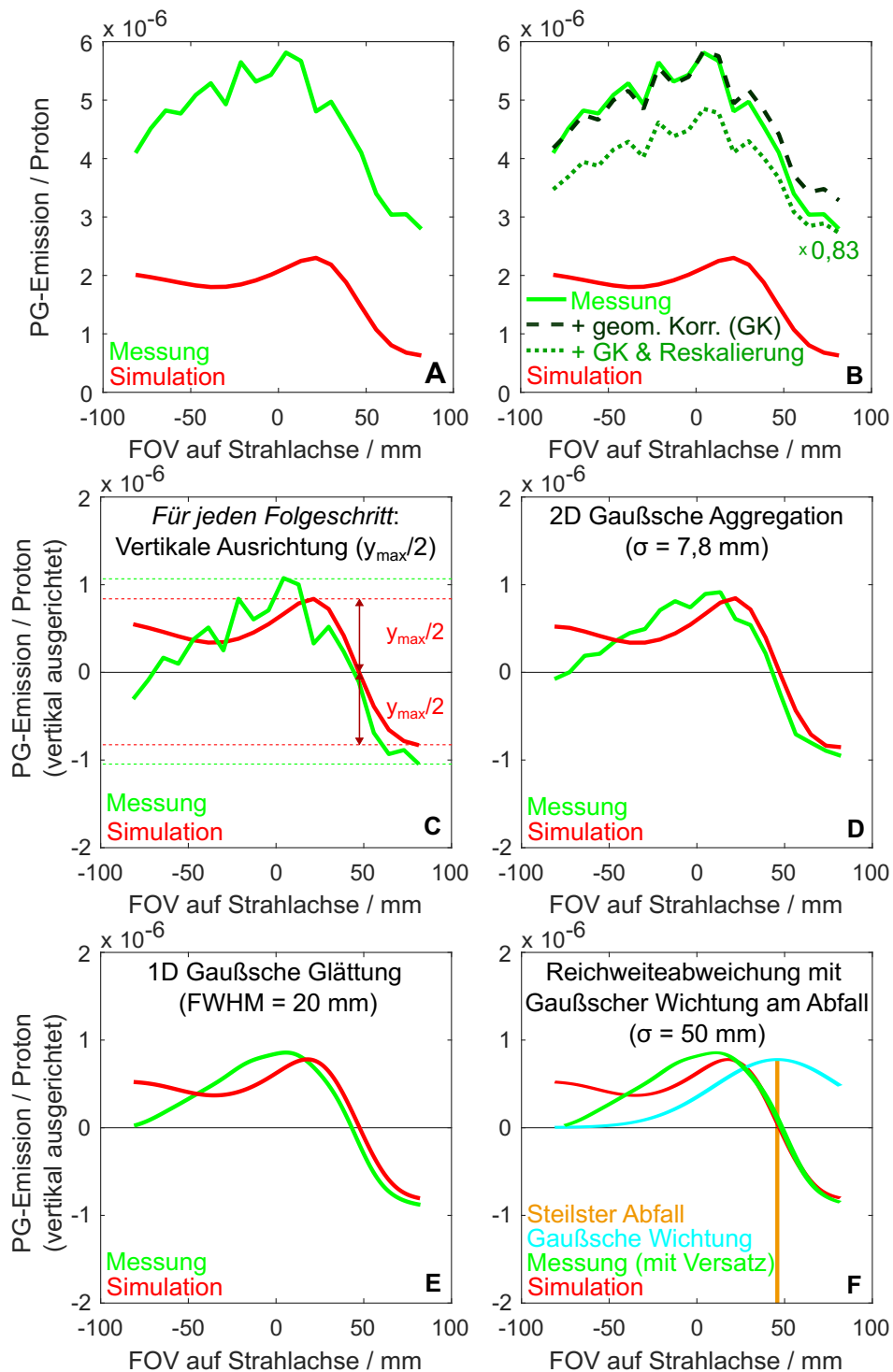


Abbildung 3.3: Visualisierung der PGI-Signalprozessierung: (A) Rohsignal, (B) geometrische Korrektur und globale Reskalierung (hier mit Faktor 0,83), (C) vertikale Ausrichtung auf die halbe Amplitude, (D) 2D Gaußsche Aggregation von Spots im selben Layer, (E) 1D Gaußsche Glättung der Signale und (F) Bestimmung der Reichweiteabweichung durch die Methode der kleinsten quadratischen Abweichung mit einer Gaußschen Wichtung am distalen Signalabfall.

3.3 Charakterisierung des PGI-Prototyps

Nach erster erfolgreicher Anwendung der PGI-Schlitzkamera am OncoRay bei Patientenbestrahlungen mittels Doppelstreuverfahren (Richter et al., 2016) und der systematischen Untersuchung im PBS-Verfahren mittels eines heterogenen Kopfphantoms (Nenoff et al., 2017) wurde zusammen mit IBA eine zweite Generation des Systems entwickelt (Abbildung 3.4). Dieses System sollte sich maßgeblich von seinem Vorgänger durch die Positionierung in der Patientenbehandlung unterscheiden, welche vor allem schneller und genauer werden sollte. Das Detektionssystem an sich blieb unverändert. Einzig der Abstand zwischen Kollimator und Detektor wurde auf 117,25 mm verkleinert und fixiert, was zu einem größeren FOV, dafür aber geringerer Auflösung des PGI-Signals führt. Der neue Trägerwagen des Detektionssystems mit einer Masse von ca. 330 kg wird mechanisch über eine Andockstation mit hoher

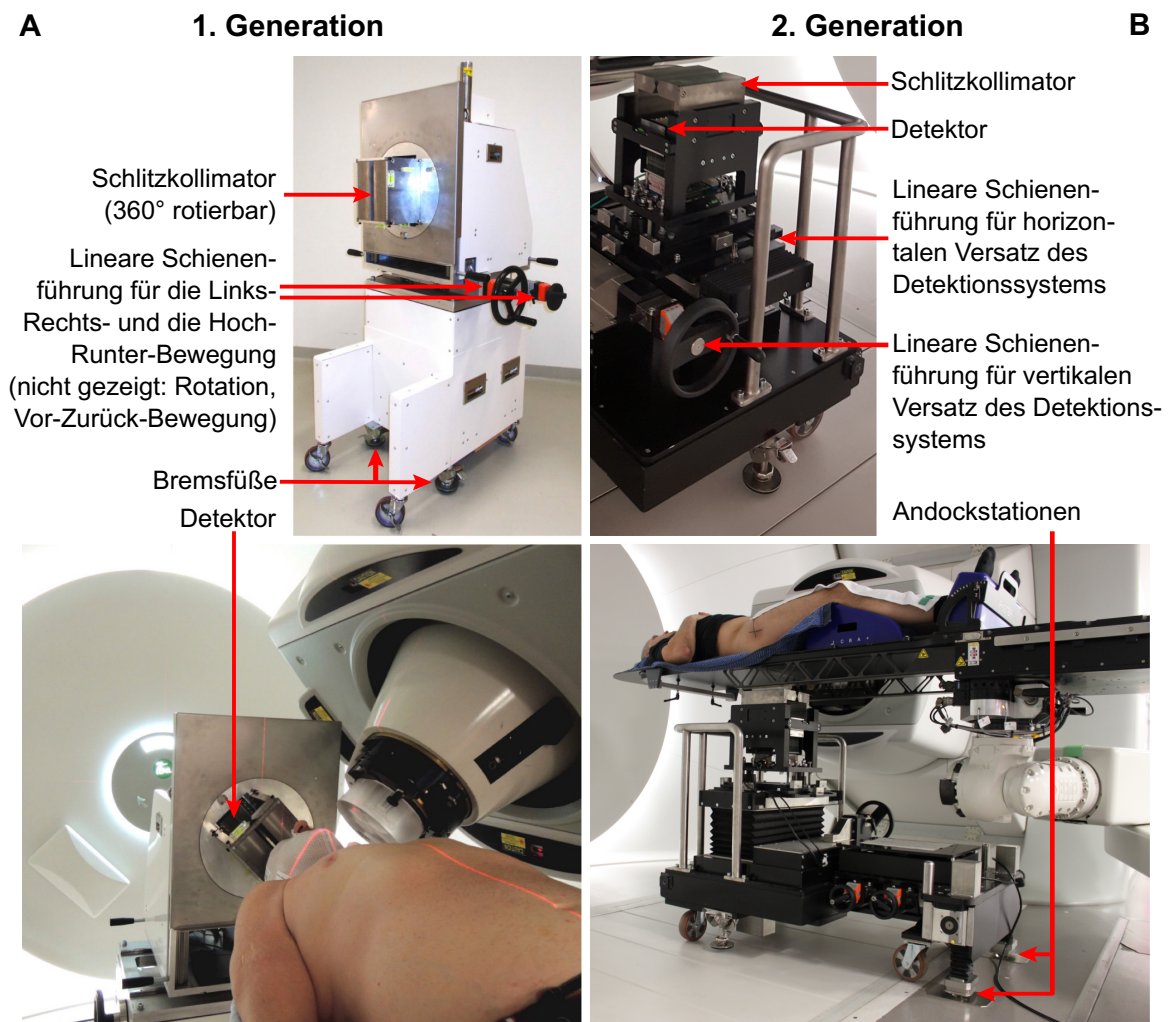


Abbildung 3.4: Vergleich zwischen den Trägerwagen der ersten (A) und zweiten (B) Generation des PGI-Systems in der klinischen Anwendung. (Quelle: In Anlehnung an Richter et al., 2016, und Khamfongkhrua, 2021)

Reproduzierbarkeit relativ zum Isozentrum positioniert (vgl. Unterkapitel 3.3.2), anstatt wie vorher mit dem Laser ausgerichtet (Priegnitz et al., 2016a). Dadurch kann das System im klinischen Einsatz sehr schnell positioniert werden (max. eine Minute Verlängerung der Behandlungszeit pro aufgenommenem Feld). Außerdem kommt sowohl in der Kalibrierung des PGI-Systems als auch in der Patientenpositionierung das gleiche planare Röntgensystem zum Einsatz, sodass systematische Unsicherheiten des Röntgensystems ignoriert werden können und eine vom Patienten und Phantom unabhängige Qualitätssicherung der Position des PGI-Systems möglich ist. Das System kann in allen Freiheitsgraden der Rotation und Translation kalibriert werden, besitzt im klinischen Einsatz jedoch nur zwei adjustierbare Freiheitsgrade: den vertikalen Abstand des Detektionssystems zum Isozentrum und den horizontalen Versatz des Detektionssystems in Strahlrichtung (parallel zum Boden). Durch diese Einschränkung war es zunächst nur möglich, Bestrahlungen mit horizontaler Gantryposition – als Bestrahlung unter 90° und 270° bezeichnet – zu monitorieren. Der Grund dafür ist, dass nur in diesem Fall die Kollimatorebene und die zentrale Strahlachse im Isozentrum parallel zueinander liegen, was eine wichtige Voraussetzung für die bisherige Simulationssoftware darstellt. Tatsächlich können Reichweiteabweichungen auch bei nicht parallel liegender Kollimatorebene und zentraler Strahlachse bestimmt werden, jedoch nimmt die Auflösung der projizierten PGI-Signale ab, je steiler die Verkippung ist. Um mehr Flexibilität beim Monitoring von Patienten zu haben, hat IBA im Bearbeitungszeitraum dieser Arbeit eine *erweiterte* Version der Simulationssoftware zur Verfügung gestellt, bei welcher die Kamera auch verkippt zur zentralen Strahlachse ausgerichtet sein kann. Die Validierung dieser neuen Version wird in Unterkapitel 3.5.2 beschrieben. Im nächsten Unterkapitel folgt zunächst, wie das System der zweiten Generation kalibriert und im Hinblick auf seine Positionierungspräzision untersucht wurde, was eine wichtige Voraussetzung für die systematische Anwendung dieses Systems darstellt.

3.3.1 Kalibrierung des Systems

Die Kalibrierung des eingesetzten PGI-Schlitzkamera-Systems kann in drei Teilschritte untergliedert werden: (1) die Temperatur und Energiekalibrierung des Detektorsystems, (2) die Kalibrierung der Kollimatorposition im angedockten Trägerwagen relativ zum Isozentrum und (3) die Kalibrierung des vollständig montierten Systems (Trägerwagen mit Kollimator und Detektionssystem) relativ zu einem bestrahlten Phantom im Isozentrum.

(1) Temperatur- und Energiekalibrierung

Die Temperatur- und Energiekalibrierung des Detektors muss in regelmäßigen Abständen wiederholt werden, um allmähliche Veränderungen der verwendeten Bauteile (z. B. durch Strahlung oder mechanische Belastung) zu kompensieren. Dabei wird die Kamera aus dem Trägerwagen ausgebaut und an der experimentellen Bestrahlungsanlage unter normierten

Bedingungen kalibriert. Zuerst wird die linear von der Temperatur abhängige Verstärkung der SiPM kalibriert, indem die Verstärkung der charakteristischen 6,13 MeV PG-Linie (anhand des dominierenden Double-Escape-Peaks) eines Wasserphantoms in Abhängigkeit von der Temperatur der SiPM (insg. zwei Messpunkte unterschiedlicher, konstanter Temperatur) bestimmt wird. Im zweiten Schritt wird das Energiespektrum jedes einzelnen Szintillationskristalls mittels des protonenbestrahlten Wasserphantoms und eines ebenso aktivierten Aluminiumtargets kalibriert. Im klinischen Einsatz wird dieses standardisierte Vorgehen in etwa aller 2–3 Monate durchgeführt, wie bereits von Priegnitz et al. (2016a) eruiert. Eine Unsicherheit durch das Ein- und Ausbauen des Detektorsystems auf die Bestimmung der Reichweiteabweichung konnte im Rahmen dieser Arbeit experimentell nicht nachgewiesen werden und ist mit der Abschätzung von $\sigma \leq 0,1 \text{ mm}$ vernachlässigbar für den weiteren Kalibrierungsprozess.

(2) Trägerwagenkalibrierung

Die Kalibrierung des Trägerwagens mit Kollimator erfolgt schrittweise: Zunächst werden die drei Freiheitsgrade der Rotation des Kollimators kalibriert. Anschließend wird die Position des Kollimators relativ zum Isozentrum der Bestrahlung über die drei Translationsfreiheitsgrade eingestellt. In Abbildung 3.5 ist ein schematischer Überblick des Koordinatensystems gegeben, auf welches diese Beschreibung Bezug nimmt. Die Kalibrierung wird mithilfe einer

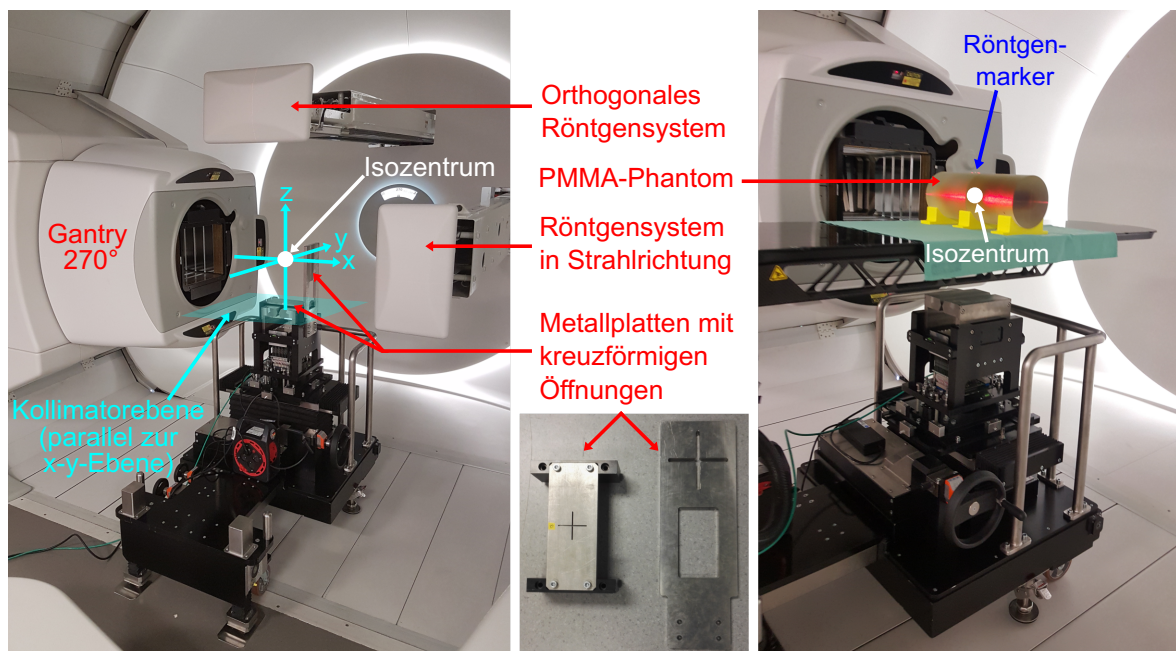


Abbildung 3.5: Koordinatensystem zur Kalibrierung des PGI-Systems unter einem Gantrywinkel von 270° (links). Zur Trägerwagenkalibrierung (2) mit dem planaren Röntgensystem werden zwei Metallplatten an den Trägerwagen geschraubt. Zur PGI-Strahlkalibrierung (3) zwischen PGI-Messung und PGI-Simulation wird ein Phantom aus Polymethylmethacrylat (PMMA) anhand eines Röntgenmarkers im Isozentrum positioniert (rechts).

Wasserwaage (Rotation um x - und y -Achse) und anhand der beiden für die Patientenpositionierung verwendeten, orthogonal angeordneten, planaren Röntgensysteme (Rotation um z -Achse sowie Translation) durchgeführt. Zur Justierung in der Translation werden an den Trägerwagen zwei Metallplatten mit kreuzförmigen Öffnungen angebracht, welche in den beiden Röntgenprojektionen die Position des auf dem Trägerwagen montierten Kollimators anzeigen. Die Rotation des Systems aufgrund der einseitigen Belastung durch eine der Metallplatten wurde empirisch korrigiert (in Strahlrichtung eine mittlere Abweichung von 0,07 mm für zehn Messungen). Der Kollimator wird unter einem Gantrywinkel von 270° kalibriert, da nur hier beide Röntgensysteme – insbesondere das System zur Justierung der Schlitzposition in Strahlrichtung – zur Bildgebung genutzt werden können. Das Ziel der Kalibrierung ist es, dass (a) die Kollimatorebene parallel zur x - y -Ebene in einem Abstand von 250 mm liegt, (b) die Schlitzblende in der Kollimatorebene parallel zur y -Achse verläuft (Rotation um z -Achse) und (c) der Mittelpunkt der Schlitzblende in x - und y -Richtung auf der z -Achse liegt. In der Bearbeitungszeit dieser Arbeit wurde die Kalibrierung des Systems insgesamt dreimal durchgeführt, da im Zuge der Einführung neuer Software für die Röntgensysteme die Position des Isozentrums des Bestrahlungsgerätes überprüft und mehrfach angepasst wurde.

Ein wichtiger Aspekt für die Kalibrierung ist die Präzision des Andockvorgangs (siehe Unterkapitel 3.3.2). Um die Streuung der Position des Systems durch die einzelnen Andockvorgänge zu berücksichtigen, soll das oben genannte Ziel der Kalibrierung (Soll-Wert) für den Mittelwert über mehrere Andockvorgänge erreicht werden (vergleichbar mit dem Monitoring mehrerer Felder). Dazu werden beim Kalibrierungsvorgang die Mittelwerte aller Freiheitsgrade für insgesamt zehn Andockvorgänge (Ist-Wert) gebildet, welche dann als Korrektur zum Soll-Wert hin verwendet werden. Aufgrund der statistischen Streuung und mechanischen Limitationen in der Einstellung der Freiheitsgrade am Trägerwagen, wird die Kalibrierung bereits akzeptiert, wenn nach der Korrektur die Mittelwerte der nächsten zehn Andockvorgänge (Ist-Wert) innerhalb von 0,4 mm in der Translation und $0,1^\circ$ in der Rotation liegen. Die Auflösung der verwendeten Messsysteme ist notwendigerweise geringer: $0,01^\circ$ für die Winkelbestimmung durch das Röntgensystem und $0,006^\circ$ durch die Wasserwaage (Skalenabstand) sowie 0,1 mm für die Bestimmung der Translation durch das Röntgensystem. Während regelmäßiger Messungen zur Qualitätssicherung des PGI-Systems (etwa aller 2–3 Monate) über den gesamten Beobachtungszeitraum zeigte sich die Stabilität der kalibrierten Position, was die Verlässlichkeit für die aufgenommenen PGI-Daten hervorhebt.

(3) PGI-Strahlkalibrierung

Die Kalibrierung des vollständigen Messsystems erfolgt anhand von am Behandlungsstrahl gemessenen PGI-Signalen. Anhand dieser Kalibrierung sollen systematische Abweichungen zwischen Simulation und Messung bestimmt werden, deren Ursache sowohl von der Simulation als auch vom Detektor stammen können. Als bestrahltes Phantom kommt ein homogener Zylinder aus Polymethylmethacrylat (PMMA) mit einem Durchmesser von 15 cm und

einer Länge von 20 cm zum Einsatz, dessen mittleres $SPR = 1,1603 \pm 0,0004$ mit einer mehrschichtigen Ionisationskammer (Giraffe, IBA Dosimetry GmbH, Schwarzenbruck, Deutschland) ermittelt wurde. Zur Positionierung des Phantoms im Isozentrum der Bestrahlung, wird auf das Phantom mithilfe eines Lineals ein Röntgenmarker in Höhe der Bragg-Peak-Position eines zentral im Phantom applizierten 160 MeV PBS-Spots angebracht. Unabhängig davon, dass die Trägerwagenkalibrierung (2) und die Positionierung des Phantoms beide mittels der planaren Röntgensysteme durchgeführt werden, kann eine globale Reichweiteabweichung (*globaler Versatz*) zwischen den gemessenen und simulierten PGI-Profilen verbleiben. Die Bestimmung dieses Versatzes ist das Ziel dieser Kalibrierung und erfolgt anhand von PGI-Messungen eines 160 MeV PBS-Spots.

Der folgende Messablauf wurde zur Bestimmung der Genauigkeit des Systems aller 1–2 Monate (insg. neunmal) wiederholt (Abbildung 3.5): Zunächst wird der berechnete Ort des Bragg-Peaks im Phantom anhand des Markers im Isozentrum positioniert. Anschließend wird das PGI-System wie in einer regulären Behandlung positioniert und es wird fünf mal das PGI-Signal eines 160 MeV PBS-Spots mit 250 MU (engl. *monitor units*) aufgenommen. Daraufhin wird die Schlitzposition des PGI-Systems (und somit der Versatz des gesamten Detektionssystems auf der Strahlachse) im Röntgenbild bestimmt. Eine Änderung der gemessenen Schlitzposition wirkt sich gleichermaßen auf die bestimmte Reichweiteabweichung mittels PGI aus. Dieser hier zu bestimmende globale Versatz zwischen diesen beiden Variablen kann korrigiert werden und zeigt sich in der Differenz der jeweiligen Messwertepaare.

In Abbildung 3.6 A sind insgesamt 20 Messpunkte für die Ermittlung des globalen Versatzes dargestellt, der mit Wichtung über die inverse Varianz der Messpunkte und die Kleinste-Quadrate-Anpassung zu $(1,12 \pm 0,28)$ mm ermittelt wurde. Die Gesamtunsicherheit ergibt sich aus der Unsicherheit der Markerposition auf dem Phantom, die mit $\sigma = 0,25$ mm abgeschätzt wurde, der Genauigkeit der SPR-Bestimmung von $\sigma = 0,06$ mm und der statistischen Unsicherheit aus der Kleinste-Quadrate-Anpassung $\sigma = 0,1$ mm. Der globale Versatz wurde hier nur für eine Protonenenergie, ein bestimmtes Material und mit der Position des Bragg-Peaks zentral über der Schlitzblende ermittelt. Auf die Abhängigkeit des globalen Versatzes vom FOV der Kamera und der Protonenenergie wird ausführlicher in Unterkapitel 3.5 eingegangen. Was die Übertragbarkeit zu anderen Materialien betrifft, so hat sich der hier ermittelte Versatz für das in der Vergleichsstudie verwendete anthropomorphe Kopfphantom aus gewebeäquivalenten Materialien bestätigt (vgl. Kapitel 6).

Unter anderen Gantrywinkeln kann sich der globale Versatz von dem ermittelten Wert unterscheiden, da sich die Position des Isozentrums im Raum aufgrund mechanischer Limitationen der Bestrahlungsanlage ändern kann. Während der Patient in die veränderte Position des Isozentrums mittels tabelliertem Korrekturvektor nachgefahren wird (Isozentrumskorrektur), behält das PGI-System seine kalibrierte Position relativ zum Isozentrum von 270° bei. Allerdings ergeben sich durch den sich mitbewegenden Gantryboden auch für die Position

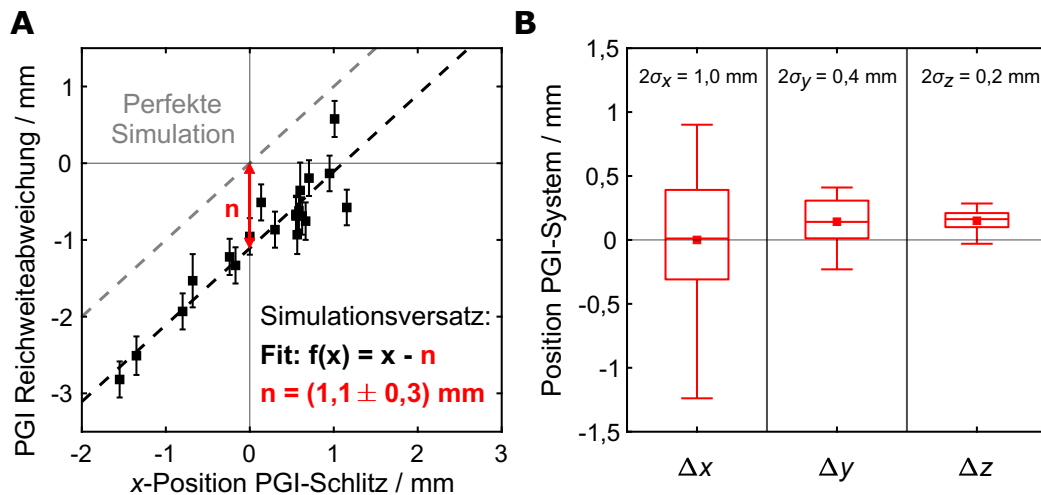


Abbildung 3.6: Charakterisierung des PGI-Systems. (A) Kalibrierung des PGI Systems zwischen Bildgebung mit orthogonalem Röntgensystem und Messung im homogenen PMMA-Phantom. Unter der Annahme eines globalen Versatzes wird dieser durch eine Kleinste-Quadrate-Anpassung der Messpunkte bestimmt. Die Unsicherheitsbalken werden durch die statistische Unsicherheit der Messung, des Röntgensystems, des Ein- und Ausbaus der Kamera und der Auflösung des Reichweitebestimmungs-Algorithmus ermittelt. (B) Box-Plot über 70 Messpunkte der Positionierungspräzision des PGI-Systems für die Abweichung vom Isozentrum entlang der Translation in Strahlrichtung Δx , parallel zum Schlitz Δy und vertikal zur Schlitzebene Δz . Der quadratische Marker zeigt den Mittelwert der jeweiligen Verteilung an.

des PGI-Systems Veränderungen, die sich vor allem anhand einer Rotation des Trägerwagens entlang der y -Achse bemerkbar machen. Zwischen dem Gantrywinkel 270° und 90° wurde bspw. im Mittel über 10 Messungen eine Rotation des Kollimators um $0,04^\circ$ gemessen, was zu einer Verschiebung der Schlitzblende um $0,7 \text{ mm}$ in Strahlrichtung führte.

Da es schwierig ist, die Summe dieser Einflüsse auf den globalen Versatz exakt zu bestimmen und außerdem unter Gantrywinkel 90° die Möglichkeit für orthogonale Röntgenbildung des PGI-Systems fehlt, wurde der globale Versatz für den Gantrywinkel 90° hier empirisch ermittelt. Der Grundgedanke hierbei ist die Betrachtung der Mittelwerte über viele Andockvorgänge, um die Unsicherheit der fehlenden Information über die exakte Position der Schlitzblende bei jeder Strahlmessung zu reduzieren. In einer Messkampagne wurde für je 20 Andockvorgänge unter 270° und 90° die Reichweiteabweichung zwischen gemessenem und simuliertem PGI-Signal für einen 160 MeV PBS-Spot bestimmt. Die Differenz der erhaltenen mittleren Reichweiteabweichungen beträgt $(2,84 \pm 0,15) \text{ mm}$, welche als Korrekturterm für den globalen Versatz unter einem Gantrywinkel von 90° dient. Dieser Wert bezieht sich auf die aktuelle Isozentrumskorrektur der Gantry; für vorherige und zukünftige Kalibrierungen muss der relative Versatz der Patientenliege zwischen 270° und 90° Gantrywinkel mit berücksichtigt werden.

Insgesamt werden also drei Werte für die Bestimmung der globalen Korrektur aufsummiert: der Ist-Wert der Schlitzposition relativ zum Isozentrum (aus der Trägerwagenkalibrierung),

Tabelle 3.1: Überblick der globalen Korrekturen in Strahlrichtung, die aus der Kalibrierung des PGI-Systems im Isozentrum und aus der Phantombestrahlung resultieren.

Zeitraum	Gantry-Winkel	Schlitzposition Ist-Wert / mm	Globaler Versatz / mm	Gantrywinkel Korrektur / mm	Σ / mm
1.11.2018 bis 24.2.2020	270° 90°	-0,39	1,12	0 2,84 - 3,23 ^a	0,73 0,34
25.2.2020 bis 13.5.2020	270° 90°	0,30	1,12	0 2,84	1,42 4,26
14.5.2020 bis 22.06.2022	270° 90°	-0,02	1,12	0 2,84	1,10 3,94

^a Differenz der Patientenverschiebung in Strahlrichtung zwischen 270° und 90° bezogen auf eine frühere Isozentrumskorrektur der Gantry.

der globale Versatz der PGI-Messung relativ zur Simulation und ein für jeden Gantrywinkel spezifischer Korrekturterm. In Tabelle 3.1 sind die verwendeten Beiträge je nach Zeitraum der Kalibrierung aufgeschlüsselt. Die Korrekturwerte wurden auf alle Felder in dieser Studie angewandt, auch auf nicht horizontale Gantrywinkel nahe zu 90° und 270° (Abweichungen bis max. $\pm 10^\circ$ in dieser Studie). Diese globalen Korrekturterme in Strahlrichtung sind das Ergebnis einer umfangreichen Kalibrierung und stellen keine Unsicherheit des PGI-Systems dar.

3.3.2 Positionierungspräzision

Die Kenntnis über die Positionierungspräzision des PGI-Systems bildet eine wichtige Grundlage für die Unsicherheitsabschätzung im klinischen Einsatz und bei der Kalibrierung des Systems. Zur Bestimmung wird das PGI-System wiederholt in der Andockstation positioniert und die relative Position zum Isozentrum mittels Röntgensystem und der am Trägerwagen befestigten Metallplatten mit kreuzförmiger Öffnungen ermittelt. Die Streuung dieser Position ist per Definition die Positionierungspräzision des Systems. Der Beobachtungszeitraum des Andockverhaltens beläuft sich auf insgesamt 13 Monate (November 2018 bis November 2019). In diesem wurden sieben Messkampagnen mit jeweils zehn Einzelmessungen, d. h. also insgesamt 70 Messungen, der Position des Systems relativ zum Isozentrum mit der gleichen Kalibrierung (vgl. Unterkapitel 3.3.1) durchgeführt. Es ergab sich eine Positionierungspräzision von $\sigma_x = 0,51$ mm in Strahlrichtung, $\sigma_y = 0,20$ mm parallel zum Schlitz und $\sigma_z = 0,08$ mm orthogonal zur Schlitzebene (Abbildung 3.6 B). Für die Rotation um die Strahlachse (x -Achse) wurde ein Wert von $\sigma = 0,026^\circ$ und um die z -Achse von $\sigma = 0,015^\circ$ gemessen. Die Rotation um die y -Achse konnte mit dem Röntgensystem nicht ermittelt werden; für zehn Einzelmessungen mit der Wasserwaage am selben Tag belief sich die Präzision auf $\sigma = 0,004^\circ$. Mit einer Positionierungspräzision im Submillimeterbereich erfüllt das PGI-System der zweiten Generation die klinischen Anforderung (Pausch et al., 2020).

3.4 Überblick zur PRIMA-Studie

Die erste klinische Anwendung einer PGI-Schlitzkamera erfolgte im Doppelstreuverfahren (Richter et al., 2016), welche im Rahmen der PRIMA-Studie (*PRompt-gamma-IMaging*: DRKS00009224, Ethikvotum: EK181042015) durchgeführt wurde. Tatsächlich wird das volle Potential der Methode aber nur im PBS-Verfahren ausgeschöpft, dessen klinische Anwendbarkeit wenig später auch demonstriert wurde (Xie et al., 2017, 2020). An diese Studien anknüpfend wird seit November 2018 die klinische PRIMA-Studie an der UPTD mit dem PGI-System der zweiten Generation im PBS-Verfahren fortgeführt. Davon wurden in dieser Arbeit 15 Prostata- und 4 Kopf-Hals-Tumor-Patienten untersucht, die im Zeitraum von November 2018 bis März 2021 aufgenommen wurden. Tabelle 3.2 und Tabelle 3.3 stellen einen Überblick der hier ausgewerteten Patienten-Datensätze dar. Die aufgenommenen PGI-Messungen sind unmittelbar mit Kontroll-CT-(kCT-)Aufnahmen in Behandlungsposition verknüpft, welche direkt vor der Bestrahlung mithilfe eines im Behandlungsraum befindlichen CT-Scanners, engl. *in-room CT*, auf Schienen durchgeführt wurden. Hierbei kam ein DECT-Scan-Protokoll (80/140 kVp) und ein baugleicher CT-Scanner (SOMATOM Definition AS; Siemens Healthineers, Forchheim, Deutschland) wie für die pCT-Scans zum Einsatz.

Allen der hier ausgewerteten Prostata-Tumor-Patienten wurde zum Schutz des Rektums ein Spacer-Ballon (BioProtect Spacer, Tzur Yigal, Israel) zwischen Rektum und Prostata

Tabelle 3.2: Überblick der mit PGI monitorierten Prostata-Tumor-Patienten in dieser Studie.

Patient	Anzahl an Messungen ^a	Besonderheiten
P1	8 ^b	Rektumballon
P2	18	
P3(1)	8	
P3(2)	6	Anpassung FOV
P4	14	Rektumballon
P5	18	
P6	10	
P7	12	
P8	2	Rektumballon
P9	18	
P10	17 ^c	
P11	14	
P12	10	
P13	14	Rektumballon
P14	18	
P15	14	

^a Summe der Anzahl an Messungen N unter 90° und 270° , wobei $N(90^\circ) = N(270^\circ)$.

^b Alle Messungen unter 270° durchgeführt.

^c Eine Messung mehr unter 90° durchgeführt.

eingesetzt. Sie erhielten eine hypofraktionierte Bestrahlung mit 20 Fraktionen, in denen neben 48 Gy für die Samenblase eine simultan integrierte Boostdosis von 60 Gy für die Prostata appliziert wurde. In allen Fällen wurde eine Feldanordnung von zwei sich gegenüberstehenden lateralen Bestrahlungsfeldern mit einer paritätischen, uniformen Einzeldosis (von 1,5 Gy pro Feld pro Fraktion für die Prostata) unter robuster Optimierung geplant. Außerdem wurde nach Ermessen der planenden Medizinphysiker und Ärzte bei vier der Prostata-Tumor-Patienten zusätzlich ein wassergefüllter Rektumballon eingesetzt (Qfix, Avondale PA, USA), der für eine konstante Füllung des Rektums unter der Therapie sorgen sollte. Für die Prostata-Tumor-Patienten-Kohorte, welche ausschließlich horizontale Bestrahlungsfelder (Gantrywinkel) aufweist, wurde die bisherige vereinfachte PGI-Simulation verwendet.

Die Kopf-Hals-Tumor-Patienten wurden entsprechend einer Standardfraktionierung mit sequenziellem Boost bestrahlt. In der ersten Teilserie wurden das elektive CTV mit 50 Gy bestrahlt, welches das primäre Ziel (Tumor und involvierte Lymphknoten) und elektive Gebiete der Lymphknoten enthält. In der zweiten Teilserie, dem sogenannten *Boost*, wurde eine Dosis von 16 Gy oder 20 Gy im Bereich des primären Tumors deponiert. Die Dosis der einzelnen Bestrahlungsfelder (2–3 pro Teilserie) wurde uniform unter Anwendung robuster Optimierung geplant. Bei allen Bestrahlungsfeldern wurde die Reichweite im Patienten durch die Verwendung eines Reichweiteverschiebungs-Blocks (aus dem Engl. *Range-Shifter*) am Strahlaustritt der Gantry reduziert, um auch oberflächennahe Spots zu applizieren. Für zwei Patienten war aufgrund der Einschränkung des PGI-Systems auf nahezu horizontale Bestrahlungsfelder nur der Boost mit PGI monitorierbar. Bei allen Kopf-Hals-Tumor-Patienten wurde unabhängig vom Gantrywinkel die erweiterte Simulationssoftware verwendet (vgl. Unterkapitel 3.5.2).

Tabelle 3.3: Überblick der mit PGI monitorierten Kopf-Hals-Tumor-Patienten in dieser Studie.

Patient (Teilserie)	Gantry-winkel	Tisch-rotation	Anzahl an Messungen	Art des Karzinoms	Besonderheiten
KH1(2)	90°	345°	1	Tonsillen	
	280°	20°	1	Tonsillen	
KH2(1)	80°	350°	5	Mund	280 mm Abstand Isozentrum-Kollimator
	280°	10°	5	Mund	280 mm Abstand Isozentrum-Kollimator
KH2(2)	270°	15°	1	Mund	300 mm Abstand Isozentrum-Kollimator
KH3(1)	270°	85°	4	Nase	
KH3(2)	270°	85°	1	Nase	
KH4(2)	275°	10°	4	Hypopharynx	

3.5 Experimentelle Studien zur PGI-Simulationsgenauigkeit

Dieser Abschnitt widmet sich Experimenten und Untersuchungen, die durch die Ergebnisse der ersten klinischen Datenaufnahmen getriggert wurden. Hierbei handelt es sich nicht um eine vollständige Betrachtung der PGI-Simulationsgenauigkeit, sondern lediglich um eine Ergänzung zu bereits vorangegangenen Studien von Smeets et al. (2012), Priegnitz et al. (2015), Priegnitz et al. (2016b), Janssens et al. (2016), Nenoff et al. (2017) und Petzoldt et al. (2018). Aufgrund des neuen Trägerwagens, der veränderten Geometrie zwischen Detektor, Kollimator und Strahlisozentrum sowie der Anwendung der geometrischen Korrektur auf die PGI-Signale, kann kein unmittelbarer Vergleich zu den genannten Studien gezogen werden.

3.5.1 Abhängigkeit vom PGI-Sichtfeld und der Protonenenergie

Das PGI-System wurde bislang nur für einen zentralen 160 MeV PBS-Spot im Zentrum des FOV des Systems mit hoher Genauigkeit kalibriert (vgl. Unterkapitel 3.3.1). Hier wird das Verhalten von Reichweiteabweichungen über das gesamte FOV und für verschiedene Protonenenergien im gleichen homogenen PMMA-Phantom untersucht.

Experiment 1: Zentrale Strahlachse 270°

Die Abhängigkeit der Reichweiteabweichung entlang der zentralen Strahlachse wurde im PMMA-Phantom durch einen PBS-Plan aus mehreren Energieschichten im Abstand von 2 MeV mit zentralen Spots untersucht. Um mögliche FOV-basierte Effekte von Effekten der Protonenenergie zu trennen, wurde zusätzlich die Position des PMMA-Phantoms (*Setup*) in vier diskreten Schritten in Strahlrichtung variiert, sodass sich im Zentrum des FOV jeweils Spots mit der Protonenenergie 120, 140, 160 oder 180 MeV befanden (Abbildung 3.7).

Im Ergebnis zeigte sich, dass die Abhängigkeit der Reichweiteabweichung von der Protonenenergie nicht eindeutig ist (Abbildung 3.7 A), d. h. die Position des Bragg-Peaks im FOV beeinflusst die gemessene Reichweiteabweichung. Stellt man die Reichweiteabweichung in Abhängigkeit von der Bragg-Peak-Position im FOV dar (Abbildung 3.7 B1), so zeigte sich ein übereinstimmender Verlauf der Reichweiteabweichungen im distalen Teil des FOV (> 0 mm) und Abweichungen zwischen den unterschiedlichen Energien im proximalen Teil des FOV (< 0 mm). Für ein ideales PGI-System würde ein konstanter Verlauf über das FOV erwartet werden. Das dargestellte Verhalten ist vermutlich durch den Einfluss von sekundärer Strahlung wie Neutronen zu erklären, deren Erzeugung von der Protonenenergie abhängt (Smeets et al., 2012) und die in der PGI-Simulation nicht berücksichtigt werden (Sterpin et al., 2015).

Um der Hypothese der Streustrahlung weiter auf den Grund zu gehen, wurde eine weitere Messung vorgenommen, bei welcher die Schlitzeblende durch einen passenden Wolfram-Einsatz verschlossen wurde. Dadurch kann derjenige Untergrund vom PGI-Signal

abgezogen werden, welcher nicht durch die Schlitzblende verläuft. Der Abzug des Untergrundes verringert die Energieabhängigkeit der Reichweiteabweichung im FOV des PGI-Systems (Abbildung 3.7 B2). Verbleibende Unterschiede könnten u. a. auf durch die Schlitzblende kommende Streustrahlung und auf den unterschiedlich stark aufgestreuten Protonenstrahl im Phantom zurückzuführen sein.

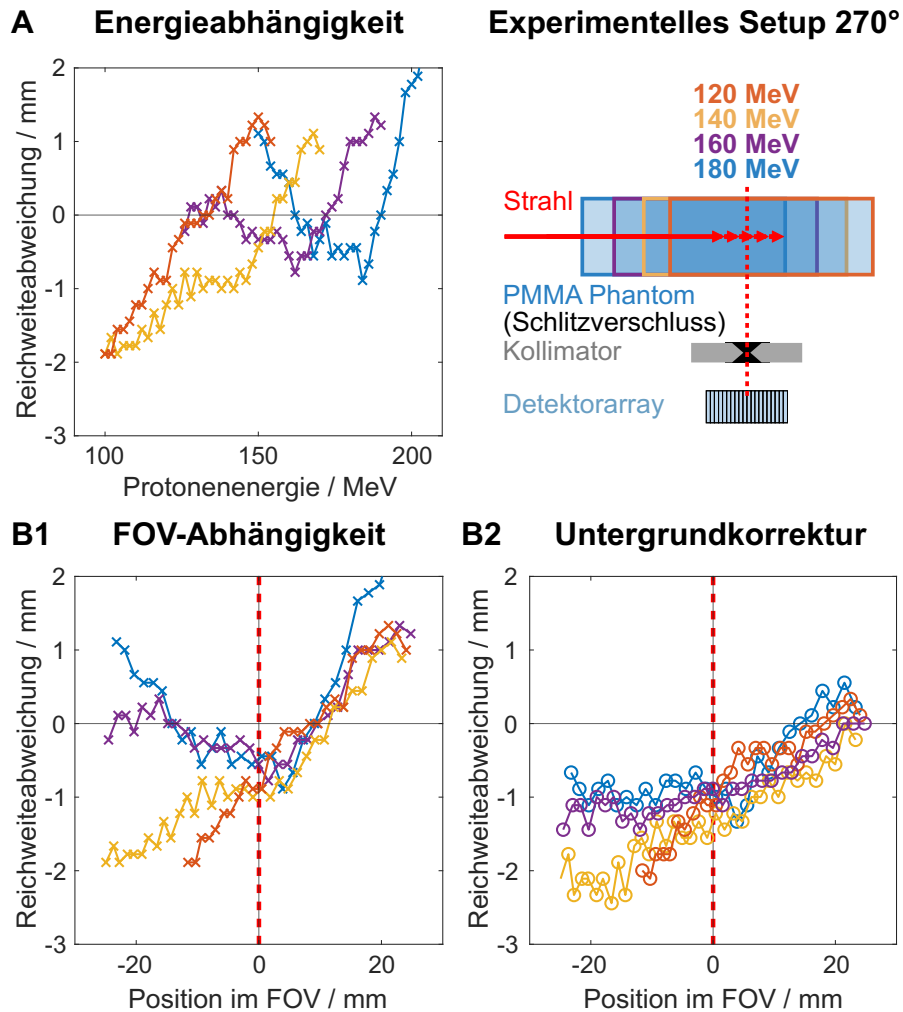


Abbildung 3.7: Einfluss von Protonenenergie (A) und FOV (B) auf die gemessenen Reichweiteabweichungen für Spots auf der zentralen Strahlachse unter 270° Gantrywinkel ohne (B1) und mit (B2) Subtraktion des Untergrundes. Die Verbindungslinien zwischen den Datenpunkten dienen der Visualisierung. Jeder Datenpunkt besitzt eine statistische Unsicherheit von $\sigma \approx 0,1$ mm.

Experiment 2: Zentrale Strahlachse 90°

Vor dem Hintergrund, dass die Reichweiteabweichungen im Patienten unter einem Gantrywinkel von 270° einen anderen Verlauf im FOV besitzen als unter 90° (dargestellt in Unterkapitel 3.5.3), wurde *Experiment 1* mit gleichen Bedingungen unter 90° wiederholt. Auch im Phantom zeigt sich ein anderes Verhalten der Reichweiteabweichungen im FOV unter einem Gantrywinkel von 90° (Abbildung 3.8). Dieses Verhalten könnte durch die leicht

veränderte Position des PGI-Trägerwagens, ein asymmetrisches Ansprechvermögen des unter 90° Gantrywinkel zur Strahlrichtung invertierten Detektorarrays und tatsächliche Unterschiede in der Bestrahlung erklärt werden. Im Bezug auf die Energieabhängigkeit der Reichweiteabweichungen im FOV zeigt sich eine ähnliche Verbesserung bei Abzug des Untergrundes wie bei einem Gantrywinkel von 270° (Abbildung 3.8 B2).

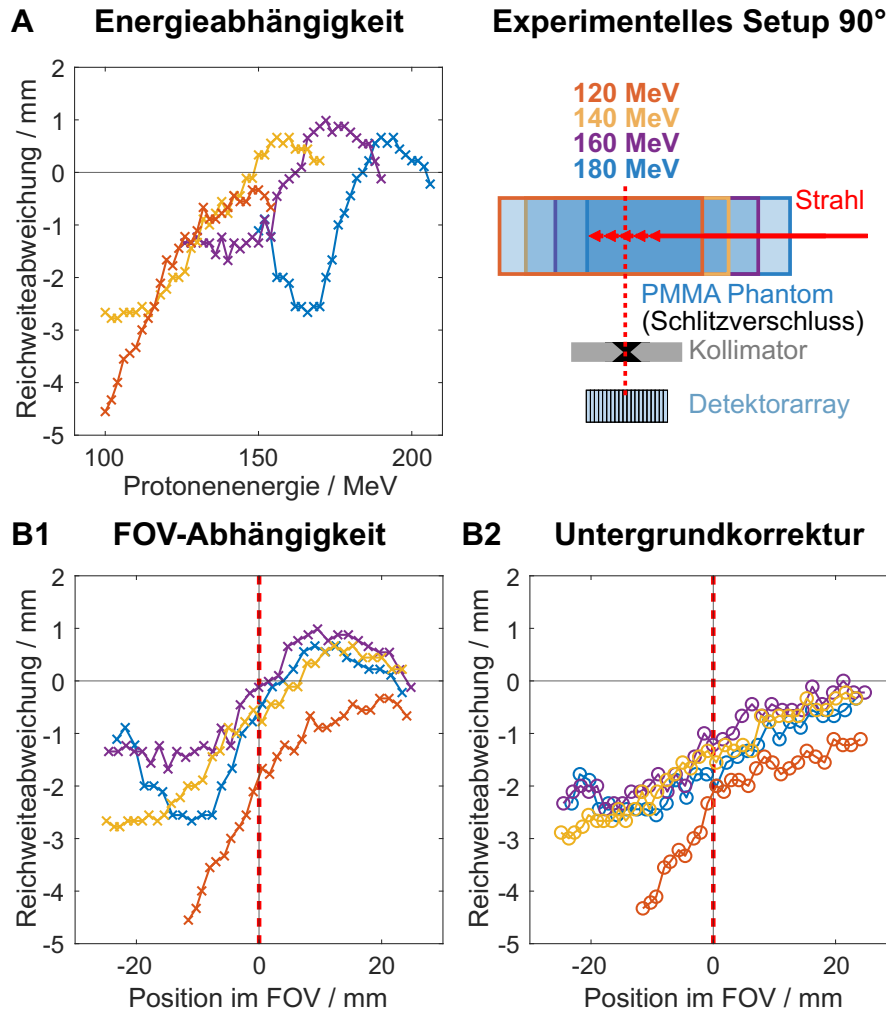


Abbildung 3.8: Einfluss von Protonenenergie (A) und FOV (B) auf die gemessenen Reichweiteabweichungen für Spots auf der zentralen Strahlachse unter 90° Gantrywinkel ohne (B1) und mit (B2) Subtraktion des Untergrundes. Die Verbindungslinien zwischen den Datenpunkten dienen der Visualisierung. Jeder Datenpunkt besitzt eine statistische Unsicherheit von $\sigma \approx 0,1$ mm. Der vertikal Versatz der Datenreihe des 120 MeV-Setups wird auf statistische Fluktuationen in der Phantompositionierung zurückgeführt.

Experiment 3: Zentrale Strahlachse 270° versus 90° – invertiertes Detektorarray

Die Hypothese eines asymmetrischen Ansprechvermögens des Detektorarrays lässt sich vergleichsweise einfach überprüfen, da das Detektorarray im Trägerwagen auch invertiert eingebaut werden kann. Der Verlauf der Reichweiteabweichungen für Bestrahlung mit nicht

invertiertem Detektorarray kann durch die Invertierung des Detektorarrays unter dem jeweils gegenüberliegenden Gantrywinkel nahezu reproduziert werden, wie exemplarisch anhand von verschiedenen Datenreihen im 180 MeV-Setup gezeigt (Abbildung 3.9 A). Der globale Versatz zwischen den Verläufen der beiden Gantrywinkel erklärt sich einerseits aus Positionierungsunsicherheiten des Phantoms und des PGI-Systems und andererseits aus den bereits in Unterkapitel 3.3.1 angesprochenen Unterschieden zwischen den Gantrywinkeln.

Qualitativ können mit diesem Experiment für das Verhalten der Reichweiteabweichungen im FOV tatsächliche Unterschiede in der Bestrahlung als Ursache ausgeschlossen und als Hauptursache ein asymmetrisches Ansprechvermögen des Detektorarrays angenommen werden. Durch Subtraktion des Untergrundes wird der Einfluss des unterschiedlichen Ansprechvermögens stark verringert, sodass die Verläufe der gemessenen Reichweiteabweichungen von invertiertem und nicht invertiertem Detektorarray je Gantrywinkel sogar annähernd gleich sind (Abbildung 3.9 B).

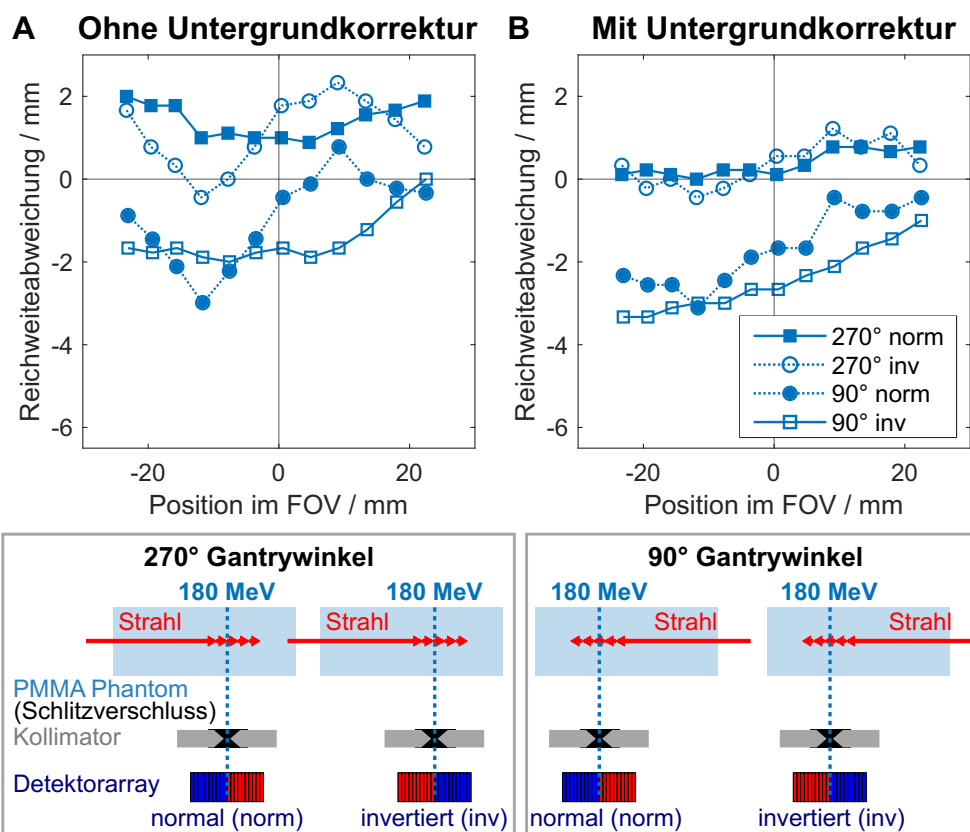


Abbildung 3.9: Einfluss der invertierten (inv) und normalen (norm), d. h. nicht invertierten, Ausrichtung des Detektorarrays auf die gemessenen Reichweiteabweichungen für Spots auf der zentralen Strahlachse unter 270° und 90° Gantrywinkel ohne (A) und mit (B) Subtraktion des Untergrundes. Das Experiment wurde analog zu Abbildung 3.7 und Abbildung 3.8 für das 180 MeV-Setup durchgeführt. Die Verbindungslinien zwischen den Datenpunkten dienen der Visualisierung. Jeder Datenpunkt besitzt eine statistische Unsicherheit von $\sigma \approx 0,1$ mm.

Experiment 4: Laterale Strahlachsen 270°

In klinischen Bestrahlungen ist das Bestrahlungsfeld lateral zur zentralen Strahlachse ausge dehnt. Zur Untersuchung des Einflusses dieser lateralen Spotpositionen auf die gemessenen PGI-Reichweiteabweichungen wurden im 160 MeV-Setup mehrere Layer mit einem quadratischen 3×3 Spotraster und Gitterabstand von 4 cm abgestrahlt. Anhand von Abbildung 3.10 lässt sich erkennen, dass die gemessenen Reichweiteabweichungen im Wesentlichen nicht von der lateralen Position in y -Richtung abhängig sind, dafür aber von der lateralen Position der Spots in z -Richtung, d. h. vom Abstand zwischen Kollimatorebene und Spotachse. Auch hierbei spielt vermutlich der Einfluss von Streustrahlung auf die PGI-Signale und das Ansprechvermögen des Detektorarrays eine Rolle.

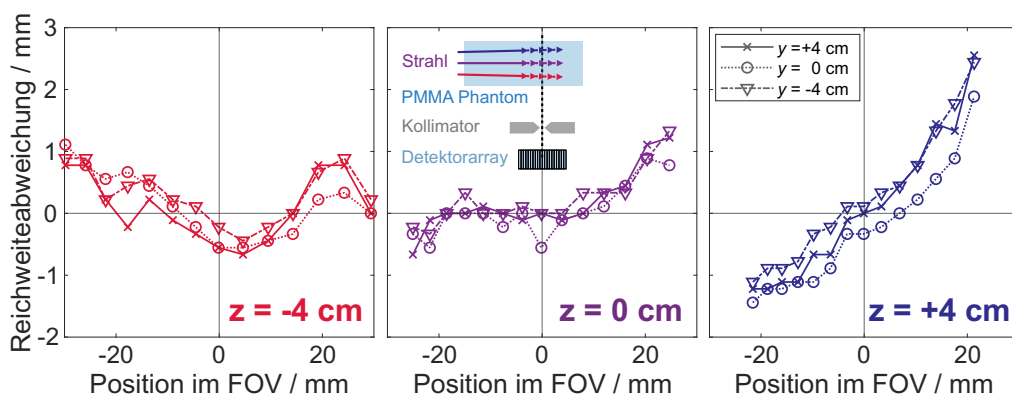


Abbildung 3.10: Einfluss der lateralen Spotposition auf die gemessenen Reichweiteabweichungen im FOV unter 270° Gantrywinkel. Das Experiment wurde analog zum 160 MeV-Setup in Abbildung 3.7 durchgeführt und durch acht laterale Spots in einem quadratischen Raster mit Gitterabstand von 4 cm erweitert. Jede Datenreihe entspricht den Ergebnissen für eine der Strahlachsen mit vertikaler Position $z = \{-4; 0; 4\}$ cm und horizontaler Position $y = \{-4; 0; 4\}$ cm. Die Verbindungslinien zwischen den Datenpunkten dienen der Visualisierung. Jeder Datenpunkt besitzt eine statistische Unsicherheit von $\sigma \approx 0,1$ mm.

3.5.2 Validierung der erweiterten Simulationssoftware

Damit das aktuelle PGI-System auch bei nicht perfekt horizontalen Einstrahlrichtungen verwendet werden kann, wurde die Simulationssoftware durch IBA erweitert. Dafür wurde die Projektion der PG-Emission entlang einer nicht parallel zur Kollimatorebene verlaufenden Strahlachse in die Simulation eingearbeitet. Die Anpassung berücksichtigt nun auch, dass die Strahlrichtung lateraler Spots nicht ganz parallel zum Zentralstrahl verläuft, wodurch sich auch für ein horizontales Bestrahlungsfeld kleinste Änderungen ergeben können. Die Rückprojektion der Profile zur Bestimmung der Reichweiteabweichung erfolgt in der erweiterten Simulationssoftware jedoch nur auf eine zum Kollimator parallele Strahlachse. Um diese Vereinfachung zu korrigieren, wurden die für einen Gantrywinkel W erhaltenen Reichweiteabweichungen in erster Näherung mit $1/|\sin(W)|$ auf die verkippte Strahlachse projiziert.

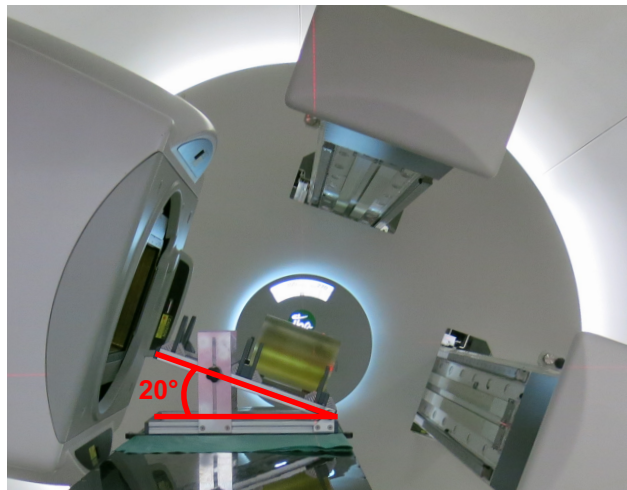


Abbildung 3.11: Winkelverstellbare Halterung für das homogene PMMA-Phantom zur Validierung der erweiterten Simulationssoftware unter einem Gantrywinkel von 290°

Zur Überprüfung der erweiterten Simulation wurden experimentelle Daten mit dem homogenen PMMA-Phantom aus Unterkapitel 3.3.1 unter verschiedenen Gantrywinkeln aufgenommen. Eine eigens dafür angefertigte winkelverstellbare Phantomhalterung ermöglichte die Bestrahlung des Phantoms unter bis zu 30° von der Horizontalen abweichenden Gantrywinkeln (Abbildung 3.11). Zielstellung war es, die PGI-Signale unter verkipptem Aufbau als Referenz für die erweiterte Simulation zu messen. Um die Gültigkeit der erweiterten Simulation zu validieren, wurden die Reichweiteabweichung zwischen den gemessenen und simulierten Signalen quantifiziert und mit den Unsicherheitseinflüssen verglichen.

Es wurden Messungen unter den Gantrywinkeln 270° – 300° und 60° – 90° mit jeweils 10° Abstand durchgeführt (in die jeweils entgegengesetzte Richtung käme es sonst zur Kollision zwischen Gantry und PGI-System). Der abgestrahlte Plan bestand aus zwölf Energieschichten von 125 – 180 MeV im Abstand von 5 MeV. Jede dieser Schichten beinhaltete neun Spots je 125 MU in einem zentrierten, quadratischen Raster mit einem Abstand von 4 cm. Herausforderung für jede Messposition war es, die Mitte des PMMA-Phantoms mittels des aufgebrauchten Röntgenmarkers genau im Strahlisozentrum zu positionieren, sodass das Phantom immer unter den gleichen Bedingungen bestrahlt wird. Zusätzlich musste für jeden Gantrywinkel die Verschiebung des Strahlisozentrums im Raum relativ zum PGI-System bei der Simulation berücksichtigt werden.

Die dargestellten Ergebnisse der Reichweiteabweichungen sind auf die Spots der zentralen Strahlachse und die Gantrywinkel 270° – 300° beschränkt (Abbildung 3.12). Qualitativ zeigen die anderen Spotpositionen und Gantrywinkel ein ähnliches Verhalten: (1) Unter der Annahme der bisherigen Simulation, d. h. ohne Berücksichtigung der Verkipfung der Kamera relativ zum Phantom, treten zwischen den simulierten PGI-Signalen und der gemessenen Referenz Reichweiteabweichungen auf, welche die zu detektierenden Reichweiteabweichungen im Patienten überdecken würden. (2) Die Abweichungen zwischen der erweiterten Simulation

und der gemessenen Referenz bewegen sich innerhalb der Effekte im FOV, die bereits in Unterkapitel 3.5.1 beobachtet wurden. Damit ist gezeigt, dass die erweiterte Simulation notwendig und innerhalb der erwarteten Genauigkeit liegt.

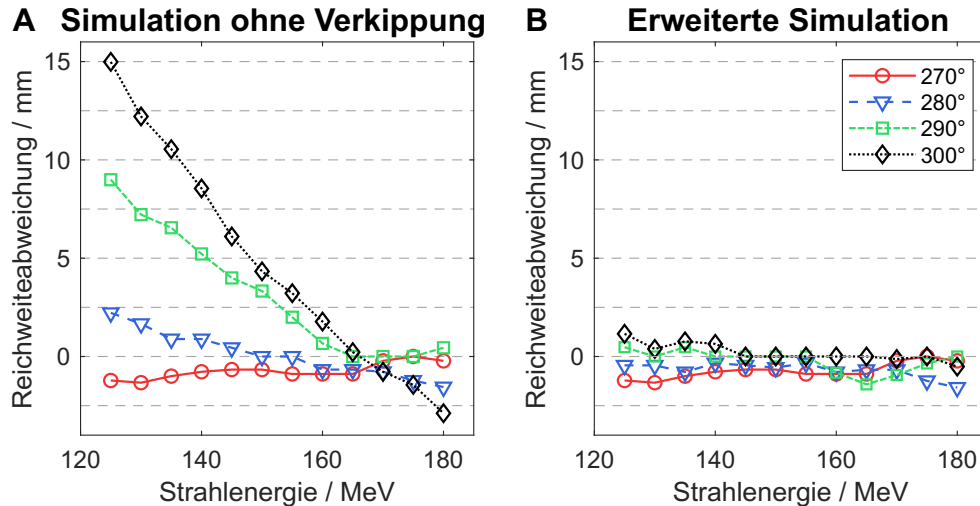


Abbildung 3.12: Validierung der erweiterten Simulation unter verschiedenen Gantrywinkeln im Vergleich zu gemessenen PGI-Profilen. Dargestellt sind die Reichweiteabweichungen unter der Annahme der bisherigen Simulation ohne die Berücksichtigung der Verkippung (A) und mittels erweiterter Simulation (B). Die Verbindungslinien zwischen den Datenpunkten dienen der Visualisierung. Jeder Datenpunkt besitzt eine statistische Unsicherheit von $\sigma \approx 0,15$ mm.

3.5.3 Abhängigkeit von der Tumorentität

Die Untersuchung der PGI-Simulationsgenauigkeit im Patienten ist mangels Grundwahrheit nur eingeschränkt möglich. Hier soll der typische Verlauf der Reichweiteabweichung für die monitorierten Prostata- und Kopf-Hals-Tumor-Patienten beim Vergleich zwischen PGI-Messung und der auf dem DECT-kCT-Scan basierenden PGI-Simulation dargestellt und erörtert werden. Durch die Verwendung der kCT-Scans können anatomische Änderungen als Ursache für Reichweiteabweichungen minimiert werden.

Prostata-Tumor-Patienten

Bei der Evaluation der PGI-Daten für Prostata-Tumor-Patienten konnten systematische Abhängigkeiten der Reichweiteabweichung beobachtet werden. Insbesondere zeigte sich bei einem Gantrywinkel von 270°, dass die gemessenen Reichweiteabweichungen für im FOV distale Spots systematisch kleiner sind als für im FOV proximale Spots (Abbildung 3.13 A). Dieser Verlauf wird jedoch unter einem Gantrywinkel von 90° nicht reproduziert. Daher liegt die Vermutung nahe, dass auch bei den PGI-Daten der Prostata-Tumor-Patienten das Ansprechvermögen der Detektoren einen relevanten Einfluss auf die Reichweiteabweichungen zu haben scheint, wie bereits durch kontrollierte Experimente im Phantom geschlossen wird. Auf der Suche nach einem Erklärungsmodell für den widersprüchlichen Verlauf der

PGI-Reichweiteabweichungen im Patienten, welches potentiell zur Verbesserung der Genauigkeit der PGI-Methode im Patienten führen könnte, wurden auf Basis der aufgenommenen Phantom- und Patientendaten heuristische Korrekturen etabliert.

Der erste Versuch einer heuristischen Korrektur der durch das Detektoransprechvermögen verursachten Unterschiede zwischen den Gantrywinkeln basiert auf dem im PMMA-Phantom aufgenommenen Verlauf der Reichweiteabweichungen über das FOV (vgl. Unterkapitel 3.5.1). Im Speziellen wurden nur die Datenreihen für das 180 MeV-Setup als Korrektur angewandt (vgl. Abbildung 3.16, Seite 50), welches für Prostata-Tumor-Patienten am ehesten die Ausrichtung des FOV relativ zum Strahl widerspiegelt. Die Anwendung der *FOV-Korrektur* zeigt, dass sich der Verlauf der Reichweiteabweichungen in Abhängigkeit vom FOV für die Gantrywinkel 270° und 90° annähert (Abbildung 3.13 B). Das asymmetrische Ansprechvermögen der Detektoren kann somit in erster Näherung für die Messung im Patienten korrigiert werden, allerdings bewirkt die Korrektur eine noch stärkere Abhängigkeit der Reichweiteabweichungen von der Position im FOV, die im Folgenden näher beleuchtet werden soll.

Nach Anwendung der FOV-Korrektur stimmen die Verläufe der Reichweiteabweichungen im FOV noch nicht perfekt überein, da es einige Patienten gibt, deren Verlauf verschoben scheint (Abbildung 3.14 A). Dies könnte darauf hindeuten, dass dem genannten Verlauf eine Abhängigkeit von der Protonenenergie zugrunde liegt, da aufgrund anatomischer Unterschiede (insbesondere der Zielvolumengröße und des Patientenumfangs) für jeden Patienten ein anderes Spektrum an Protonenenergien zum Einsatz gekommen ist. Die Darstellung

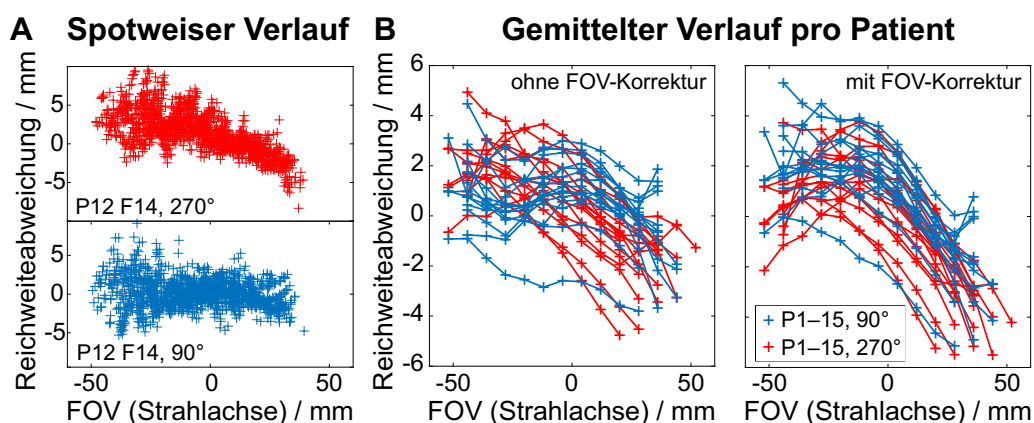


Abbildung 3.13: Gantrywinkelabhängige Darstellung des Verlaufs der Reichweiteabweichung zwischen PGI-Messung und kCT-Scan basierter PGI-Simulation über die Position des Bragg-Peaks im FOV (hier Projektion auf die Strahlachse, d. h. effektiv mit einem Faktor zwei skaliert). (A) Spotweise Verteilung der Reichweiteabweichungen für die repräsentative Fraktion F14 des Patienten P12 unter Gantrywinkel 90° und 270° . (B) Über Intervalle von je 8 mm entlang des FOV gemittelter Verlauf der Reichweiteabweichungen aller aufgenommenen Fraktionen pro Patient und Gantrywinkel ohne (links) und mit (rechts) FOV-Korrektur (vgl. Abbildung 3.16). Die Verbindungslinien zwischen den Datenpunkten dienen der Visualisierung.

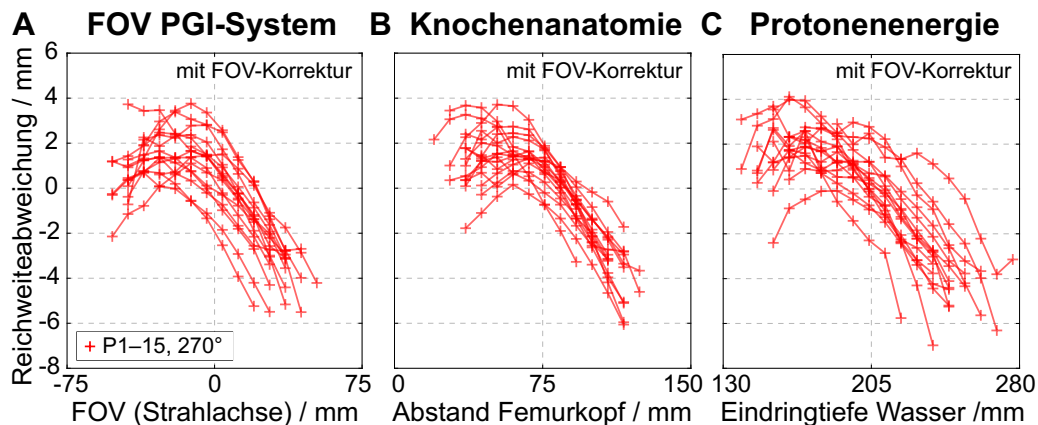


Abbildung 3.14: Verlauf der Reichweiteabweichung zwischen PGI-Messung und kCT-Scan basierter PGI-Simulation über (A) die Position des Bragg-Peaks im FOV (hier Projektion auf die Strahlachse, d. h. effektiv mit einem Faktor zwei skaliert), (B) den Abstand des Bragg-Peaks zu einer manuell definierten Referenzebene am Femurkopf (vgl. Abbildung 3.16) und (C) die wasseräquivalente Eindringtiefe (Protonenenergie). In den dargestellten Datenpunkten wurden die Reichweiteabweichungen aller aufgenommenen Fraktionen pro Patient unter dem Gantrywinkel 270° entlang des FOV über Intervalle von je 8 mm gemittelt. Die Verbindungslinien zwischen den Datenpunkten dienen der Visualisierung. Auf alle Reichweiteabweichungen wurde die FOV-Korrektur angewandt (vgl. Abbildung 3.16).

der Reichweiteabweichungen über die wasseräquivalente Eindringtiefe (Protonenenergie) der Spots bringt jedoch keine deutliche Übereinstimmung der Verläufe verschiedener Patienten (Abbildung 3.14 C). Hingegen zeigte sich bei der Untersuchung von Zusammenhängen zwischen den Reichweiteabweichungen und der im Pelvis-Bereich ausgeprägten und von den Protonen durchdrungenen Knochenstruktur ein konsistentes Verhalten über alle Patienten. Hierbei wurde der Abstand in Strahlrichtung zwischen Bragg-Peak-Position und einer manuell definierten Referenzebene am Femurkopf (vgl. Abbildung 3.16, Seite 50) als Parameter mit guter Korrelation zu den Reichweiteabweichungen identifiziert (Abbildung 3.14 B). Es wurde nur der Verlauf der Reichweiteabweichungen für dem Gantrywinkel 270° betrachtet, da dieser aufgrund der dort durchgeführten Kalibrierung des PGI-System verlässlicher ist. Bei qualitativer Betrachtung der Verläufe der Reichweiteabweichungen in Abhängigkeit der drei untersuchten Parametern ist die Streuung zwischen verschiedenen Patienten gegenüber dem Abstand zum Femurkopf am geringsten und zeigt die größte Symmetrie.

Mögliche Erklärungen für eine Abhängigkeit der Reichweiteabweichungen von der Knochenstruktur nach deren Relevanz sortiert sind:

- (1) Ungenauigkeiten in der PGI-Simulationssoftware z. B. bei der Coulomb-Streuung und den zugrundeliegenden Wechselwirkungsquerschnitten und die Nichtbeachtung anderer Sekundärteilchen (vgl. Sterpin et al., 2015),
- (2) Ungenauigkeiten in der Materialzusammensetzung menschlichen Gewebes nach Woodard und White (1986), welches die PG-Emission beeinflusst,

- (3) inhärente Limitationen in der Bestimmung der Reichweiteabweichung bei großen Dichteunterschieden (vgl. Priegnitz et al., 2015), hier Knochen vs. Weichgewebe,
- (4) Ungenauigkeit in der SPR-Vorhersage des Gewebes.

Die ersten beiden Punkte können anhand der vorliegenden Daten nicht genauer untersucht werden. Hierzu wären systematische Studien in einem anthropomorphen Pelvis-Phantom und ggf. biologischen Gewebe notwendig, die außerhalb des Rahmens dieser Arbeit liegen. Außerdem müsste genauer untersucht werden, ob es notwendig ist, aufgrund inhärenter Limitationen Spots von der Analyse auszuschließen (vgl. Janssens et al., 2016). Bisher gab es dazu jedoch noch keine Anhaltspunkte. Die Genauigkeit der SPR-Vorhersage im Gewebe kann – nicht zuletzt aufgrund der Verwendung des in zahlreichen Studien systematisch validierten DirectSPR-Verfahrens (Wohlfahrt & Richter, 2020) – als Hauptursache ausgeschlossen werden: (a) Die Hypothese einer systematischen SPR-Abweichung würde zu einem eindeutigen Trend aller Patienten über die wasseräquivalente Eindringtiefe führen, welcher nicht beobachtet wurde (Abbildung 3.14 C). (b) Die Hypothese, dass das SPR beispielsweise nur für Knochen unterschätzt wird, kann ausgeschlossen werden, da sich keine Korrelation mit der Dicke oder dem integralen SPR von Knochengewebe für die Reichweiteabweichung der Spots gezeigt hat. Außerdem wäre in diesem Fall ein konstanter Versatz der Reichweiteabweichungen zu erwarten, da die Spots in der Regel nicht direkt im Knochengewebe gestoppt werden.

Kopf-Hals-Tumor-Patienten

Im Vergleich zu den Prostata-Tumor-Patienten zeichnen sich die Kopf-Hals-Tumor-Patienten durch eine deutlich geringere Eindringtiefe der Protonen und eine andere Anatomie, bspw. weniger zu durchdringendes Knochengewebe und eine veränderte elementare Zusammensetzung des Weichgewebes, aus. Diese Gründe könnten dazu beitragen, dass für die Kopf-Hals-Tumor-Patienten der für die Prostata-Tumor-Patienten beschriebene Trend nicht beobachtet wurde, wie sich an exemplarischen Behandlungsfeldern unter 90° und 270° erkennen lässt (Abbildung 3.15). Bestrahlungsfelder unter nicht horizontalen Gantrywinkeln wurden hierbei nicht betrachtet, da Unsicherheiten der erweiterten Simulationssoftware den zu untersuchenden Verlauf überlagern könnten. Im Vergleich zu den hypofraktionierten Bestrahlungen der Prostata-Tumor-Patienten ist für die normal fraktionierten Kopf-Hals-Bestrahlungen, bei denen die Dosis meist auf 2–3 gleichgewichtete Bestrahlungsfelder aufgeteilt ist, das statistische Rauschen in den PGI-Messungen erkennbar größer. Hinzu kommen systematische Ungenauigkeiten in der Bestimmung der Reichweiteabweichung bei laterale Dichtegradierten an luftgefüllten Kavitäten und Hohlräumen (Priegnitz et al., 2016b). Außerdem sind die Zielvolumina innerhalb der Kopf-Hals-Kohorte sehr unterschiedlich, was sich anhand der Lage des Karzinoms und der Anordnung der Bestrahlungsfelder zeigt (Tabelle 3.3). Einen weiteren Unterschied stellt die Verwendung eines Range-Shifters am Strahlaustritt der Gantry dar, dessen Einfluss auf die PGI-Signale noch nicht systematisch untersucht wurde.

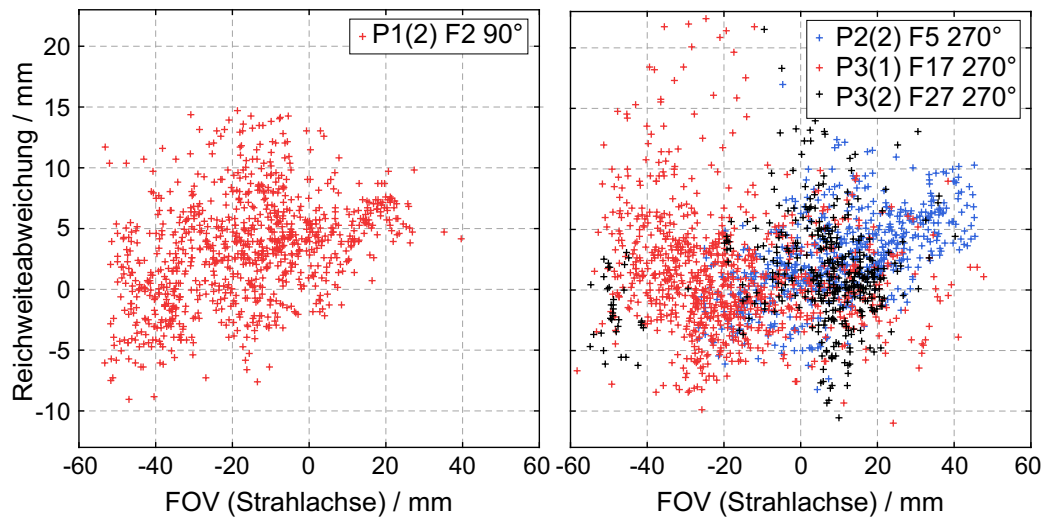


Abbildung 3.15: Gantrywinkelabhängige Darstellung der Reichweiteabweichung zwischen PGI-Messung und kCT-Scan basierter PGI-Simulation über der Position des Bragg-Peaks im FOV (hier Projektion auf die Strahlachse, d. h. effektiv mit einem Faktor zwei skaliert) für exemplarische Bestrahlungsfractionen (F) unter 90° und 270° .

3.5.4 Schlussfolgerungen

In diesem Unterkapitel wurde die Genauigkeit der PGI-Simulation im Bezug auf die Messung mit dem vorliegenden PGI-Detektionssystem in einem homogenen PMMA-Phantom und im Patienten untersucht. Die erörterten Zusammenhänge beruhen auf den seit über sieben Jahren gesammelten Daten und Erfahrungen mit dem PGI-System durch Patienten- und Phantom-Messungen, auch schon vor Beginn der vorliegenden Arbeit.

Im Falle des PMMA-Phantoms wurde ein von der Position des Bragg-Peaks im FOV abhängiger Verlauf der Reichweiteabweichungen festgestellt. Hierbei stellt eine Restvariabilität des Ansprechvermögens der Detektoren trotz aufwendiger Energiekalibrierung die Hauptursache für den nicht-konstanten Verlauf der Reichweiteabweichungen im FOV der PGI-Kamera dar. Darüber hinaus sind die Reichweiteabweichungen im proximalen Teil des FOV zusätzlich von der Protonenenergie abhängig, was maßgeblich auf Unterschiede in der Entstehung und Verteilung von sekundärer Strahlung im Phantom zurückzuführen ist. Hervorzuheben ist jedoch, dass für Spots unterschiedlicher Energie in der Mitte des FOV (hier 120, 140, 160 und 180 MeV) – im Rahmen der statistischen Unsicherheit der PGI-Messung ($\sigma \approx 0,1$ mm) und der Genauigkeit beim Verfahren der Patientenliege (erfahrungsgemäß $\sigma \approx 0,1-0,2$ mm) – keine Abweichungen der detektierten Reichweiteabweichungen beobachtet wurden (vgl. Abbildung 3.7). Somit konnte die Gültigkeit des in der Kalibrierung des PGI-Systems ermittelten globalen Versatzes (vgl. Tabelle 3.1) für eine Bandbreite an relevanten Protonenenergien gezeigt werden.

Die detektierten Reichweiteabweichungen zu der PGI-Simulation auf den kCT-Scans der Patienten ergaben bislang kein konsistentes Bild. Bei den Prostata-Tumor-Patienten zeigte sich ein deutlich von den Phantommessungen abweichender Verlauf der Reichweiteabweichungen, was auf weitere systematische Einflüsse hinweist. Die Unterschiede zwischen Bestrahlungsfeldern unter 270° und 90° konnten mithilfe der Phantomdaten in erster Näherung reduziert werden, sodass die Annahme getroffen werden kann, dass dieses Verhalten im Phantom auf den Patienten übertragbar ist. Der verbleibende Verlauf der Reichweiteabweichungen lässt sich am besten durch eine Abhängigkeit von der pelvisspezifischen Knochenstruktur erklären. Dafür werden momentan Defizite in der PGI-Simulation als Hauptursache angenommen, weil die Signaturen untypisch für ein Problem seitens der umfassend validierten SPR-Vorhersage mit DirectSPR sind.

Die Datengrundlage der Kopf-Hals-Tumor-Patienten ist deutlich geringer und inhomogener (mehrere Gantrywinkel, unterschiedliche Bestrahlungsgebiete) als bei den Prostata-Tumor-Patienten, wodurch die Untersuchung systematischer Effekte kaum möglich ist. Allerdings weist der Verlauf der Reichweiteabweichungen von den untersuchten Bestrahlungsfeldern keine Ähnlichkeit zu den Prostata-Tumor-Patienten auf. Dadurch wird die These gestützt, dass nicht das DirectSPR-Verfahren systematische Abweichungen verursacht, sondern die PGI-Simulation. Auf Grundlage aller hier aufgeführten Erkenntnisse werden in den anschließenden Untersuchungen dieser Arbeit die nachfolgenden heuristischen Korrekturen angewandt (Tabelle 3.4).

Tabelle 3.4: Überblick der heuristischen Korrekturen für die PGI-Auswertung der jeweiligen Untersuchung im Rahmen dieser Arbeit.

Untersuchung	Korrekturen
(1) Validierung der CT-basierten Reichweitevorhersagegenauigkeit (Kap. 4)	keine Korrektur (Heuristischer Unsicherheitsbeitrag)
(2) Erkennung anatomischer Veränderungen bei Prostata-Tumor-Bestrahlungen (Kap. 5)	FOV-Korrektur & Pelvis-Korrektur
(3) Erkennung anatomischer Veränderungen bei Kopf-Hals-Tumor-Bestrahlungen (Kap. 5)	keine Korrektur
(4) Vergleichsstudie PGI vs. PGS mittels anthropomorphem Kopfphantom (Kap. 6)	keine Korrektur

(1) Für die Bestimmung der absoluten Genauigkeit der Reichweitevorhersage in Prostata-Tumor-Patienten (Kapitel 4) wurden keine Korrekturen angewandt und stattdessen ein *heuristischer Unsicherheitsbeitrag* anhand der Variation der im Patienten gemessenen Reichweiteabweichungen im FOV abgeleitet. Es scheint bis zur abschließenden Klärung der Ursache des systematischen Verlaufs angemessen zu sein, einen konservativen Ansatz zu verfolgen, da die Reichweiteabweichungen den Endpunkt für die Bestimmung der Vorhersagegenauigkeit darstellen.

(2) Um die Detektierbarkeit von anatomischen Veränderungen mittels PGI in Prostata-Tumor-Patienten zu quantifizieren (Kapitel 5), wurden zwei heuristische Korrekturen eingeführt: (a) eine *FOV-Korrektur* auf Grundlage des 180 MeV-Setups im PMMA-Phantom und (b) eine *Pelvis-Korrektur* aufgrund der vermuteten Abhängigkeit der Reichweiteabweichungen vom Abstand zur pelvisspezifischen Knochenstruktur (Abbildung 3.16). Bei den beiden Korrekturen wird angenommen, dass der Verlauf der Reichweiteabweichungen im FOV oder in Abhängigkeit vom Abstand zum Femurkopf konstant sein müsste. Genauer gesagt, werden die entsprechenden Bilddaten vom Phantom (mit gemessenem SPR-Werten) und vom Patienten (mithilfe von DirectSPR bestimmten SPR-Werten) als SPR-Grundwahrheit angenommen. Da es hier im Gegensatz zur Bestimmung der absoluten Genauigkeit der Reichweitevorschau um die Detektion relativer Veränderungen geht – vergleichbar mit der Reichweiteabweichung zwischen zwei gemessenen PGI-Signalen – ist die Korrektur vermuteter systematischer Simulationsungenauigkeiten zielführend. Die FOV-Korrektur wurde durch eine kubische Anpassungsfunktion angenähert, die dabei aber keine physikalische Bedeutung trägt. Für die Pelvis-Korrektur wurde in jedem Patienten eine Referenzebene senkrecht zur zentralen Strahlachse definiert, welche den Femurkopf tangential berührt. Der Abstand zwischen dieser Ebene und der Bragg-Peak-Position in Strahlrichtung wurde als Parameter für die Korrektur angewandt. Hierbei sind die Daten aller Patienten zum Einsatz gekommen, wenngleich einzelne wenige Fraktionen mit vermuteten intrafraktionellen Änderungen und damit großer Streuung der Reichweiteabweichung ausgeschlossen wurden, da diese die gewählte lineare Anpassung sonst beeinflusst hätten.

(3) Um die Detektierbarkeit von anatomischen Veränderungen mittels PGI in Kopf-Hals-Tumor-Patienten zu quantifizieren (Kapitel 5), wurden *keine* der Korrekturen angewandt. Die oben vorgeschlagene FOV-Korrektur kann für die Kopf-Hals-Tumor-Patienten nicht zum Einsatz kommen, da (a) in den Bestrahlungsfeldern ein Range-Shifter verwendet wurde, (b) auch Bestrahlungen unter nicht-horizontalen Gantrywinkeln monitoriert wurden und (c) in manchen Bestrahlungen ein veränderter Abstand zwischen Kollimator und Isozentrum eingestellt werden musste. Zu den genannten Einflussfaktoren müssten noch weitere Untersuchungen in Phantomen erfolgen, die über den Rahmen dieser Arbeit hinausgehen.

(4) Bei der Untersuchung des PGI-Systems im anthropomorphen Kopfphantom (Kapitel 6) wurde für die Bestimmung der absoluten Genauigkeit gegenüber der Grundwahrheit *keine* der Korrekturen angewandt. Für die Bestimmung der relativen Genauigkeit gegenüber Szenarien mit lokalen Reichweiteverschiebungen wurde hingegen die im Referenzfall gemessenen Reichweiteabweichungen von denen der Szenarien abgezogen (effektiv also Messung gegen Messung verglichen). Dadurch wird das Potential zur Detektion relativer Reichweiteverschiebungen nicht durch mögliche systematische Abweichungen der PG-Simulationen überlagert.

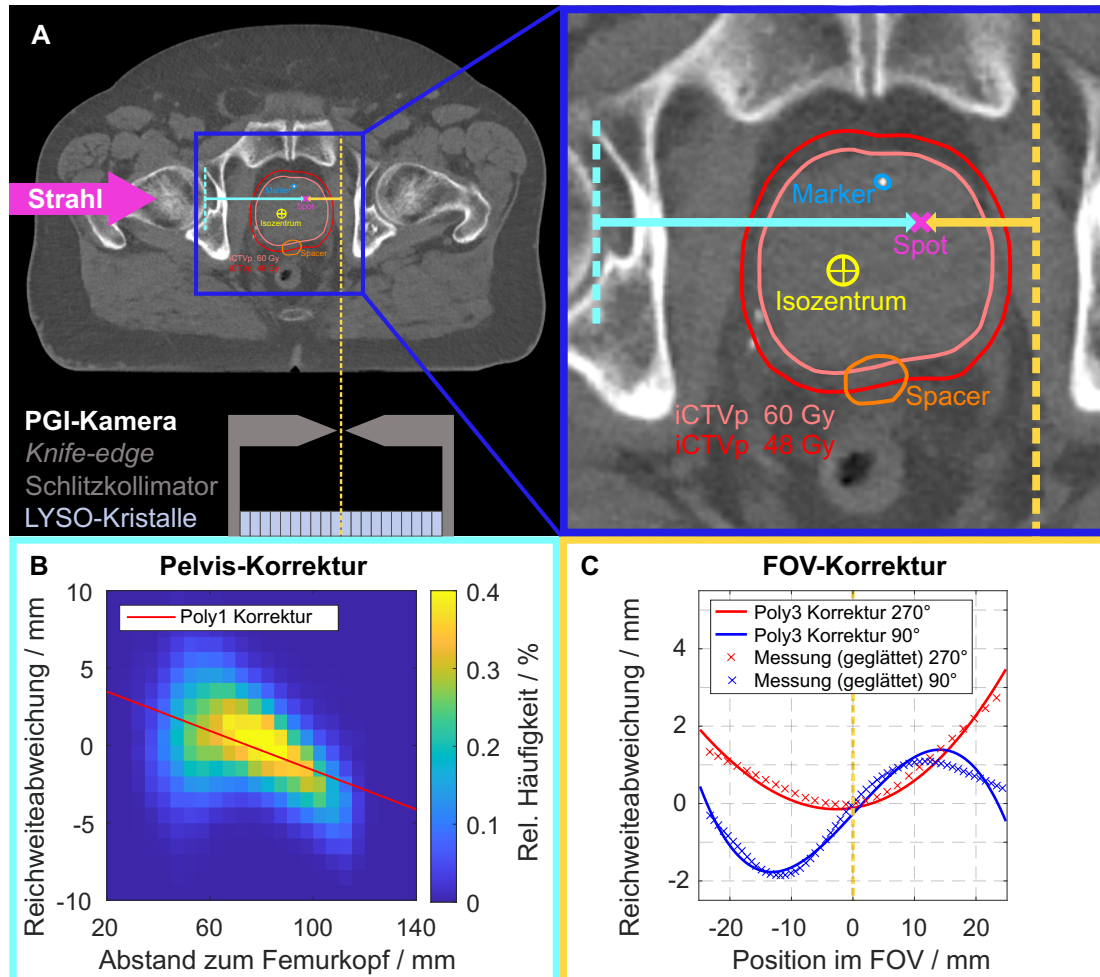


Abbildung 3.16: Heuristische Korrekturen der detektierten Reichweiteabweichung in Prostata-Tumor-Patienten. (A) Exemplarische kCT-Schicht mit schematischer Visualisierung der Korrekturparameter für einen exemplarischen Spot. Die CT-Schicht zeigt die Konturen des internen, primären Zielvolumens (iCTVp) sowie eines implantierten Marker (PolyMark™, CIVCO Radiotherapy, Orange City, U.S.) zur Positionierungskontrolle und den Spacerballon zum Schutz des Rektums. (B) Verteilung der Reichweiteabweichungen in Abhängigkeit zur Knochenstruktur, die durch eine manuell definierte Ebene senkrecht zur Strahlrichtung und tangential zum Femurkopf repräsentiert wird. Die heuristische Korrektur basiert auf einer linearen Anpassungsfunktion (Polynom ersten Grades: Poly1) über die ganze Prostata-Kohorte. (C) Abhängigkeit der Reichweiteabweichungen von der Position des Spots im FOV für den jeweiligen Gantrywinkel. Die FOV-Korrektur basiert auf einer kubischen Anpassungsfunktion (Polynom dritten Grades: Poly3), dessen Grundlage die im PMMA-Phantom gemessenen Daten für das 180 MeV-Setup sind (vgl. Abbildung 3.7 und Abbildung 3.8).

4 Validierung der CT-basierten Reichweitevorhersage mittels PGI

Das OncoRay war maßgebend an der Entwicklung von DECT-basierten Verfahren für die Reichweitevorhersage von Protonen im Patienten beteiligt und hat diese erstmals zur routinemäßigen klinischen Anwendung gebracht (Peters et al., 2022; Wohlfahrt et al., 2020). Dadurch besteht die Möglichkeit, mehrere klinisch verifizierte Verfahren zur Reichweitevorhersage im Rahmen von Patientenbehandlungen an der UPTD zu validieren.

Der Vergleich von Reichweitevorhersagemethoden ohne eine Grundwahrheit wurde bereits *in silico* durchgeführt und zeigte deutliche Unterschiede zwischen dem klassischen Standard-HLUT-Verfahren und dem DECT-basierten DirectSPR-Verfahren (Wohlfahrt et al., 2017b). Die Überlegenheit von DECT-basierten Verfahren – insbesondere von DirectSPR – in der Bestimmung des SPR und der daraus folgenden Protonenreichweite konnte ebenfalls in umfassenden Arbeiten in homogenen (Hünemohr et al., 2013) und heterogenen Phantomen (Wohlfahrt et al., 2018) sowie in biologischen Geweben von Tieren demonstriert werden (Bär et al., 2018; Möhler et al., 2018; Taasti et al., 2017; Xie et al., 2018). Eine *in vivo* Validierung im Patienten wurde jedoch bislang noch nicht erbracht werden, da es an entsprechenden Möglichkeiten gefehlt hat, das genaue SPR bzw. die tatsächliche Protonenreichweite im Patienten zu bestimmen. Durch PGI-basierte Reichweiteverifikation kann diese Lücke geschlossen werden.

In den Untersuchungen von Nenoff et al. (2017) konnte gezeigt werden, dass das angewandte PGI-System die absolute Reichweite von Protonen in einem heterogenen Kopfphantom millimetergenau bestimmen kann. Beim Vergleich mit den deutlich größeren Unsicherheiten in der Vorhersage der Reichweite (Unterkapitel 2.3.1) ist es daher naheliegend, PGI für die Validierung von Reichweitevorhersagemethoden im Patienten anzuwenden. Hierzu musste zunächst der Einfluss von zusätzlichen Reichweiteunsicherheiten durch anatomische Veränderungen des Patienten im Vergleich zur Bestrahlungsplanung konzeptionell minimiert werden. Das Ergebnis ist eine neuartige Methode zur systematischen Validierung der Reichweitevorhersage in der Protonentherapie mittels PGI. In diesem Kapitel werden die Ergebnisse aus einer bereits veröffentlichten Publikation vorgestellt (Berthold et al., 2021).

4.1 Konzept der Validierung

Der angestrebte Hauptanwendungszweck von PGI-Reichweiteverifikation in der Protonentherapie ist die unabhängige Kontrolle der Bestrahlung. Dazu detektiert das System Reichweiteabweichungen gegenüber der auf dem pCT-Scan berechneten Protonenreichweite. Das Ziel dieser Validierungsstudie hingegen ist die systematische Bestimmung und Gegenüberstellung der Genauigkeit verschiedener klinisch genutzter Verfahren der CT-basierten Reichweitevorhersage mittels PGI. Hierbei besteht die Herausforderung darin, dass die Ursachen für Reichweiteunsicherheiten vielfältig sind, weshalb detektierte Reichweiteabweichungen im Patienten nicht eindeutig auf Unsicherheiten in der Reichweitevorhersage zurückgeführt werden können. Damit interfraktionelle anatomische Veränderungen ausgeschlossen werden können, basiert die Validierung der Reichweitevorhersage auf *in-room* kCT-Scans vom bereits positionierten Patienten. Beim kCT-Scan ist lediglich die Bildgebungs-dosis um 60% geringer als im pCT-Scan, was jedoch auch aufgrund der iterativen Bildrekonstruktion und komplexer Rauschfilterung keinen Einfluss auf die Reichweitevorhersage hat (Wohlfahrt, 2018). Verglichen werden insgesamt drei Verfahren, die auf dem durchgeführten DECT-Scan basieren (vgl. Unterkapitel 2.2.1): (1) ein Standard-HLUT-Verfahren und (2) ein auf DirectSPR basierendes adaptiertes HLUT-Verfahren, deren Grundlage ein aus dem DECT-Scan generierter monoCT-Scan ist, sowie (3) das DirectSPR-Verfahren (vgl. Unterkapitel 2.2.1).

Durch die Anwendung von PGI während PBS-Bestrahlungen kann die Detektion von Reichweiteabweichungen spotgenau, also mit hoher örtlicher Auflösung im Patienten erfolgen. In Abhängigkeit von der im Spot applizierten Protonenanzahl p , hat die mit PGI gemessene Reichweiteabweichung r eine hohe statistische Streubreite (Präzision $\sigma_r \propto 1/\sqrt{p}$, vgl. Perali et al., 2014: $\sigma_r \approx 2$ mm für $p = 1,4 \cdot 10^8$). Daher ist eine spotweise Beurteilung der Reichweitevorhersage nur eingeschränkt möglich, sodass die gewichtete mittlere Reichweiteabweichung μ (mit inverser Varianz gewichtet) über alle monitorierten Bestrahlungsfelder f und alle Spots pro Bestrahlungsfeld $i(f)$ ein geeigneteres Maß für die Genauigkeit der Reichweitevorhersage darstellt:

$$\mu = \frac{\sum_f \sum_{i(f)} r_i / \sigma_{r,i}^2}{\sum_f \sum_{i(f)} 1 / \sigma_{r,i}^2} \approx \frac{\sum_f \sum_{i(f)} r_i \cdot p_i}{\sum_f \sum_{i(f)} p_i}. \quad (4.1)$$

Die untersuchte Patientenkohorte stellt einen Teil der mit PGI monitorierten Prostata-Tumor-Patienten dar. Konkret handelt es sich um die Bestrahlungsfelder der Patienten P3(2)–7 (vgl. Tabelle 3.2), also insgesamt 30 Fraktionen. Aufgrund der Notwendigkeit der Kalibrierung des PGI-Systems unter einem Gantrywinkel von 270° (vgl. Unterkapitel 3.3.1) und der damit verbundenen Genauigkeit, wurden nur Bestrahlungsfelder unter diesem Gantrywinkel für die Validierung berücksichtigt.

4.2 Gesamtabschätzung der Validierungsunsicherheit

Umgang mit Unsicherheiten

Um aus den Beiträgen einzelner Unsicherheiten u_i auf eine Gesamtunsicherheit u_{total} zu schließen, wurde der Ansatz laut GUM (engl. *Guide to the expression of uncertainty in measurement*, Joint Committee for Guides in Metrology / Working Group 1, 2008) angewandt. Bei diesem werden Typ-A-Unsicherheiten, welche als Standardabweichungen aus der Messwiederholung gewonnen werden, und Typ-B-Unsicherheiten, die auf einer Abschätzung gegebener Informationen beruhen, miteinander zu einer Gesamtunsicherheit kombiniert. Hierfür werden Typ-B-Unsicherheiten entsprechend der angenommenen Wahrscheinlichkeitsdichtefunktion (WDF) als statistische Unsicherheit interpretiert. In dieser Arbeit wurden als WDF eine Normalverteilung (insbesondere für alle Typ-A-Unsicherheiten), eine Dreiecksverteilung oder eine Rechteckverteilung mit entsprechenden Gewichtungsfaktoren a_i von 1, $\frac{1}{\sqrt{6}}$ bzw. $\frac{1}{\sqrt{3}}$ verwendet. Die Gewichtungsfaktoren ergeben sich aus der Wurzel der Varianz für die jeweiligen WDF. Die Gesamtunsicherheit folgt mittels

$$u_{\text{total}} = \sqrt{\sum_i (a_i \cdot u_i)^2}. \quad (4.2)$$

Unsicherheitsquellen für die Validierung der Reichweitevorschau mittels PGI

Als Grundvoraussetzung für die Validierung von Vorhersagemethoden der Protonenreichweite gilt, dass die Validierungsunsicherheit geringer ist als die Unsicherheiten in der Reichweitevorschau. Im Folgenden sollen daher die Unsicherheitsquellen der gewichteten mittleren Reichweiteabweichung, die als Maß für die absolute Genauigkeit der Reichweitevorschau dient, identifiziert, quantifiziert und mittels GUM zu einer Gesamtunsicherheit u_μ kombiniert werden:

(I) Positionierungspräzision des PGI-System im Isozentrum:

Eine Verschiebung der PGI-Kamera relativ zum sich im Isozentrum befindenden Patienten bewirkt einen Versatz aller gemessenen Reichweiteabweichungen für ein bestrahltes Feld. Die Positionierungspräzision wurde bereits im Zuge der Kalibrierung des PGI-Systems systematisch untersucht (vgl. Unterkapitel 3.3.2). Für die maßgebliche Unsicherheit in Strahlrichtung x ergab sich der Wert $\sigma_x = 0,51$ mm. Andere Unsicherheiten der Positionierung können als Beiträge höherer Ordnung vernachlässigt werden.

(II) pCT-zu-kCT Registrierung:

Die aufgenommenen kCT-Scans werden rigide mit dem pCT-Scan registriert, um die eigentliche Positionierung des Patienten von den medizinisch-technischen Radiologieassistenten (MTRAs) mit dem orthogonalen, planaren Röntgensystem nachzubilden. Die daraus

stammende Unsicherheit wurde gemeinsam mit erfahrenen MTRAs und Medizinphysikern abgeschätzt. Hierbei wurden auch die vom Autor durchgeführten Registrierungen exemplarisch durch MTRAs überprüft. Das Ergebnis dieser Abschätzung ist eine dreiecksverteilte WDF mit der Breite von $u_i = \pm 1$ mm. Daraus ergibt sich ein Unsicherheitsbeitrag laut GUM von $\sigma_{CT} = a_i \cdot u_i = \frac{u_i}{\sqrt{6}} = 0,41$ mm.

(III) Standardfehler der gewichteten mittleren Reichweiteabweichung:

Die statistische Streubreite der gemessenen Reichweiteabweichungen aufgrund von Poissonrauschen wird als Fehler des Mittelwerts (Standardfehler) in der Unsicherheitsabschätzung berücksichtigt. Der Standardfehler wurde mithilfe eines Bootstrapping-Verfahrens bestimmt, da die spotweisen Reichweiteabweichungen aufgrund der unterschiedlichen Spotgewichte keine gemeinsame statistische Streubreite besitzen und weiterhin aufgrund der Spotaggregation nicht unabhängig voneinander sind. Der Standardfehler ergibt sich als Streubreite der gewichteten mittleren Reichweiteabweichungen, die im Bootstrapping-Verfahren für insgesamt 10 000 zufällige Stichproben aus der Grundgesamtheit an Spots (Ziehen mit Zurücklegen) für jede Reichweitevorhersagemethode bestimmt wurde. Dabei enthielt jede Stichprobe genauso viele Spots wie die Grundgesamtheit.

(IV): Genauigkeit der Trägerwagenkalibrierung:

Aus Unterkapitel 3.3.1 geht hervor, dass die Lage des PGI-Systems zum Isozentrums in der Kalibrierung als globaler Korrekturterm für die Reichweiteunsicherheit berücksichtigt wird. Die hier ausgewertete Patientenkohorte wurde im ersten Kalibrierungszeitraum mit einer mittleren Abweichung zum Isozentrum von $-0,39$ mm aufgenommen (vgl. Tabelle 3.1). Die Ungenauigkeit dieser systematischen Korrektur ergibt sich als Standardfehler des Mittelwerts $\sigma_x / \sqrt{N} = 0,06$ mm mit $N = 70$ Einzelmessungen und $\sigma_x = 0,51$ mm (vgl. Unterkapitel 3.3.2).

(V): Genauigkeit der PGI-Strahlkalibrierung:

Als Genauigkeit der PGI-Strahlkalibrierung soll die Unsicherheit des globalen Versatzes zwischen gemessenen und simulierten PGI-Signalen bezeichnet werden, der für einen Zentral-spot mit 160 MeV im Zentrum des FOV bestimmt wurde (vgl. Abbildung 3.6 A). Die ermittelte Unsicherheit nach GUM beträgt $\sigma = 0,28$ mm.

(VI): Simulationsgenauigkeit:

Über die Genauigkeit der PGI-Strahlkalibrierung hinausgehend, wurden in Unterkapitel 3.5 einige weitere Aspekte aufgezeigt, welche die Simulationsgenauigkeit beeinflussen: das Detektoransprechvermögen, die Abhängigkeit von Protonenenergie und Bragg-Peak-Position der Spots im FOV sowie die Abhängigkeit von der Tumorentität und damit von der Gewebeszusammensetzung. Die Vielzahl der Faktoren, welche die Simulationsgenauigkeit beeinflussen, erschwert eine quantitative Korrektur der Daten. Daher wurden in diesem Teil der Arbeit, welcher sich mit der Bestimmung der absoluten Genauigkeit der Reichweitevorhersage

befasst, die daraus resultierenden Unsicherheiten mithilfe von Spotfiltern minimiert und heuristisch anhand der Patientenmessung abgeschätzt, was im Folgenden näher beschrieben wird. Maßgeblich diente hierzu die Änderung der Reichweiteabweichung im FOV des PGI-Systems. Auf Grundlage der Experimente im homogenen PMMA-Phantom wurde zunächst das effektive FOV der Auswertung auf den distalen Bereich [0; 5] cm auf der zentralen Strahlachse eingeschränkt, um den Einfluss der Protonenenergie zu minimieren (vgl. Unterkapitel 3.5.1). Weiterhin wurden PGI-Signale selektiert, die nicht mindestens einen 5 mm breiten Plateaubereich nach dem Signalabfall zeigten, wie bereits in einer Vorarbeit untersucht (Khamfongkhrua et al., 2020). Anhand des Verlaufs der gewichteten mittleren Reichweiteabweichung der verbleibenden Spots, wurde die Summe der weiteren Einflussfaktoren über die ausgewertete Patientenkohorte quantifiziert. Dazu wurden in fünf aufeinanderfolgenden Intervallen im verbleibenden FOV mit je 8 mm Breite die gewichteten mittleren Reichweiteabweichungen ermittelt, d. h. auf insg. 4 der 5 cm. Das Intervall [4; 5] cm wurde nicht berücksichtigt, da es lediglich 6% der Gesamtprotonenanzahl im FOV enthielt. Der Standardfehler des Mittelwerts der fünf Intervalle dient als heuristische Unsicherheitsabschätzung von Einflüssen der Simulationsgenauigkeit und ist von der jeweiligen Reichweitevorschagemethode abhängig.

Die Unsicherheitsbeiträge (I) bis (III) sind statistische Unsicherheiten, deren Einfluss auf die gewichtete mittlere Reichweiteabweichung durch weitere Messungen im Patienten verringert werden kann. Insbesondere kann der Einfluss von Unsicherheiten in der Positionierung des Patienten oder des PGI-Systems auf den Mittelwert durch die Anzahl von aufgenommenen Fraktionen N_f mit einem Faktor $1/\sqrt{N_f}$ verringert werden. Demgegenüber sind die Unsicherheitsbeiträge (IV) bis (VI) systematische Unsicherheiten, die sich aus den verwendeten Messmethoden und Messgeräten abschätzen lassen.

Anhand der quantifizierten Unsicherheitsquellen (Tabelle 4.1) wurde mittels GUM eine Gesamtunsicherheit für die Validierung der jeweiligen Vorhersagemethode bestimmt. Da die Unsicherheitsbeiträge III und VI unmittelbar aus der Verteilung der Reichweiteabweichungen berechnet werden, hängt auch die Gesamtunsicherheit von der verwendeten Reichweitevorschagemethode ab. Zusammengefasst ergab sich eine Gesamtunsicherheit u_μ von 1,1 mm (2σ) für DirectSPR, 0,9 mm (2σ) für die adaptierte HLUT und ebenso 0,9 mm (2σ) für die Standard-HLUT.

Im Einzelnen ergab sich für den Standardfehler des Mittelwerts mittels Bootstrapping-Verfahren (III) ein Wert von 0,04 mm (2σ) für das DirectSPR-Verfahren und je 0,03 mm (2σ) für die beiden HLUT-basierten Ansätze. Für die Simulationsgenauigkeit (VI), welche über Streuung der mittleren Reichweiteabweichung in Subintervallen des FOV bestimmt wurde, konnte ein Wert von 0,88 mm (2σ) für DirectSPR, 0,59 mm (2σ) für die adaptierte HLUT und 0,67 mm (2σ) für die Standard-HLUT bestimmt werden. Dieser Beitrag macht demnach die größte Unsicherheit der Validierung aus.

Tabelle 4.1: Zusammenfassung aller Unsicherheitsbeiträge der gewichteten mittleren Reichweiteabweichung μ zur Validierung verschiedener Reichweitevorhersagemethoden.

Unsicherheitsquelle	WDF	Unsicherheit der Validierung (2σ) / mm		
		DirectSPR	Adaptierte HLUT	Std.-HLUT
Stat. Unsicherheiten				
<i>Fraktionsweise:</i>				
I Positionierungspräzision ^a	Normal	1,02 / $\sqrt{N_f}$	1,02 / $\sqrt{N_f}$	1,02 / $\sqrt{N_f}$
II CT-Registrierung ^a	Dreieck	0,82 / $\sqrt{N_f}$	0,82 / $\sqrt{N_f}$	0,82 / $\sqrt{N_f}$
<i>Spotweise:</i>				
III Standardfehler Mittelwert	Normal	0,04	0,03	0,03
Syst. Unsicherheiten				
IV Trägerwagenkalibrierung	Normal	0,12	0,12	0,12
V PGI-Strahlkalibrierung	Normal	0,56	0,56	0,56
VI Simulationsgenauigkeit	Normal	0,88	0,59	0,67
Gesamtunsicherheit u_μ (2σ)		1,1	0,9	0,9

^a Anzahl der Fraktionen: $N_f = 30$

4.3 Ergebnisse der Validierung

Die spotweise ausgewerteten Fraktionen aller Patienten wurden nach Wichtung mit der Protonenanzahl zu einem Histogramm kombiniert (Abbildung 4.1), dessen gewichteter Mittelwert μ (Gleichung 4.1) ein Maß für die Genauigkeit der Reichweitevorhersage darstellt. Für die jeweils untersuchten Methoden ergab sich eine Vorhersagegenauigkeit mit Gesamtunsicherheit u_μ (2σ) von ($0,0 \pm 1,1$) mm für DirectSPR, ($0,7 \pm 0,9$) mm für die adaptierte HLUT und ($3,8 \pm 0,9$) mm für die Standard-HLUT. Bezogen auf die wasseräquivalente Protonenreichweite jedes Spots, ergab sich für die untersuchten Reichweitevorhersagemethoden im Mittel eine relative Abweichung und Gesamtunsicherheit (2σ) von ($0,0 \pm 0,5$)% für DirectSPR, ($0,3 \pm 0,4$)% für die adaptierte HLUT und ($1,8 \pm 0,4$)% für die Standard-HLUT. Im Mittel ist die wasseräquivalente Protonenreichweite aller einbezogenen Spots 21,8 cm gewesen.

Wie man anhand des Histogramms von HU-Werten der ausgewerteten Spots erkennen kann (Abbildung 4.2), wird in der hier durchgeführten Validierung der CT-basierten Reichweitevorhersage das ganze Spektrum von SPR-Werten im Pelvis-Bereich abgedeckt. Aus der Verteilungsbreite der Reichweiteabweichungen (Gaußsche Standardabweichung von 2,7 mm für DirectSPR, 2,4 mm für die adaptierte HLUT und 2,6 mm für die Standard-HLUT) kann keine weitere Schlussfolgerung über die Genauigkeit der Reichweitevorhersage getroffen werden, da diese vorrangig von der hohen statistischen Streubreite der spotweisen PGI-Messung abhängt. Insgesamt tragen 16 310 Spots mit ihrer Reichweiteabweichung zum gewichteten Histogramm bei, welche sich im Durchschnitt auf 544 (maximale Streubreite: 359–700) Spots pro Fraktion aufteilen. Die Variabilität der Anzahl an ausgewerteten Spots pro Fraktion

ergibt sich aus den unterschiedlich großen Zielvolumina der verschiedenen Patienten sowie der Einstellung des FOV, das mit dem derzeitigen PGI-System nur in 5 mm Schritten entlang der Strahlachse variiert werden kann.

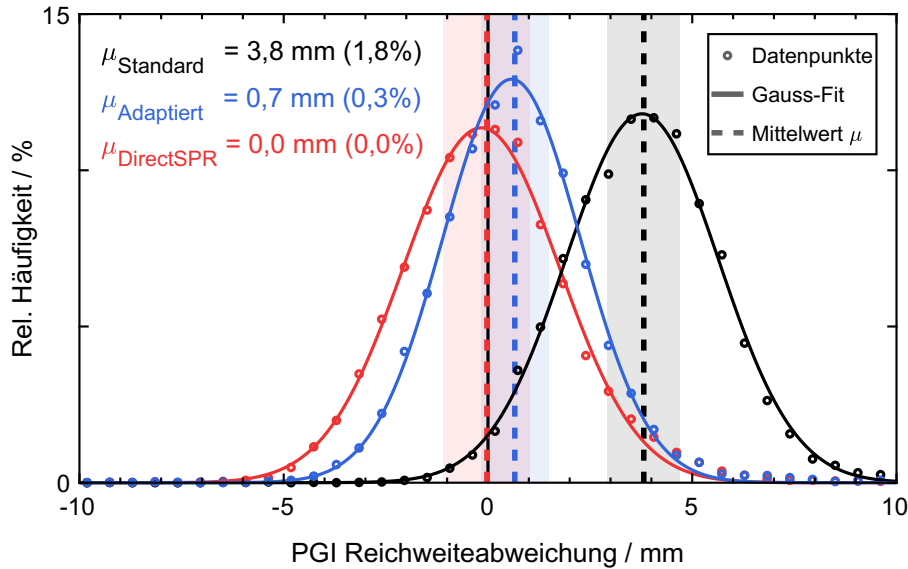


Abbildung 4.1: Gewichtetes Histogramm der mit PGI gemessenen spotweisen Reichweiteabweichungen im Bezug auf die drei untersuchten Reichweitevorhersagemethoden für alle Fraktionen und Patienten in dieser Auswertung. Die schattierten Bereiche entsprechen der Gesamtunsicherheit des Mittelwerts $\pm u_{\mu}(2\sigma)$ mit allen statistischen und systematischen Unsicherheitsbeiträgen. Die an die Datenpunkte angepasste Gauß-Funktion dient der Visualisierung und hebt hervor, dass die Daten annähernd normalverteilt sind.

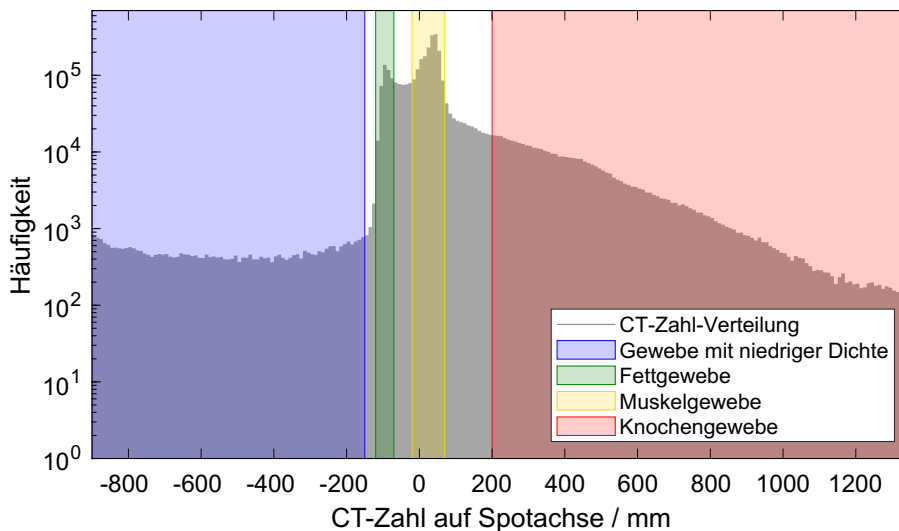


Abbildung 4.2: CT-Zahl Histogramm der Voxel auf den Strahlachsen aller für die Genauigkeit der Reichweitevorhersagemethoden ausgewerteten Spots. Die farbig schattierten Bereiche markieren Geweberegionen wie sie bei Wohlfahrt et al. (2020) definiert sind. Alle nicht hinterlegten CT-Zahl-Regionen stellen Bereiche mit gemischter Zusammensetzung dar (z. B. von Knochen- und Muskelgewebe).

4.4 Diskussion

Erstmals konnte die Vorhersage der Protonenreichweite in klinischen Prostata-Tumor-Bestrahlungen mit hoher Genauigkeit validiert werden. Die Grundlage dafür war die Anwendung eines hochgenauen Reichweiteverifikationssystems und die Verfügbarkeit von kCT-Scans als anatomische Grundwahrheit in Behandlungsposition.

Einordnung der Bedeutung der Validierung

Die hier durchgeführte Validierung bestätigt die Überlegenheit DECT-basierter Reichweitevorschauverfahren im Vergleich zum SECT-basierten Standard-HLUT-Verfahren. Die größte Genauigkeit wurde mit dem DirectSPR-Verfahren erreicht, das im Mittel keine Abweichungen der Protonenreichweite zwischen Vorhersage und PGI-Messung zeigte. Gleichwohl bedeutsam ist es, dass die ermittelten Gesamtunsicherheiten der Validierung durch das PGI-System mit ungefähr 1 mm wesentlich kleiner sind, als die in der Klinik gängigen Unsicherheiten der Reichweitevorschau (vgl. Unterkapitel 2.3.1). Dadurch wird eine aussagekräftige Validierung der Vorhersauverfahren überhaupt erst ermöglicht. Demzufolge stützt das Ergebnis dieser Validierung die Reduktion von Reichweiteunsicherheiten in der UPTD seit der routinemäßigen klinischen Anwendung des DirectSPR-Verfahrens ab 2019 ($1,7\% \pm 2$ mm im Kopf- und $2,0\% \pm 2$ mm im Pelvis-Bereich). Darüber hinaus könnte eine routinemäßige Anwendung von PGI dazu beitragen, die Reichweiteunsicherheiten potentiell noch weiter zu reduzieren.

Innerhalb der ermittelten Gesamtunsicherheit stimmt auch die mittlere Reichweiteabweichung beim Verfahren mit adaptierter HLUT – das im Kern auf DirectSPR basiert – mit einer perfekten Reichweitevorschau überein. Demgegenüber zeigte das Standard-HLUT-Verfahren eine relevante Abweichung der Protonenreichweite von $(1,8 \pm 0,4)\%$. Mit Blick auf eine hohe Variabilität bei der CT-basierten SPR- und Reichweitevorschau zwischen europäischen Protonentherapiezentren (Peters et al., 2021), unterstreicht dieses Ergebnis die Gültigkeit aktueller Reichweiteunsicherheiten beim Verfahren mit Standard-HLUT. Die Abweichung zum DirectSPR-Verfahren ist qualitativ und quantitativ übereinstimmend mit einer *in silico* Studie (Wohlfahrt et al., 2017b), welche die beiden Ansätze auf Dosisenebene, allerdings ohne Grundwahrheit, miteinander vergleicht.

Zusammengefasst komplementiert und bestätigt die dargestellte Untersuchung der Protonenreichweite im Patienten die umfassende Validierung von DECT-basierter Reichweitevorschau im anthropomorphen Kopfphantom (Wohlfahrt et al., 2018) und biologischem Gewebe (Möhler et al., 2018) sowie die umfangreiche Unsicherheitsabschätzung des DirectSPR-Verfahrens für die klinische Anwendung (Peters et al., 2022).

Rolle von PGI bei der Validierung

In dieser Studie konnte erstmals gezeigt werden, dass ein Reichweiteverifikationssystem erfolgreich zur Validierung von Methoden zur CT-basierten Reichweitevorhersage genutzt werden kann. Konzeptionell wird bei der Validierung die gewichtete mittlere Reichweiteabweichung der Spots in einem Behandlungsfeld betrachtet, da die statistische Unsicherheit der spotweisen PGI-Reichweiteabweichung potentielle Ungenauigkeiten der SPR-Vorhersage deutlich überlagert. Die Positionierung zwischen PGI-System und Patienten ist jedoch eine verbleibende statistische Unsicherheit der gewichteten mittleren Reichweiteabweichung eines einzelnen Feldes, welche eine Kombination aus der Positionierungspräzision des PGI-Systems (I) und der Präzision der CT-Registrierung (II) darstellt ($\sigma = \sqrt{\sigma_x^2 + \sigma_{CT}^2} = 0,65 \text{ mm}$). Um den Einfluss dieser Unsicherheit zu minimieren, wurden die Daten aus mehreren Feldern zusammengefasst, da die Unsicherheit des Mittelwerts über mehrere Felder N_f mit dem Faktor $1/\sqrt{N_f}$ gegenüber der Unsicherheit eines einzelnen Feldes reduziert wird. Durch die Zusammenfassung von insgesamt 30 Bestrahlungsfeldern konnte der Einfluss der Positionierung in dieser Validierung demnach um den Faktor 0,18 reduziert werden.

Mit der PGI-Methode kann das integrale SPR entlang der Strahlachse bis zur Eindringtiefe der Protonen für klinische Bestrahlungsfelder validiert werden. Ein Nachteil hieran ist, dass es prinzipiell zur Kompensation von Über- und Unterschätzungen des SPR im Strahlengang kommen kann. Als weitere Methoden zur Validierung gibt es z. B. die Protonenradiographie und das sogenannte *range probing* (vgl. Unterkapitel 2.3.2), bei welchen jedoch zusätzliche Dosis durch Protonen im Patienten eingetragen wird. Außerdem ermöglichen diese Methoden auch nur eine integrale Beurteilung der SPR-Vorhersage und sind noch stärker von Kompensationseffekten betroffen, da die Protonen den Patienten vollständig durchdringen müssen. Dafür können die Protonen direkt gemessen werden, wodurch statistische Unsicherheiten eine geringere Rolle spielen als bei PGI.

Limitationen

Die hier durchgeführte Validierung ist auf Prostata-Tumor-Behandlungen beschränkt. Unterschiede in der Zusammensetzung und Heterogenität des bestrahlten Gewebes anderer Tumorentitäten können theoretisch einen Einfluss auf die Übertragbarkeit der Ergebnisse haben. Für die Simulation der PG-Emission wurden die tabellierten Daten menschlichen Gewebes (Pelvis) von Woodard und White (1986) verwendet, welche immer noch die umfassendste und genaueste Zusammenstellung für diesen Zweck darstellen. In dieser Validierungsstudie wurden bewusst Prostata-Bestrahlungen gewählt, da für diese Bestrahlungen ein relativ homogenes Zielvolumen und eine für die Protonentherapie vergleichsweise hohe Eindringtiefe charakteristisch ist. Hierbei ergeben sich vergleichsweise große absolute Reichweiteunsicherheiten von etwa 7–10 mm (je nach Vorhersagemethode), wohingegen die Simulationsgenauigkeit von PGI im betrachteten FOV kaum von der Protonenenergie

abhängt (vgl. Unterkapitel 3.5.1). Dadurch ist die relative Aussagekraft der Validierung mittels PGI in Bezug auf die klinischen Reichweiteunsicherheiten für Felder mit einer größeren Eindringtiefe höher.

Als weitere Limitation der Studie sind potentielle intrafraktionelle Änderungen zwischen der Aufnahme des kCT-Scans und der eigentlichen Bestrahlung zu nennen. Solche Änderungen können sich beispielsweise auf die Füllung von Rektum, Darm und Blase beziehen, aber auch die teils ruckhafte Bewegung der Patientenliege kann Veränderungen der Patientenlagerung hervorrufen. Letzter Grund ist der Anlass für eine weitere Überprüfung der Patientenposition mit dem planaren Röntgensystem im Isozentrum direkt nach dem kCT-Scan. Damit können klinisch relevante rigide Abweichungen ausgeschlossen werden. Für nicht rigide Änderungen wird angenommen, dass diese zufällig und nicht systematisch auftreten, sodass deren Einfluss auf die gewichtete mittlere Reichweiteabweichung über mehrere Fraktionen hinweg vernachlässigbar ist.

Direkten Einfluss auf den Mittelwert haben systematische Effekte zwischen PGI-Simulation und -Messung. Im Unterkapitel 3.5 wurden als Ursachen für Abweichungen das Untergrundsignal (insbesondere durch Sekundärstrahlung, wie Neutronen), das Ansprechvermögen der Detektoren und die Gewebezusammensetzung (d. h. die Tumorentität) identifiziert. Der Einfluss des Untergrundes auf die PGI-Profile ist von der Position des Bragg-Peaks im FOV abhängig. In der Praxis kann der Untergrund in der Patientennmessung nicht bestimmt werden, sodass das effektive FOV in dieser Studie auf den distalen Teil beschränkt wurde, wo die Abhängigkeit von der Protonenenergie vernachlässigbar ist. Zudem beeinflusst das Ansprechvermögen der Detektoren die Reichweiteabweichungen teils deutlich, wie aus Abbildung 3.9 hervorgeht. Es zeigt sich, dass dieser Einfluss für PGI-Aufnahmen mit normaler Ausrichtung des Detektorarrays unter 90° Gantrywinkel deutlich stärker ausgeprägt ist, als für Aufnahmen unter 270° . Deshalb und weil das PGI-System nur unter 270° kalibriert werden kann, wurden für diese Studie nur die Felder unter 270° Gantrywinkel berücksichtigt. Zum Zeitpunkt der Durchführung und Auswertung dieser Untersuchung war noch nicht klar, dass die in Unterkapitel 3.5 vorgeschlagene FOV-Korrektur die Reichweiteabweichungen unter 270° und 90° zur besseren Übereinstimmung bringt und somit womöglich gerechtfertigt gewesen wäre. Andererseits kann nicht ausgeschlossen werden, dass der leicht veränderte Untergrund im Patienten eine andere Abhängigkeit bewirkt und eine Korrektur im Patienten somit die Daten verfälschen würde. Zudem müsste der Einfluss von der lateralen Spotposition auf die Reichweiteabweichungen für eine genaue Korrektur noch systematischer untersucht werden. Wendet man dennoch die in Unterkapitel 3.5 beschriebene FOV-Korrektur auf die hier gezeigten Daten an, ist die maximale Änderung der mittleren Reichweiteabweichung nur so groß, dass sie mit den heuristischen Unsicherheitsabschätzungen der Simulationsgenauigkeit (VI) in Einklang ist. Durch weitere Untersuchungen der PGI-Simulationsgenauigkeit z. B. im Pelvis-Phantom, könnte dieser Unsicherheitsbeitrag potentiell noch reduziert werden.

5 Detektionsfähigkeit anatomischer Veränderungen mittels PGI

Schwer vorhersehbare anatomische Veränderungen über den Verlauf der Behandlung stellen aktuell eine der größten Herausforderungen für die Genauigkeit in der Protonentherapie dar. Für die klinische Realisierung einer OAPT, in welcher die Bestrahlung optimalerweise während einer Fraktion oder sogar in Echtzeit (online) an die aktuelle Anatomie des Patienten anpasst wird, spielen 3D-Bildgebung des Patienten, schnelle Software zur Planadaption und die online Qualitätssicherung der Bestrahlung eine zentrale Rolle. Als Teil eines solchen online-adaptiven Arbeitsablaufs sind daher Systeme zur Behandlungsverifikation gefragt, die Abweichungen in der Dosisapplikation des (potentiell geänderten) Bestrahlungsplans während der Bestrahlung unabhängig im Patienten überprüfen können und damit zur Qualitätssicherung beitragen (Albertini et al., 2020; Paganetti et al., 2021a, 2021b). In diesem Kapitel soll das in Kapitel 3 eingeführte PGI-System im Hinblick auf sein Potential zur Erkennung von Behandlungsabweichungen bei Prostata- und Kopf-Hals-Tumor-Patienten untersucht werden. Es wird Bezug genommen auf in Konferenzen vorgestellte Ergebnisse (Berthold et al., 2022a, 2023), sowie während dieser Arbeit betreute studentische Abschlussarbeiten (Hübinger, 2021; Piplack, 2021).

5.1 Prinzipieller Aufbau der Studie

Das Potential von PGI zur Erkennung von Behandlungsabweichungen soll anhand der Detektierbarkeit von anatomischen Veränderungen retrospektiv untersucht werden. Durch die Verfügbarkeit von kCT-Scans im Behandlungsraum konnten sowohl anatomische als auch dosimetrische Abweichungen der Behandlung gegenüber der geplanten Bestrahlung genau bestimmt und somit eine Grundwahrheit für die Behandlungsverifikation abgeleitet werden. Zur Bestimmung von dosimetrischen Abweichungen der Behandlung wurden spotweise integrierte Tiefendosisprofile (ITD-Profile) für die Dosisverteilungen auf den pCT- und kCT-Scans extrahiert, indem die berechnete 3D-Dosisverteilung jedes einzelnen Spots senkrecht zur Strahlachse aufsummiert wurde. Zwischen den ITD-Profilen eines Spots auf dem pCT- und kCT-Scan konnte so eine dosisbasierte ITD-Reichweiteabweichung ermittelt werden. Auf ähnliche Weise wurden die PGI-Reichweiteabweichungen für jeden Spot zwischen der

entsprechenden PGI-Messung und der auf dem pCT-Scan basierenden PGI-Simulation ermittelt. Darüber hinaus wurde ein sogenanntes *Clustering* zur Bildung von *Clustern* auf die spotweisen PGI-Reichweiteabweichungen angewandt, um den Einfluss statistischer Messunsicherheiten zu reduzieren. Im Rahmen einer Teiluntersuchung wurde dann zunächst die Pearson-Korrelation zwischen ITD- und PGI-Reichweiteabweichungen für Spots und Cluster quantifiziert. Der primäre Endpunkt der Untersuchung ist jedoch eine binäre Klassifikation für jedes applizierte Bestrahlungsfeld, ob relevante oder nicht relevante anatomische Veränderungen während der Bestrahlung vorgelegen haben (*relevante* oder *nicht relevante* Felder). Dafür wurde eine CT-basierte feldweise Grundwahrheit mithilfe von ITD-Profilen definiert, gegenüber der das Klassifikationspotential eines unabhängigen PGI-basierten Klassifikationsmodells getestet wurde (Abbildung 5.1).

Die Definition einer relevanten anatomischen Veränderung im Bestrahlungsfeld sollte sich idealerweise an den klinischen Anforderungen orientieren. Wegen der derzeit verwendeten klinischen Sicherheitssäume bei der Bestrahlungsplanoptimierung tritt die Notwendigkeit einer Planadaption, die sich nach der Einhaltung der klinischen Zielvorgaben für DVH-Parameter richtet, aber nur selten auf. Lediglich bei zwei der insgesamt fünfzehn Prostata-Tumor-Patienten und bei keinem der vier Kopf-Hals-Tumor-Patienten wurde während des Monitorings mit PGI eine Adaption durchgeführt. Jedoch wird man im Rahmen einer OAPT die Sicherheitssäume weiter reduzieren wollen, um noch mehr Normalgewebe zu schonen. Daher wurde in dieser Arbeit bewusst ein neuer, auch für kleine Dosisänderungen sensitiver Standard für die Notwendigkeit einer Planadaption definiert, für den zunächst kein Pendant in der Literatur zu finden ist. Die hier verwendete feldweise Grundwahrheit basiert im Kern auf einer manuellen Klassifikation der Dosisverteilungen auf dem pCT- und kCT-Scan, anhand welcher dann eine automatische, auf den ITD-Reichweiteabweichungen basierende Klassifikation etabliert wurde.

Für die Klassifikation der PGI-Daten wurde ein feldweises Klassifikationsmodell entwickelt, welches anhand der Grundwahrheit optimiert wurde. Diese Optimierung erfolgte getrennt für Prostata- und Kopf-Hals-Tumor-Patienten und wurde zunächst mit einer Trainingskohorte von 131 Feldern (10 Prostata-Tumor-Patienten) bzw. 11 Feldern (3 Kopf-Hals-Tumor-Patienten) durchgeführt. Abschließend wurde das entwickelte PGI-Modell an einer Testkohorte mit 70 Feldern (5 unabhängige Prostata-Tumor-Patienten) und 11 Feldern (3 Kopf-Hals-Tumor-Patienten, davon einer unabhängig) getestet, wobei die Sensitivität und Spezifität zur Erkennung von interfraktionellen Veränderungen quantifiziert wurde.

5.2 Grundwahrheit auf Basis von CT- und Dosisinformationen

Zur Bestimmung der Grundwahrheit anatomischer Abweichungen im Patienten gegenüber der Bestrahlungsplanung dienen die mit einem in-room CT-Scanner aufgenommenen

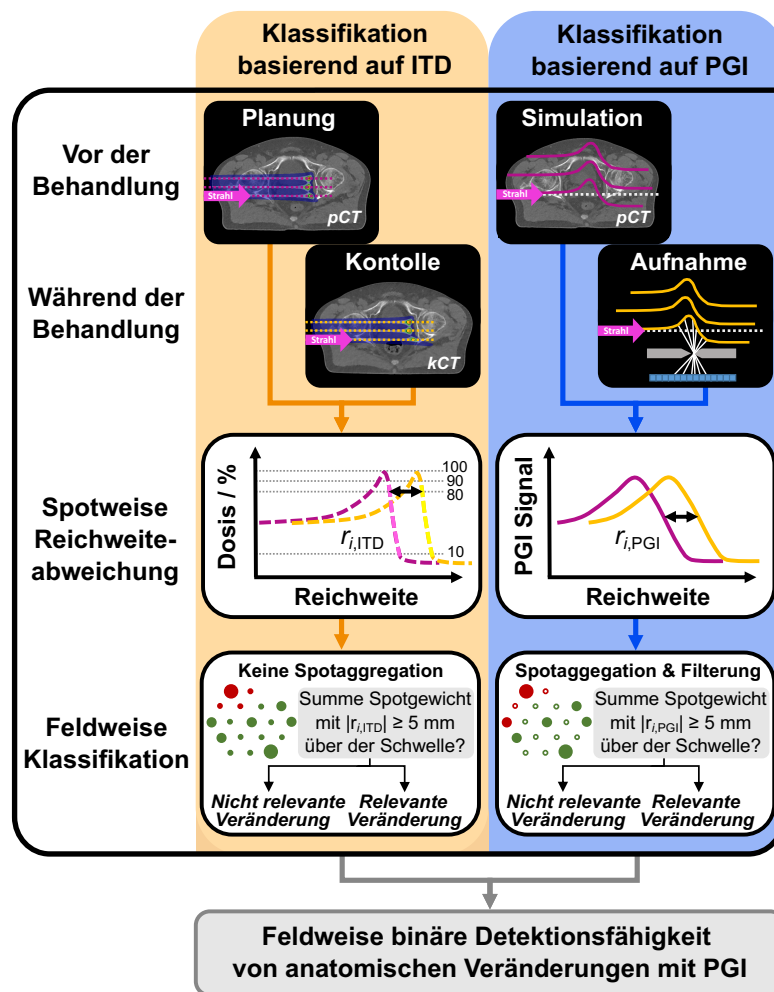


Abbildung 5.1: Studiendesign zur Bestimmung der Detektierbarkeit feldweiser anatomischer Veränderungen mit einem PGI-System. Dabei wird eine auf ITD-Profilen basierende feldweise Grundwahrheit mit einem feldweisen PGI-Klassifikationsmodell verglichen.

kCT-Scans. Wie auch für die Untersuchung im vorangegangenen Kapitel müssen die aufgenommenen kCT-Scans mit dem pCT-Scan rigide registriert werden, um die Positionierung des Patienten durch die MTRAs während der Behandlung nachzubilden. Alle Prostata-Tumor-Patienten wurden vom Autor dieser Arbeit registriert; die Kopf-Hals-Tumor-Patienten wurden im Rahmen der betreuten Physik-Masterarbeit von Lisa Hübinger registriert. In beiden Fällen wurde Rücksprache mit MTRAs genommen, welche die Registrierungen exemplarisch und für besonders komplizierte Fälle geprüft haben. Die Registrierung wurde in der RayStation mit dem Fokus auf der Überlagerung der starren Knochenstruktur im Patienten durchgeführt. Aufgrund von nicht-rigiden Änderungen im Bereich des Unterkiefers und der Wirbelsäule musste in manchen Fällen die Registrierung im Kopf-Hals-Bereich unter Berücksichtigung des Bestrahlungsvolumens abgewogen werden. Außerdem wurden Artefaktregionen von z. B. Metallimplantaten in einigen kCT-Scans von Kopf-Hals-Tumor-Patienten durch erwartetes Gewebe überschrieben. Dies erfolgte nach Rücksprache mit klinischen Medizinphysikern, die ähnliche Überschreibungen in den pCT-Scans durchgeführt hatten.

Zur Dosisberechnung wurde die DirectSPR-Methode verwendet, die nachweislich am genauesten in Prostata-Tumor-Patienten ist (vgl. Kapitel 4). Für Kopf-Hals-Tumor-Patienten war dies nicht immer möglich, da zur Reduktion von Bewegungsartefakten die pCT-Scans prinzipiell nur als SECT durchgeführt wurden. Um dennoch von der verbesserten Reichweitenvorhersage durch DirectSPR zu profitieren, wurde eine adaptierte HLUT für die SECT-Scans verwendet (vgl. Wohlfahrt et al., 2020). Da die kCT-Scans hingegen als DECT aufgenommen wurden, konnte hier je nach Ausprägung von Bewegungsartefakten die DirectSPR-Methode (fehleranfälliger bei Bewegung) oder die adaptierte HLUT auf den monoCT--Datensätzen (letztere bei zwei Feldern) zum Einsatz kommen. Darauf basierend wurde abschließend für alle pCT- und registrierten kCT-Scans die Dosisverteilung der jeweiligen Bestrahlungsfelder einheitlich mit der RayStation 9B berechnet (vgl. Unterkapitel 2.2.2).

Mithilfe von PGI können prinzipiell solche anatomischen Änderungen detektiert werden, die eine Änderung der Dosisverteilung und somit der Protonenreichweite im Patienten bewirken. Die allein auf CT-Scans basierte Definition einer feldweisen Grundwahrheit für das Auftreten relevanter anatomischer Veränderungen, die auch tatsächlich mit PGI erkannt werden können, erfolgte schrittweise: (1) Zunächst wurde eine manuelle Auswertung und Klassifizierung der Dosisverteilungen auf pCT- und kCT-Scan mit zusätzlicher Identifizierung der anatomischen Ursachen für die Dosisänderungen in den beiden CT-Scans durchgeführt. (2) Die manuelle Klassifizierung dient als Referenz für die Etablierung eines automatischen Klassifizierungsansatzes der Dosisverteilungen anhand der spotweisen ITD-Reichweiteabweichungen. Zusätzlich getestet wurde im Rahmen dieser Arbeit noch ein weiterer automatischer Klassifizierungsansatz auf Grundlage einer Fehlerrate, die durch eine Gamma-Analyse zwischen den Dosisverteilungen auf pCT- und kCT-Scan bestimmt wurde (Low et al., 1998). Dieser volumenbasierte Ansatz korrelierte insbesondere bei relevanten lokalen Änderungen schlecht mit der manuellen Klassifikation und wurde daher nicht weiter verfolgt. Die beiden anderen Klassifizierungsansätze werden im Folgenden noch detaillierter beschrieben.

5.2.1 Manuelle Klassifizierung

Die manuelle Klassifizierung basiert auf der Beurteilung der Dosisdifferenz zwischen den Dosisverteilungen auf pCT- und kCT-Scan innerhalb der RayStation-Software, die es auch ermöglicht, manuelle Liniendosisprofile auszuwerten. Der Arbeitsablauf für jedes Bestrahlungsfeld kann wie folgt zusammengefasst werden:

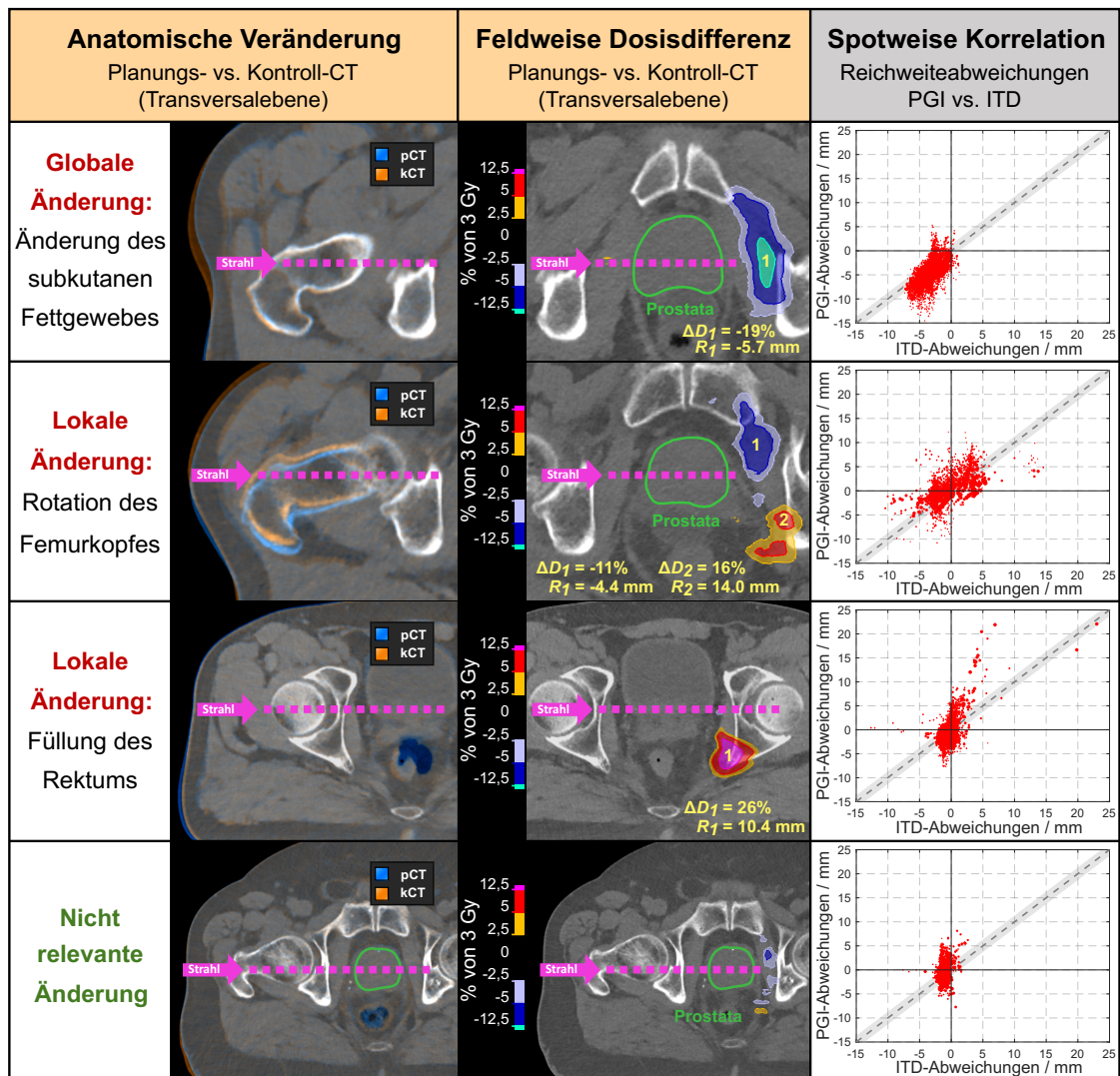
- (1) In der Dosisdifferenzverteilung werden für alle Regionen n (typischerweise 0–6 pro Feld) mit erkennbaren *lokalen* (Teile des Feldes sind betroffen) oder *globalen* (das ganze Feld ist betroffen) Dosisdifferenzen ΔD_n relativ zur Fraktionsdosis bestimmt (Prostatatumore: 3 Gy, Kopf-Hals-Tumore: 2 Gy).

- (2) Durch die maximale Dosisdifferenz in jeder dieser Regionen wird manuell ein Liniendosisprofil durch die Gesamtdosisverteilung in Strahlrichtung gelegt (bei gleichzeitiger Abweichung des Gantrywinkels von der Horizontalen und einer Rotation der Patientenliege nicht immer exakt möglich). Anhand dessen wird zwischen pCT- und kCT-Scan eine Reichweiteabweichung R_n bei 80% des distalen Dosisabfalls auf dem pCT-Scan ermittelt. Um nicht in den Bereich der lateralen Penumbra für das Bestrahlungsfeld zu kommen, wo die Reichweiteabweichungen in Strahlrichtung nicht wohldefiniert sind, werden Liniendosisprofile nur ausgewertet, wenn die Dosis auf dem pCT 50% der maximalen Dosis im Bestrahlungsfeld übersteigt. Falls keine der Liniendosisprofile in einer Region diese Bedingung erfüllt, wird die ganze Region von der Betrachtung ausgeschlossen.
- (3) Anatomische Veränderungen in Strahlrichtung werden in der Fusionsansicht von pCT- und kCT-Scan identifiziert und den detektierten Reichweiteabweichungen zugeordnet.
- (4) Behandlungsfelder werden auf Grundlage der Ausprägung lokaler oder globaler Reichweiteabweichungen $|R_n|$ mit den Dosisdifferenzen $|\Delta D_n|$ als relevante bzw. nicht relevante Felder klassifiziert (Abbildung 5.2 und 5.3). Hierbei ist ein Feld bereits dann relevant, wenn eine Region n als relevant eingeordnet wird.

Die Klassifizierungsmatrix ergab sich aus einem iterativen Prozess während der Auswertung der Patientendaten und durch die Diskussion mit den klinischen Medizinphysikern. Dabei sollten vor allem Reichweiteabweichungen $|R_n| \geq 5$ mm in Kombination mit großen relativen Dosisänderungen $|\Delta D_n| \geq 12,5\%$ der Fraktionsdosis als relevante Änderungen deklariert werden, um potentiell OARs in Strahlrichtung (nicht der Fall bei Prostata-Tumor-Bestrahlungen) zu schonen. Außerdem wurden globale Reichweiteänderungen, welche nur in der Prostata auftraten, bereits ab $|R_n| \geq 3$ mm als relevant deklariert, da dieser Wert den angenommenen Positionierungsunsicherheiten bei der Patientenbestrahlung entspricht. Kleine Dosisabweichungen $|\Delta D_n| < 12,5\%$ der Fraktionsdosis werden prinzipiell als nicht relevant erachtet, selbst bei großen Reichweiteabweichungen z. B. an flachen distalen Dosisgradienten. Die hier dargestellte Definition von Feldern mit relevanten und nicht relevanten anatomischen Veränderungen liefert eine manuelle Klassifikation, die mehr als die Hälfte der monitorierten Felder als relevant klassifiziert und damit voraussichtlich sensitiv genug für eine potentielle Reduktion von Sicherheitssäumen im Rahmen der OAPT ist.

5.2.2 Klassifizierung auf Grundlage von integrierten Tiefendosisprofilen

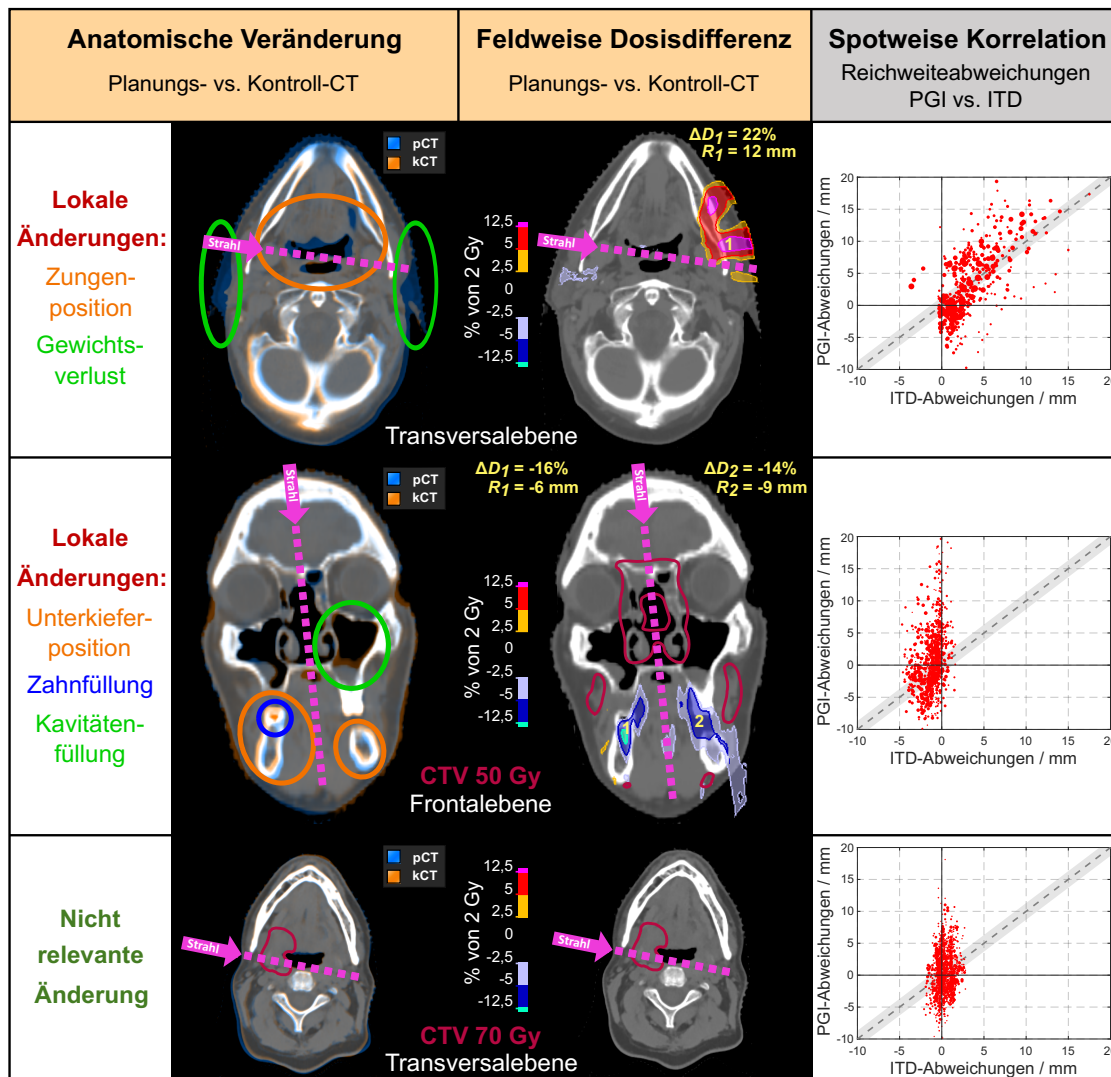
Mithilfe von integrierten Tiefendosisprofilen konnte zunächst die dosisbasierte Reichweiteabweichung auf Spotebene bestimmt werden. Hierbei wurde RayStation-intern durch die radiale Integration der 3D-Dosisverteilung eines Spots i auf seine jeweilige Strahlachse ein 1D-Tiefendosisprofil generiert (Auswertungstool wurde bereitgestellt von Erik Traneus,



Manuelle Klassifikationsmatrix

$ \Delta D_n \backslash R_n $	< 3 mm	3–5 mm	≥ 5 mm
< 12,5%	nicht relevant		
$\geq 12,5\%$	nicht relevant	Lokal: nicht relevant Global: relevant	relevant

Abbildung 5.2: Repräsentative Beispiele anatomischer Veränderungen für Prostata-Tumor-Patienten (linke Spalte) und deren Einfluss auf die Dosisdifferenzverteilung (mittlere Spalte), in welcher die Dosisdifferenzen ΔD_n und Reichweiteabweichungen R_n für entsprechende Regionen n exemplarisch dargestellt sind. Die manuelle Klassifikation der Bestrahlungsfelder basiert auf der unten dargestellten Klassifikationsmatrix und entspricht der ungünstigsten Klassifikation einer Region im Feld. Weiterhin ist die spotweise Korrelation zwischen PGI- und ITD-Reichweiteabweichungen dargestellt (rechte Spalte). Das grau schattierte Konfidenzband um die Eins-zu-eins-Korrelation (gestrichelte Linie) markiert den 2σ -Bereich der erwarteten Kamerapositionierungs- und Registrierungsunsicherheit ($2\sigma = 1,3$ mm, vgl. Unterkapitel 4.2).



Manuelle Klassifikationsmatrix

$ \Delta D_n \backslash R_n $	< 3 mm	3–5 mm	$\geq 5 \text{ mm}$
< 12,5%	nicht relevant		
$\geq 12,5\%$	nicht relevant	Lokal: nicht relevant Global: relevant	relevant

Abbildung 5.3: Repräsentative Beispiele anatomischer Veränderungen für Kopf-Hals-Tumor-Patienten (linke Spalte) und deren Einfluss auf die Dosisdifferenzverteilung (mittlere Spalte), in welcher die Dosisdifferenzen ΔD_n und Reichweiteabweichungen R_n für entsprechende Regionen n exemplarisch dargestellt sind. Man beachte, dass durch die potentielle Rotation der Patientenliege bei gleichzeitiger Abweichung von horizontalen Gantrywinkeln die angezeigte Strahlrichtung nicht in der dargestellten Ebene liegen muss. Die manuelle Klassifikation der Bestrahlungsfelder basiert auf der unten dargestellten Klassifikationsmatrix und entspricht der ungünstigsten Klassifikation einer Region im Feld. Weiterhin ist die spotweise Korrelation zwischen PGI- und ITD-Reichweiteabweichungen dargestellt (rechte Spalte) mit Eins-zu-eins-Korrelation und 2σ -Konfidenzband (vgl. Abbildung 5.2).

RaySearch Laboratories). Anhand der ITD-Profile auf pCT- und kCT-Scan konnten spotweise ITD-Reichweiteabweichungen $r_{i,ITD}$ ermittelt werden. Hierfür wurden zwei verschiedenen Ansätze verwendet: (1) über die Differenz des $R80$, d. h. die Reichweite der Protonen bei 80% des Abfalls der ITD-Profile und (2) über die Bestimmung der horizontalen Verschiebung zwischen den ITD-Profilen, welche die quadratische Summe der Profildifferenzen im Bereich von 90–10% des distalen Profileabfalls auf dem pCT-Scan minimiert (Abbildung 5.1). Der letztere Ansatz ist der robustere bei der Analyse von ITD-Profile mit deutlichen Inhomogenitäten lateral zur Strahlrichtung und wurde insbesondere für die Analyse der Kopf-Hals-Tumor-Patienten etabliert. Bei den Prostata-Tumor-Patienten wurde historisch gewachsen der erste Ansatz verwendet, welcher aufgrund weniger oft vorkommender starker Dichtegradienten lateral zur Strahlrichtung (bspw. durch luftgefüllte Hohlräume oder Kavitäten) nur eine geringe Einschränkung darstellt.

Die ITD-Reichweiteabweichungen dienen als Grundlage für die Klassifikation einer feldweisen Grundwahrheit. Im Gegensatz zur PGI-basierten Auswertung, wurde hierbei keine Spotaggregation (vgl. Unterkapitel 3.2.2) angewandt, damit die feldweise Grundwahrheit – der primäre Endpunkt dieser Untersuchung – frei von Limitationen der PGI-Methode ist. Demgegenüber wurde die Spotaggregation in einer Teiluntersuchung zur Quantifizierung der Pearson-Korrelation und der Differenzverteilung zwischen ITD- und PGI-Reichweiteabweichungen auch auf die ITD-Profile angewandt, um einen direkten Vergleich auf Spot- und Clusterebene zu ermöglichen.

Für den primären Endpunkt dieser Untersuchung wurde eine ITD-basierte Klassifikation (kurz: ITD-Klassifikation) entwickelt, welche auf zwei Parametern beruht. Anhand des ersten Parameters werden zunächst Spots mit relevanten Abweichungen (*relevante Spots*) definiert, wofür analog zur Gesamtdosisverteilung bei der manuellen Klassifikation ein konstanter Schwellwert mit $|r_{i,ITD}| \geq 5 \text{ mm}$ für die ITD-Reichweiteabweichungen ohne Spotaggregation festgelegt wurde. Im Gegensatz zur manuellen Klassifikation erfolgt hier keine weitere Unterscheidung zwischen lokalen und globalen Änderungen. Der zweite Parameter dient zur Abgrenzung zwischen relevanten und nicht relevanten Feldern, indem die summierten Spotgewichte w_i aller relevanten Spots ins Verhältnis zum Gesamtgewicht des Bestrahlungsfeldes gesetzt werden, sodass sich der Parameter

$$W_{\text{rel,ITD}} = \frac{\sum_{i(|r_{i,ITD}| \geq 5 \text{ mm})} w_i}{\sum_i w_i} \begin{cases} < T_{\text{rel,ITD}}: \text{ nicht relevant} \\ \geq T_{\text{rel,ITD}}: \text{ relevant} \end{cases} \quad (5.1)$$

mit $i \in \{1, \dots, N\}$ ergibt, wobei N die Anzahl der Spots im Bestrahlungsfeld ist. Dem Ansatz entsprechend werden Felder als relevant oder nicht relevant klassifiziert, wenn der Parameter $W_{\text{rel,ITD}}$ über oder unter einem gewissen Schwellwert $T_{\text{rel,ITD}}$ liegt. Dieser Schwellwert wurde unter Berücksichtigung der Korrelation mit der manuellen Klassifikation bestimmt, da (1) hier erstmals ein ITD-basierter Klassifizierungsansatz Anwendung findet und

(2) der manuellen Klassifikation bereits viele Detailkenntnisse über die Behandlungsabweichungen und deren Ursachen zu Grunde liegen. Die Anpassung des Schwellwertes $T_{rel,ITD} \in [0; 100]$ in 0,1% Schritten wurde getrennt für Prostata-Tumor-Patienten und Kopf-Hals-Tumor-Patienten durchgeführt und hatte dabei das Ziel, den Matthews Korrelationskoeffizienten (MKK) zwischen manueller und ITD-Klassifikation zu maximieren (Chicco & Jurman, 2020; Matthews, 1975; Yule, 1912). Im binären Fall ist der MKK definiert als

$$MKK = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{(n_{11} + n_{01})(n_{10} + n_{00})(n_{11} + n_{10})(n_{01} + n_{00})}}, \quad (5.2)$$

wobei n_{xy} für die Anzahl der Elemente (hier: Felder) mit manueller Klassifikation $x \in \{0, 1\}$ und ITD-Klassifikation $y \in \{0, 1\}$ steht (hier: 0 - nicht relevant, 1 - relevant). Zur besseren Einordnung des MKK wurde weiterhin die Korrekturklassifikationsrate, engl. *accuracy* (ACC), bestimmt als

$$ACC = \frac{n_{11} + n_{00}}{n_{11} + n_{00} + n_{10} + n_{01}}. \quad (5.3)$$

5.2.3 Ergebnis der Etablierung einer CT-basierten Grundwahrheit

Auswertung für Patienten mit Prostata-Tumor

Bei der manuellen Auswertung der CT-Scans konnten drei Hauptursachen anatomischer Veränderungen in Prostata-Tumor-Patienten identifiziert werden, die zu relevanten Reichweiteabweichungen der Protonen führen: (a) (gasförmige) Füllung des Rektums, (b) Rotation des Oberschenkels und (c) Variationen des lateralen Weichgewebes (z. B. subkutanes Fettgewebe) aufgrund der Patientenpositionierung oder Veränderungen des Patientengewichts (Abbildung 5.2). Die Rotation des Oberschenkels trat hierbei am seltensten auf, bewirkte aber fast immer auch eine Verschiebung des lateralen Weichgewebes, welche die häufigste Ursache für anatomische Veränderungen darstellte. In anderen Studien zu Prostata-Tumor-Behandlungen wurden vergleichbare Ursachen anatomischer Veränderungen genannt (Moatebbed et al., 2016; Trofimov et al., 2011; Wang et al., 2011). Jedem als relevant klassifizierten Feld konnte mindestens eine dieser anatomischen Veränderungen zugeordnet werden. Häufig kam es auch zur Überlagerung mehrerer Ursachen.

Die Anpassung zwischen der feldweisen manuellen und ITD-basierten Klassifikation erreicht eine maximale Korrelation (MKK = 0.64) für einen Schwellwert von $T_{rel,ITD} = 1,5\%$ (Abbildung 5.4), d. h. ein Feld wird als relevant klassifiziert, wenn die Summe der Gewichte aller Spots mit $|r_{i,ITD}| \geq 5$ mm mindestens 1,5% des kumulativen Spotgewichts aller Spots im Feld beträgt. Damit wird zwischen manueller und ITD-Klassifikation eine gute Übereinstimmung von ACC = 81% erzielt. Die etablierte ITD-Klassifikation wird daher als finale Grundwahrheit für den Vergleich mit den PGI-Klassifikationsmodellen verwendet.

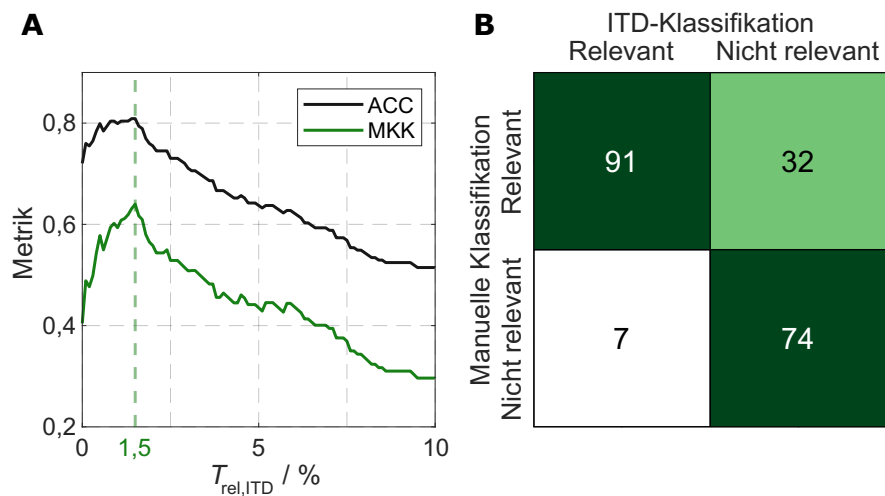


Abbildung 5.4: Vergleich zweier Klassifizierungsansätze zur Festlegung einer Grundwahrheit für anatomische Veränderungen bei den Prostata-Tumor-Patienten. (A) Optimierung der ITD-Klassifikation gegenüber der manuellen Klassifikation anhand der Maximierung der gewählten Metrik (Matthews Korrelationskoeffizient (MKK) bzw. Korrektklassifikationsrate ACC) bei Variation des Schwellwerts ($T_{rel,ITD}$). (B) Resultierende Wahrheitsmatrix bei optimalem Schwellwert $T_{rel,ITD} = 1,5\%$.

Auswertung für Patienten mit Tumoren im Kopf-Hals-Bereich

Bei der manuellen Auswertung der CT-Scans konnten drei Hauptursachen anatomischer Veränderungen in Kopf-Hals-Tumor-Patienten identifiziert werden, die zu relevanten Reichweiteabweichungen der Protonen führen: (a) Gewebeschrumpfung, (b) Verteilung von Luft in Mund- und Nasenhöhle (z. B. die Zungen- oder Gaumenposition) und (c) Rotation und Verkipfung von Wirbelsäule und Unterkiefer (Abbildung 5.3). Jedem als relevant klassifizierten Feld konnte mindestens eine dieser anatomische Veränderungen zugeordnet werden. Häufig kam es allerdings auch zur Überlagerung mehrerer Ursachen.

Die Anpassung zwischen der feldweisen manuellen und ITD-basierten Klassifikation erreicht eine maximale Korrelation (MKK = 0.63) für einen Schwellwert von $T_{rel,ITD} = 4,5\%$ (Abbildung 5.5), d. h. ein Feld wird als relevant klassifiziert, wenn die Summe der Gewichte aller Spots mit $|r_{i,ITD}| \geq 5$ mm mindestens 4,5% des kumulativen Spotgewichts aller Spots im Feld beträgt. Damit wird zwischen manueller und ITD-Klassifikation eine gute Übereinstimmung von ACC = 77% erzielt. Die etablierte ITD-Klassifikation wird daher als finale Grundwahrheit für den Vergleich mit den PGI-Klassifikationsmodellen verwendet.

5.3 Etablierung einer Klassifikation auf Basis von PGI-Daten

Die Etablierung eines PGI-basierten feldweisen Klassifikationsmodells unabhängig von der dosisbasierten Grundwahrheit ist eine essentielle Voraussetzung dieser Untersuchung im Hinblick auf die klinische Anwendung des Modells, bei der in-room CT-Scans nicht zur Verfügung stehen. Neben einem spotbasierten Ansatz wurde für die Klassifizierung ein

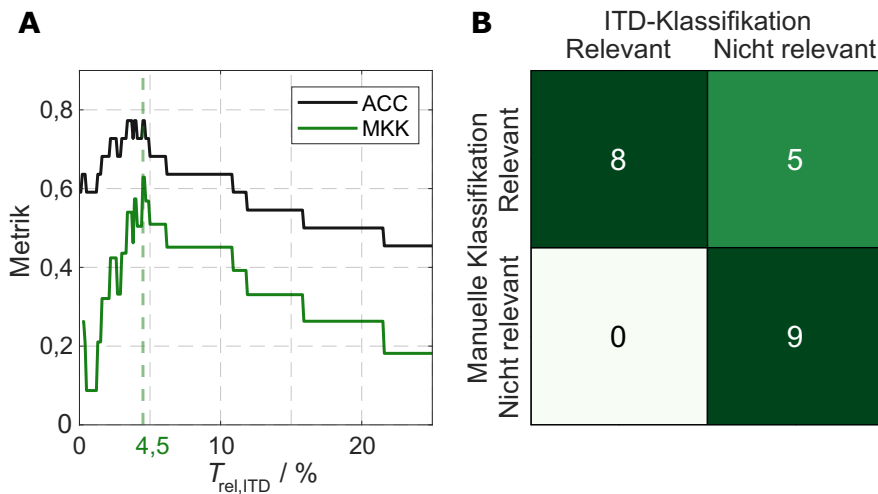


Abbildung 5.5: Vergleich zweier Klassifizierungsansätze zur Festlegung einer Grundwahrheit für anatomische Veränderungen bei den Kopf-Hals-Tumor-Patienten. (A) Optimierung der ITD-Klassifikation gegenüber der manuellen Klassifikation anhand der Maximierung der gewählten Metrik (Matthews Korrelationskoeffizient (MKK) bzw. Korrektklassifikationsrate ACC) bei Variation des Schwellwerts ($T_{rel,ITD}$). (B) Resultierende Wahrheitsmatrix bei optimalem Schwellwert $T_{rel,ITD} = 4,5\%$.

Cluster-Algorithmus entwickelt, welcher PGI-Reichweiteabweichungen benachbarter Spots kombiniert, um den Einfluss von Signalrauschen weiter abzuschwächen und damit die Verlässlichkeit der PGI-Reichweiteabweichungen für die Klassifikation zu erhöhen. Im Folgenden sollen zunächst der Cluster-Algorithmus und danach die auf Spots und Clustern basierenden Klassifikationsmodelle beschrieben werden.

5.3.1 Verarbeitung der PGI-Daten mittels Cluster-Algorithmus

Die Spotgewichtung, d. h. die applizierte Anzahl an Protonen pro Spot, wird bei der PBS-Bestrahlungsplanung variiert, um eine möglichst homogene Dosisverteilung im Zielvolumen zu erreichen. Für die auf PG basierende Reichweite- bzw. Bestrahlungsverifikation bedeutet dies gleichzeitig eine Herausforderung, da die Präzision solcher Methoden mit der Anzahl der applizierten Protonen skaliert (vgl. Perali et al., 2014). Zusätzlich zu der in dieser Arbeit verwendeten Spotaggregation innerhalb eines Layers, bei welcher die spotweisen PGI-Signale noch vor der Bestimmung einer Reichweiteabweichung summiert werden (vgl. Unterkapitel 3.2.2), sollen durch einen 3D-Cluster-Algorithmus Reichweiteabweichungen über benachbarte Spots gemittelt werden. Der Cluster-Algorithmus ist so konzipiert, dass die entstehenden Cluster eine möglichst gleiche Präzision besitzen, d. h. eine gleiche Anzahl an applizierten Protonen. Hierzu werden zunächst die Spots anhand ihrer Bragg-Peak-Position im Patienten in ein uniformes Voxelraster mit Basislänge 1 cm (Prostata) bzw. 1,6 cm (Kopf-Hals) zusammengeführt. Wie in Abbildung 5.6 dargestellt werden daraufhin zusammenhängende Voxel kombiniert, die eine definierte kumulative Protonenanzahl erreichen (Prostata: $\geq 3 \cdot 10^9$, Kopf-Hals: $\geq 0,75 \cdot 10^9$), um dann ein Cluster bestehend aus maximal 8 Voxeln

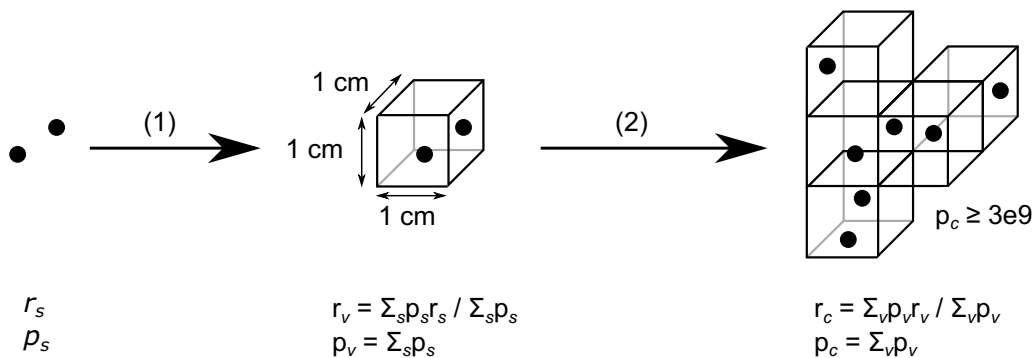


Abbildung 5.6: Visualisierung des Cluster-Algorithmus mit den Parametern für Prostata-Tumor-Patienten. (1) Spots (Index s) werden auf Grundlage ihrer berechneten Bragg-Peak-Position in ein $1 \times 1 \times 1$ cm Voxelraster zusammengeführt. Pro Voxel (Index v) werden die spotweisen PGI-Reichweiteabweichungen r_s nach Gewichtung durch die Protonenanzahl p_s gemittelt sowie die kumulative Protonenanzahl p_v bestimmt. (2) Benachbarte Voxel werden zu Clustern kombiniert, die eine kumulative Protonenanzahl $\geq 3 \cdot 10^9$ aufweisen müssen. Pro Cluster (Index c) werden die voxelweisen PGI-Reichweiteabweichungen r_v nach Gewichtung durch die Protonenanzahl p_v gemittelt sowie die kumulative Protonenanzahl p_c bestimmt.

zu bilden (effektiv ein $2 \times 2 \times 2$ Voxel-Block). Je nach Anordnung der Spots und deren Spotgewichte im Patienten, kann es passieren, dass Voxel keinem Cluster zugeordnet werden können und damit verworfen werden. Die prinzipiell variablen Cluster-Parameter wurden aufgrund von Erfahrungswerten so festgelegt, dass $\leq 10\%$ der aggregierten Spotgewichte zusätzlich verworfen werden müssen. Der Cluster-Algorithmus bearbeitet die Voxelmatrix linear und deterministisch, um dabei Cluster ohne eine gewisse Vorzugsrichtung (z. B. in Strahlrichtung) zu bilden. Die resultierenden Cluster sind variabel in Volumen und Form, sodass diese potentiell die örtliche Auflösung besser erhalten als bei einer Erhöhung der Voxelgröße.

Als Teiluntersuchung wurde für alle Cluster, wie auch auf Spotebene, die Pearson-Korrelation und die Differenzverteilung zwischen den clusterweisen ITD- und PGI-Reichweiteabweichungen bestimmt. Die Reichweiteabweichung für jedes Cluster wurde dabei als mit der Protonenanzahl gewichteter Mittelwert der ITD- bzw. PGI-Reichweiteabweichungen für jeweils dieselben Spots berechnet (Abbildung 5.6).

5.3.2 Definition von spot- oder clusterbasierten Klassifikationsmodellen

Analog zur feldweisen ITD-Klassifikation wurde anhand der spot- oder clusterweisen PGI-Reichweiteabweichungen jeweils ein PGI-basiertes feldweises Klassifikationsmodell erstellt.

Beim *spotbasierten* Klassifikationsmodell wird das relative Spotgewicht $W_{rel,PGI,Spot}$ aller für die PGI-Auswertung aggregierten und gefilterten Spots mit Reichweiteabweichungen $|r_{s,PGI}| \geq 5$ mm als Parameter für die Optimierung mit der ITD-basierten Grundwahrheit bestimmt. Ein Feld wird als relevant oder nicht relevant klassifiziert, wenn der Parameter

$W_{\text{rel,PGI,Spot}}$ über oder unter einem gewissen Schwellwert $T_{\text{rel,PGI,Spot}}$ liegt, wobei $W_{\text{rel,PGI,Spot}}$ gegeben ist durch

$$W_{\text{rel,PGI,Spot}} = \frac{\sum_{s(|r_{s,\text{PGI}}| \geq 5 \text{ mm})} w_s}{\sum_s w_s} \begin{cases} < T_{\text{rel,PGI,Spot}}: \text{ nicht relevant} \\ \geq T_{\text{rel,PGI,Spot}}: \text{ relevant} \end{cases} \quad (5.4)$$

mit $s \in \{1, \dots, N_f\}$ und N_f der Anzahl gefilterter Spots im Bestrahlungsfeld.

Beim *clusterbasierten* Klassifikationsmodell wird die relative Anzahl an ermittelten Clustern $W_{\text{rel,PGI,Cluster}}$ mit Reichweiteabweichungen $|r_{c,\text{PGI}}| \geq 5 \text{ mm}$ als Parameter für die Optimierung mit der ITD-basierten Grundwahrheit bestimmt. D. h. Cluster werden hier vereinfachend als gleichgewichtig angesehen, obwohl das kumulative Spotgewicht in verschiedenen Clustern sich prinzipiell unterscheiden kann. Ein Feld wird als relevant oder nicht relevant klassifiziert, wenn der Parameter $W_{\text{rel,PGI,Cluster}}$ über oder unter einem gewissen Schwellwert $T_{\text{rel,PGI,Cluster}}$ liegt, wobei $W_{\text{rel,PGI,Cluster}}$ gegeben ist durch

$$W_{\text{rel,PGI,Cluster}} = \frac{\sum_{c(|r_{c,\text{PGI}}| \geq 5 \text{ mm})} 1}{N_c} \begin{cases} < T_{\text{rel,PGI,Cluster}}: \text{ nicht relevant} \\ \geq T_{\text{rel,PGI,Cluster}}: \text{ relevant} \end{cases} \quad (5.5)$$

mit $c \in \{1, \dots, N_c\}$ und N_c der Anzahl an Clustern im Bestrahlungsfeld.

Zur Beurteilung der Klassifikationsmodelle wird die ROC-Kurve (Operationscharakteristik eines Beobachters, engl. *receiver operating characteristic*) und ihre Fläche unter der Kurve, engl. *area under the curve* (AUC), durch Variation des jeweiligen Schwellwertes $T_{\text{rel,PGI,Spot}} \in [0; 100]\%$ oder $T_{\text{rel,PGI,Cluster}} \in [0; 100]\%$ in 0,5% Schritten ermittelt. Die Variationen der AUC, der Sensitivität und der Spezifität wurde mittels Bootstrapping-Verfahren für insgesamt 10 000 zufällige Stichproben aus der Grundgesamtheit an Feldern (Ziehen mit Zurücklegen) untersucht. Dabei enthielt jede Stichprobe genauso viele Felder wie die Grundgesamtheit. Das 95% Konfidenzintervall ($KI_{95\%}$) wird als Intervall zwischen den Quantilen 2,5 und 97,5 der entsprechenden Verteilung bestimmt.

Für die Etablierung der Modelle sind zwei Schritte notwendig: (1) die Bestimmung des optimalen Schwellwertes $T_{\text{rel,PGI,Spot}}$ bzw. $T_{\text{rel,PGI,Cluster}}$ basierend auf einer Trainingskohorte und (2) die Evaluation von Sensitivität und Spezifität des Klassifikationsmodells mit optimalem Schwellwert für eine unabhängige Testkohorte. Dafür wurden die Bestrahlungsfelder und Patienten entsprechend Tabelle 5.3 (Seite 77) und Tabelle 5.6 (Seite 81) aufgeteilt. Bei der Aufteilung der Kohorten wurde der Anteil relevanter und nicht relevanter Felder in der ITD-basierten Grundwahrheit berücksichtigt. Zusätzlich wurde bei den Prostata-Tumor-Patienten auf den Zeitpunkt des Monitorings der Patienten geachtet, sodass die Trainings- und Testkohorte jeweils Daten aus dem gesamten Beobachtungszeitraum enthalten. Trotz des geringen Umfangs der Kohorte an Kopf-Hals-Tumor-Patienten konnte bei der Aufteilung der Kohorten

erreicht werden, dass kein Bestrahlungsfeld (d. h. Strahlanordnung) sowohl in Trainings- als auch Testkohorte vorkommt und ein Patient in jeder Kohorte völlig unabhängig ist.

Der optimale Schwellwert $T_{rel,PGI,Spot}$ oder $T_{rel,PGI,Cluster}$ des jeweiligen Klassifikationsmodells wird anhand der Trainingskohorte über die Maximierung der Übereinstimmungswahrscheinlichkeit, engl. *concordance probability* (CP), bestimmt (Liu, 2012):

$$\text{Sensitivität} = \frac{TP}{TP+FN}, \quad (5.6)$$

$$\text{Spezifität} = \frac{TN}{TN+FP}, \quad (5.7)$$

$$\text{CP} = \text{Sensitivität} \times \text{Spezifität} = \frac{TP \cdot TN}{(TP+FN)(TN+FP)}. \quad (5.8)$$

Hierbei bezieht sich TP (engl. *true positive*) und TN (engl. *true negativ*) auf die Anzahl der durch das PGI-Modell richtig klassifizierten relevanten und nicht relevanten Behandlungsfelder gegenüber der ITD-basierten Grundwahrheit. Folglich bezieht sich FP (engl. *false positive*) und FN (engl. *false negative*) auf die Anzahl an Behandlungsfeldern, die vom PGI-Modell fälschlicherweise als relevant oder nicht relevant klassifiziert wurden.

5.4 Ergebnisse der PGI-Detektionsfähigkeit

5.4.1 Auswertung für Patienten mit Prostata-Tumor

Für die Auswertung der Prostata-Tumor-Patienten wurden Spots herausgefiltert, deren Bragg-Peak-Position nicht im effektiven FOV des PGI-Systems liegt, welches auf der zentralen Strahlachse des Bestrahlungsplans einen Bereich von [-5; 5] cm relativ zur Projektion der Schlitzeblende umfasst. Als zusätzliche Einschränkung werden nur PGI-Signale von Spots mit mindestens $5 \cdot 10^7$ applizierten Protonen und einem 5 mm breiten Plateaubereich nach dem Signalabfall betrachtet, wie bereits in einer Vorarbeit untersucht (Khamfongkhrua et al., 2020). Neben diesen Spotfiltern (Tabelle 5.1) werden die in Unterkapitel 3.5.4 eingeführten heuristischen Korrekturen angewandt. In Abbildung 5.2 sind Beispiele für die spotweise Korrelation von PGI und ITD-Reichweiteabweichungen einzelner Felder gezeigt.

Auf Spot- und Clusterebene konnte über alle monitorierten Felder eine gute Korrelation zwischen den aggregierten und gefilterten PGI- und ITD-Reichweiteabweichungen mit einem Pearson-Korrelationskoeffizienten von $\rho = 0,57$ ($p < 0,01$) bzw. $\rho = 0,67$ ($p < 0,01$) gezeigt werden (Abbildung 5.7). Es ergab sich eine mittlere gewichtete Abweichung zwischen PGI und ITD-Reichweiteabweichungen von 0,5 mm für Spots und Cluster sowie eine gewichteten Standardabweichung von 2,6 mm für Spots und 2,1 mm für Cluster. Diese Streuung zwischen den PGI- und ITD-Reichweiteabweichungen verdeutlicht, dass es eine feldweise Klassifikation der PGI-Daten benötigt, um Abweichungen zuverlässig erkennen zu können.

Tabelle 5.1: Auswirkung der Verwendung von Spotfiltern und des Clusterings auf den Datensatz der 15 Prostata-Tumor-Patienten. Pro Bestrahlungsfeld werden Median, Minimum (Min) und Maximum (Max) für die entsprechenden Verarbeitungsschritte im Bezug auf die applizierten Monitoreinheiten (MU), die Spot-, Voxel- und Clusteranzahl sowie für die Zielvolumina angegeben.

Parameter	Median (Min–Max)
MU pro Behandlungsfeld	445,1 (375,4–515,5)
MU nach FOV-Filter	97,9% (89,1%–99,9%)
MU nach FOV- und Protonenanzahl-Filter	96,2% (88,0%–99,0%)
MU nach Clustering	94,0% (84,2%–98,9%)
Spots pro Behandlungsfeld	1595 (1298–4188)
Spots nach FOV-Filter	95,0% (83,9%–99,8%)
Spots nach FOV- und Protonenanzahl-Filter	85,4% (81,0%–94,5%)
Spots nach Clustering	81,2% (70,5%–94,1%)
Voxel pro Behandlungsfeld	392 (293–465)
Genutzter Voxelanteil in Clustering	92,4% (74,9%–100,0%)
Genutzter MU-Anteil in Clustering	98,0% (91,0%–100,0%)
Cluster pro Behandlungsfeld	88 (62–159)
Prostatavolumen / cm ³	42,4 (29,3–65,3)
CTV _{60 Gy} -Volumen / cm ³	78,2 (54,5–108,1)
CTV _{48 Gy} -Volumen / cm ³	143,4 (101,4–180,3)

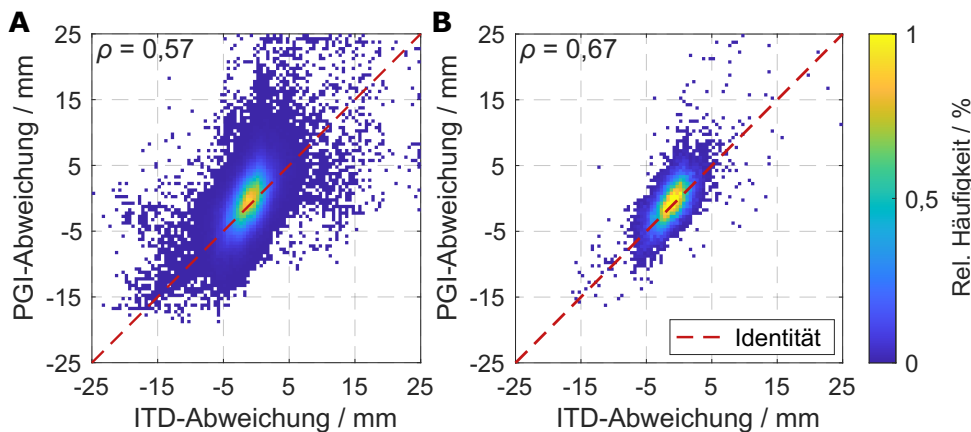


Abbildung 5.7: Spotweise (A) und clusterweise (B) Korrelation von PGI- und ITD-Reichweiteabweichungen in einem 2D Histogramm über alle Prostata-Tumor-Patienten. Die rote, gestrichelte Linie markiert eine Eins-zu-eins-Korrelation zwischen den beiden Variablen.

Für die Untersuchung des feldweisen PGI Klassifikationspotentials (primärer Endpunkt) sind die ermittelten AUC-Werte für das spot- und clusterbasierte PGI-Klassifikationsmodell mit der Hypothese einer zufälligen Klassifikation (AUC = 0,5) *nicht* verträglich. Mit einer AUC von etwa 0,7 und 0,8 auf der Trainings- bzw. Testkohorte (Tabelle 5.2 und Abbildung 5.8) zeigt sich sowohl für das spot- als auch das clusterbasierte PGI-Klassifikationsmodell der klinische Nutzen der PGI-Behandlungsverifikation für die automatische Detektion von anatomischen Veränderungen während der Bestrahlung. Für die Optimierung der Modelle auf der Trainingskohorte wurden die optimalen Schwellwerte $T_{rel,PGI,Spot} = 6,0\%$ und $T_{rel,PGI,Cluster} = 1,5\%$

Tabelle 5.2: Ergebnisse für die AUC, Sensitivität (Sens) und Spezifität (Spez) der spot- und clusterweisen PGI-Klassifikationsmodelle gegenüber der ITD-basierten Grundwahrheit auf der Trainings- und Testkohorte von Prostata-Tumor-Patienten.

	Spotbasiertes Modell		Clusterbasiertes Modell	
	Training	Test	Training	Test
AUC [KI_{95%}]	0,72 [0,63; 0,81]	0,80 [0,68; 0,90]	0,69 [0,60; 0,77]	0,82 [0,71; 0,91]
Sens [KI_{95%}] / %	63 [51; 75]	74 [58; 89]	68 [56; 79]	81 [66; 94]
Spez [KI_{95%}] / %	76 [65; 86]	79 [66; 92]	67 [55; 78]	72 [57; 85]

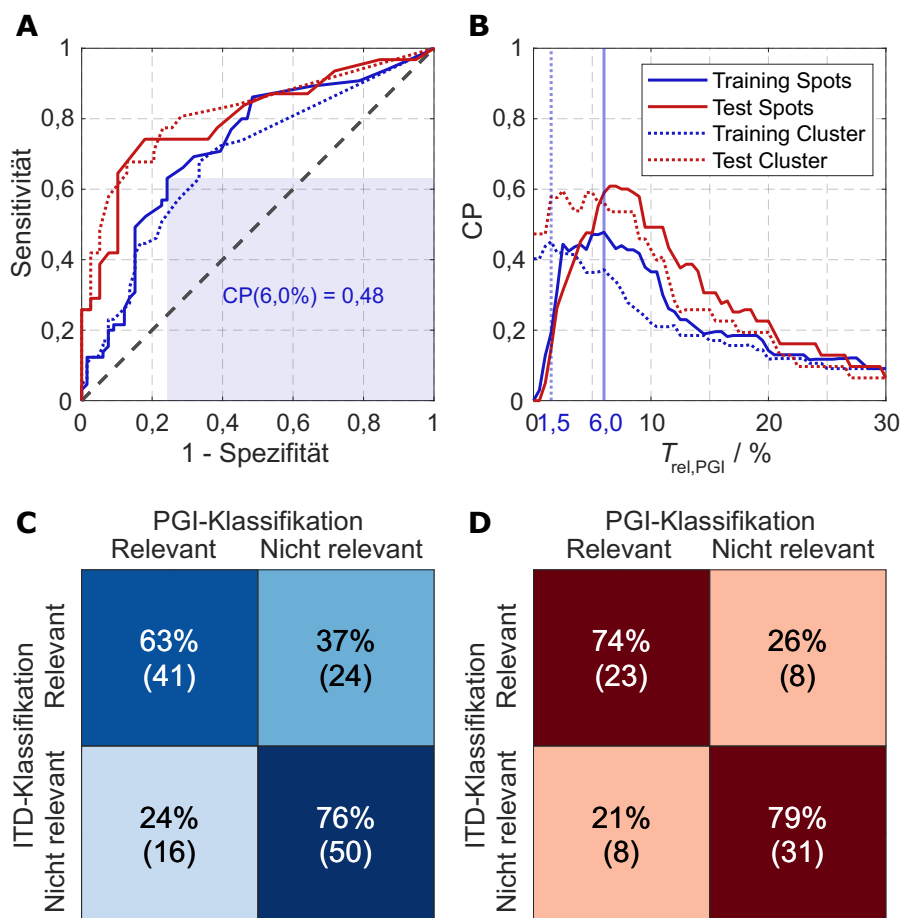


Abbildung 5.8: Ergebnisse für die Optimierung der PGI-Klassifikationsmodelle gegenüber der ITD-basierten feldweisen Grundwahrheit für Prostata-Tumor-Patienten. (A) ROC-Kurve für das spot- und clusterbasierte Klassifikationsmodell unterteilt in Trainings- und Testkohorte. Das Rechteck unter der ROC-Kurve für die Trainingskohorte im spotbasierten Modell visualisiert den optimalen Schwellwert mit maximaler Übereinstimmungswahrscheinlichkeit (CP = Sensitivität × Spezifität). (B) Verlauf der Übereinstimmungswahrscheinlichkeit in Abhängigkeit vom Schwellwert $T_{rel,PGI}$ für Trainings- und Testkohorte im spot- und clusterbasierten Modell. Wahrheitsmatrizen des spotbasierten Klassifikationsmodells für die Trainingskohorte (C) und Testkohorte (D), welche sich aus der Anwendung des optimalen Schwellwertes $T_{rel,PGI, Spot} = 6,0\%$ ergeben.

für das spot- bzw. clusterbasierte PGI-Klassifikationsmodell bestimmt. Anhand von Abbildung 5.8 B lässt sich erkennen, dass sich die Übereinstimmungswahrscheinlichkeit (CP) im Bereich des Maximums für das jeweilige Modell in einem stabilen Plateaubereich befindet, sodass eine Verringerung der Intervallschritte nur zu einer Überanpassung an den Trainingsdatensatz führen würde. Mit diesen Schwellwerten konnte die Sensitivität und die Spezifität auf Trainings- und Testkohorte für das jeweilige PGI-Klassifikationsmodell quantifiziert werden (Tabelle 5.2). Für beide Modelle war die Sensitivität und Spezifität auf der unabhängigen Testkohorte größer als auf der Trainingskohorte und ergab mindestens 72%. Da sich das spot- und clusterbasierte PGI-Klassifikationsmodell nicht wesentlich in ihrem Klassifikationspotential unterscheiden, wird das einfachere spotbasierte Klassifikationsmodell als PGI-Klassifikationsmodell für den klinischen Einsatz vorgeschlagen.

In Tabelle 5.3 ist die Aufteilung der Patienten in Trainings- und Testkohorte sowie die entsprechende ITD-Klassifikation und optimierte spotbasierte PGI-Klassifikation mit Sensitivität und Spezifität je Patient dargestellt. Es zeigt sich, dass eine mögliche Ursache für das höhere Klassifikationsvermögen des PGI-Modells auf der Testkohorte in der Aufteilung der einzelnen Patienten liegt. Beispielsweise traten bei den Patienten P6 und P15 nur nicht relevante bzw. relevante Veränderungen auf, wohingegen die Verteilung von relevanten und nicht relevanten Feldern für die restlichen Patienten wesentlich ausgeglichener ist. Ein weiterer Grund könnte sein, dass alle vier Prostata-Tumor-Patienten mit Rektumballon (38 Felder, vgl. Tabelle 3.2) zufälligerweise nur der Trainingskohorte zugeordnet wurden. Es gibt jedoch keinen objektiven Grund, warum gerade für diese Patienten anatomische Veränderungen weniger gut erkannt werden sollten, zumal sich die Häufigkeit der beobachteten anatomischen Veränderungen in der manuellen Klassifikation zwischen Patienten mit und ohne Rektumballon nicht merklich unterscheidet. Schlussendlich ist bei der Betrachtung des 95% Konfidenzintervalls für die AUC von Trainings- und Testkohorte auch eine statistische Ursache als Erklärung für das unterschiedliche Klassifikationsvermögen wahrscheinlich.

Tabelle 5.3: Aufteilung der mit PGI monitorierten Prostata-Tumor-Patienten nach Trainings- und Testkohorte für die Optimierung des PGI-Modells und Überblick zur zugrundeliegenden ITD-Klassifikation und optimierten spotbasierten PGI-Klassifikation (NR = nicht relevant, R = relevant) sowie die Betrachtung der Sensitivität und Spezifität pro Patient (vgl. Tabelle 3.2).

Kohorte Patient	Training										Test				
	P1	P2	P4	P5	P7	P8	P10	P12	P13	P14	P3(1+2)	P6	P9	P11	P15
Monitorierte Felder	8	18	14	18	12	2	17	10	14	18	14	10	18	14	14
Feldweise ITD	6	9	10	5	5	-	7	6	9	9	10	10	12	7	-
Klassifikation	2	9	4	13	7	2	10	4	5	9	4	-	6	7	14
Feldweise PGI	6	9	7	10	8	2	5	5	10	12	12	7	8	12	-
Klassifikation	2	9	7	8	4	-	12	5	4	6	2	3	10	2	14
Spezifität / %	83	78	60	60	80	-	57	83	78	100	100	70	58	100	-
Sensitivität / %	50	78	75	46	42	0	90	100	40	66	50	-	83	29	100

5.4.2 Auswertung für Patienten mit Tumoren im Kopf-Hals-Bereich

Für die Auswertung der Kopf-Hals-Tumor-Patienten wurden die gleichen Spotfilter wie bei den Prostata-Tumor-Patienten angewandt (vgl. Unterkapitel 5.4.1), wobei die Schwelle für die applizierte Protonenanzahl in einem Spot auf $\geq 2,5 \cdot 10^7$ reduziert wurde, um das Clustering der Spots zu vereinfachen. Zudem wurde ein zusätzlicher Spotfilter eingeführt (Tabelle 5.4), der nur Spot mit einer Bragg-Peak-Position innerhalb einer um 5 mm reduzierten Außenkontur des Patienten berücksichtigt. Damit wird sichergestellt, dass nicht vollständig im Patienten stoppende Spots von der Auswertung ausgeschlossen werden (Khamfongkhrua et al., 2020). Heuristische Korrekturen wie für die Prostata-Tumor-Patienten (vgl. Unterkapitel 3.5.4) kommen hierbei *nicht* zur Anwendung, da viele der Bestrahlungsfelder unter nicht horizontalem Gantrywinkel appliziert wurden und ein Range-Shifter zum Einsatz kam.

Auf Spot- und Clusterebene konnte über alle monitorierten Bestrahlungsfelder eine mäßige Korrelation zwischen den aggregierten und gefilterten PGI- und ITD-Reichweiteabweichungen mit einem Pearson-Korrelationskoeffizienten von $\rho = 0,34$ ($p < 0,01$) bzw. $\rho = 0,51$ ($p < 0,01$) gezeigt werden (Abbildung 5.9). Es ergab sich eine mittlere gewichtete Abweichung zwischen den PGI- und ITD-Reichweiteabweichungen von 1,5 mm für Spots und Cluster mit einer gewichteten Standardabweichung von 4,5 mm für Spots und 3,1 mm für Cluster. Damit ist die Streuung zwischen den PGI- und ITD-Reichweiteabweichungen bei den Kopf-Hals-Tumor-Patienten noch einmal deutlich höher als bei den Prostata-Tumor-Patienten, was u. a. auf die insgesamt geringere Dosis der Spots pro Fraktion und damit die statistischen Unsicherheiten der PGI-Messung zurückzuführen ist. In Abbildung 5.3 sind Beispiele für die spotweise Korrelation von PGI- und ITD-Reichweiteabweichungen einzelner Felder gezeigt.

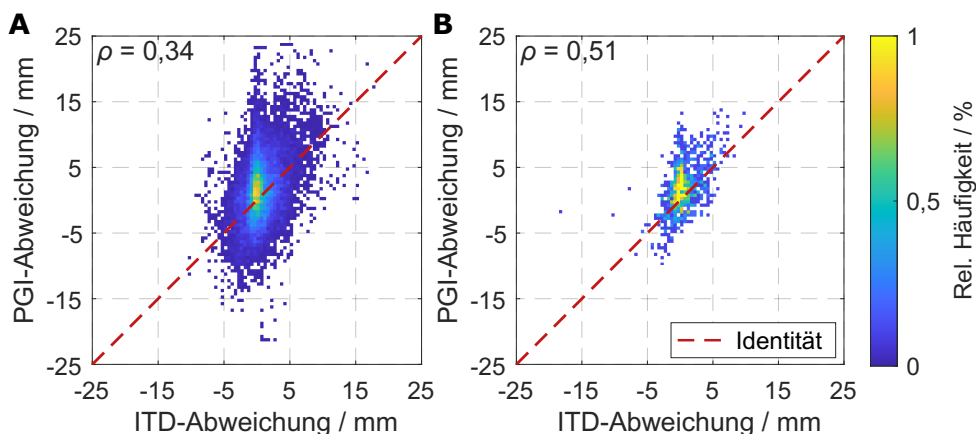


Abbildung 5.9: Spotweise (A) und clusterweise (B) Korrelation von PGI- und ITD-Reichweiteabweichungen in einem 2D Histogramm über alle Kopf-Hals-Tumor-Patienten. Die rote, gestrichelte Linie markiert eine Eins-zu-eins-Korrelation zwischen den beiden Variablen.

Tabelle 5.4: Auswirkung der Verwendung von Spotfiltern und des Clusterings auf den Datensatz der 4 Kopf-Hals-Tumor-Patienten. Pro Bestrahlungsfeld werden Median, Minimum (Min) und Maximum (Max) für die entsprechenden Verarbeitungsschritte im Bezug auf die applizierten Monitoreinheiten (MU), die Spot-, Voxel- und Clusteranzahl sowie für die Zielvolumina angegeben.

Parameter	Median (Min–Max)
MU pro Behandlungsfeld	462,2 (282,1–671,1)
MU nach FOV-Filter	45,9% (30,3%–79,5%)
MU nach FOV- und Kontur-Filter	31,8% (20,0%–58,0%)
MU nach FOV-, Kontur- und Protonenanzahl-Filter	31,2% (16,7%–55,2%)
MU nach Clustering	30,8% (16,6%–55,2%)
Spots pro Behandlungsfeld	2756 (986–3108)
Spots nach FOV-Filter	50,5% (48,0%–88,4%)
Spots nach FOV- und Kontur-Filter	37,1% (35,0%–73,0%)
Spots nach FOV-, Kontur- und Protonenanzahl-Filter	32,2% (20,6%–52,6%)
Spots nach Clustering	30,0% (20,5%–52,6%)
Voxel pro Behandlungsfeld	128 (60–205)
Genutzter Voxelanteil in Clustering	98,4% (95,9%–100,0%)
Genutzter MU-Anteil in Clustering	99,7% (98,5%–100,0%)
Cluster pro Behandlungsfeld	36,5 (27–87)
CTV _{50 Gy} -Volumen / cm ³	139,6 (127,4–151,8)
CTV _{Boost} -Volumen / cm ³	66,9 (38,7–88,3)

Aufgrund des geringen Stichprobenumfangs der mit PGI monitorierten Bestrahlungen in Kopf-Hals-Tumor-Patienten zeigt sich für die Untersuchung des feldweisen PGI-Klassifikationspotentials (primärer Endpunkt) beim spot- und clusterbasierten PGI-Klassifikationsmodell kein Unterschied zur Hypothese einer zufälligen Klassifikation mit $AUC = 0,5$ (Tabelle 5.5 und Abbildung 5.10). Für die Optimierung der Modelle auf der Trainingskohorte wurden die optimalen Schwellwerte $T_{rel,PGI,Spot} = 28,0\%$ und $T_{rel,PGI,Cluster} = 20,0\%$ für das spot- bzw. clusterbasierte PGI-Modell bestimmt. Anhand von Abbildung 5.10 B lässt sich erkennen, dass diese Optima zumindest für Trainings- und Testkohorte übereinstimmen, wenngleich der Wert der Übereinstimmungswahrscheinlichkeit sich zwischen beiden Kohorten deutlich unterscheidet. Mit diesen Schwellwerten konnte die Sensitivität und die Spezifität auf Trainings- und Testkohorte für das jeweilige PGI-Klassifikationsmodell quantifiziert werden (Tabelle 5.5). Wie sich aus dem Verlauf der ROC-Kurven erwarten ließ (Abbildung 5.10 A), schwankt die Sensitivität und Spezifität der jeweiligen Modelle zwischen Trainings- und Testkohorte teilweise deutlich. Um eine statistisch gesicherte Aussage über das PGI-Klassifikationspotential zur Erkennung anatomischer Veränderungen bei Kopf-Hals-Tumor-Bestrahlungen treffen zu können, bedarf es grundsätzlich eines größeren Stichprobenumfangs. Da sich auch hier, wie für die Prostata-Tumor-Patienten, das spot- und clusterbasierte PGI-Klassifikationsmodell nicht wesentlich in ihrem Klassifikationspotential unterscheiden, wird das einfachere spotbasierte Klassifikationsmodell als PGI-Klassifikationsmodell für den potentiellen klinischen Einsatz vorgeschlagen.

Tabelle 5.5: Ergebnisse für die AUC, Sensitivität (Sens) und Spezifität (Spez) der spot- und clusterweisen PGI-Klassifikationsmodelle gegenüber der ITD-basierten Grundwahrheit auf der Trainings- und Testkohorte von Kopf-Hals-Tumor-Patienten.

	Spotbasiertes Modell		Clusterbasiertes Modell	
	Training	Test	Training	Test
AUC [KI_{95%}]	0,54 [0,11; 0,90]	0,75 [0,38; 1,00]	0,50 [0,11; 0,88]	0,86 [0,50; 1,00]
Sens [KI_{95%}] / %	50 [0; 1]	100 [50 ^a ; 100]	50 [0; 1]	100 [50 ^a ; 100]
Spez [KI_{95%}] / %	71 [43; 100]	71 [43; 100]	71 [43; 100]	86 [57; 100]

^a Maximale untere Grenze für die Annahme einer zufälligen Klassifikation durch das PGI-Klassifikationsmodell.

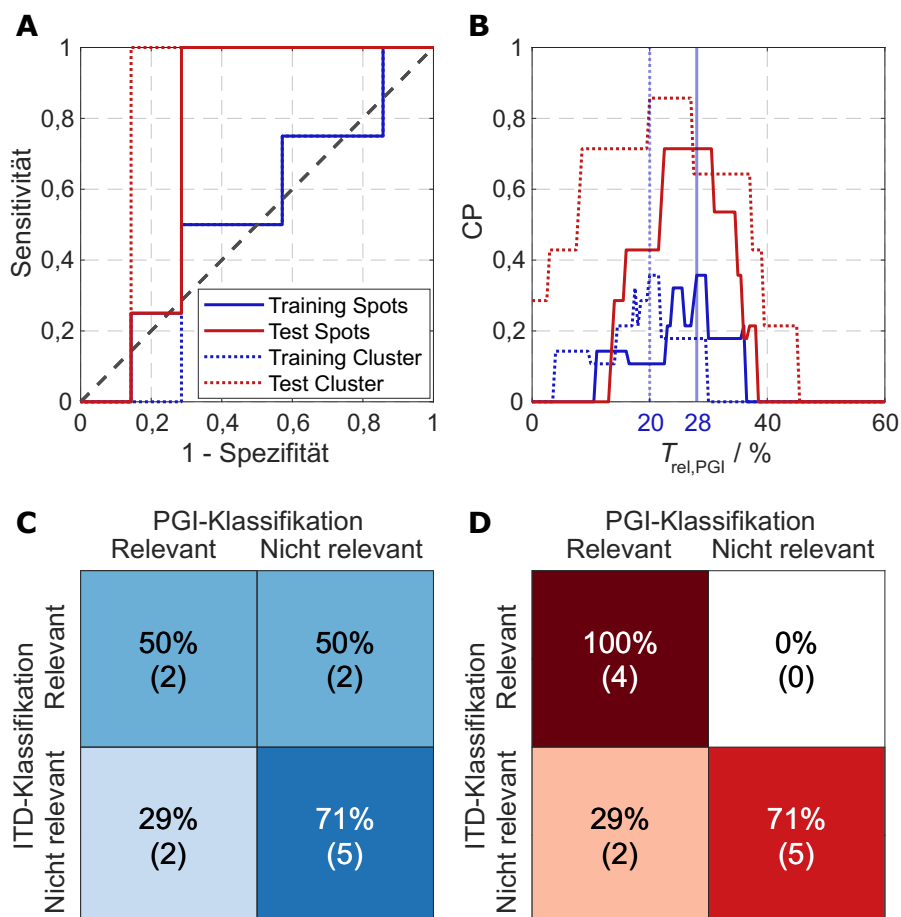


Abbildung 5.10: Ergebnisse für die Optimierung der PGI-Klassifikationsmodelle gegenüber der ITD-basierten feldweisen Grundwahrheit für Kopf-Hals-Tumor-Patienten. (A) ROC-Kurve für das spot- und clusterbasierte Klassifikationsmodell unterteilt in Trainings- und Testkohorte. (B) Verlauf der Übereinstimmungswahrscheinlichkeit (CP = Sensitivität × Spezifität) in Abhängigkeit vom Schwellwert $T_{rel,PGI}$ für Trainings- und Testkohorte bzw. spot- und clusterbasiertes Modell. Wahrheitsmatrizen des spotbasierten Klassifikationsmodells für die Trainingskohorte (C) und Testkohorte (D), welche sich aus der Anwendung des optimalen Schwellwertes $T_{rel,PGI, Spot} = 28,0\%$ ergeben.

Tabelle 5.6: Aufteilung der mit PGI monitorierten Kopf-Hals-Tumor-Patienten nach Trainings- und Testkohorte für die Optimierung des PGI-Modells und Überblick zur zugrundeliegenden ITD-Klassifikation und optimierten spotbasierten PGI-Klassifikation (NR = nicht relevant, R = relevant) (vgl. Tabelle 3.3).

Kohorte Patient Gantrywinkel	Training				Test			
	KH1(2) 280°	KH2(1) 80°	KH3(1) 270°	KH3(2) 270°	KH1(2) 90°	KH2(1) 280°	KH2(2) 270°	KH4(2) 275°
Monitorierte Felder	1	5	4	1	1	5	1	4
Feldweise ITD- Klassifikation	NR R	1 2	3 1	- 1	1 -	5 -	1 -	- 4
Feldweise PGI- Klassifikation	NR R	1 4	1 -	4 -	1 1	5 -	- 1	- 4

In Tabelle 5.6 ist die Aufteilung der Patienten in Trainings- und Testkohorte sowie die entsprechende ITD-Klassifikation und optimierte spotbasierte PGI-Klassifikation je Patient dargestellt. Hieran wird noch einmal deutlich, dass die Ursache für das höhere Klassifikationsvermögen des PGI Modells in der Testkohorte im Allgemeinen auf den geringen Umfang der beiden Kopf-Hals-Kohorten zurückzuführen ist. Dabei stellt die Aufteilung der einzelnen Patienten einen zusätzlichen Einflussfaktor dar. Die Entscheidung, einzelne Bestrahlungsfelder bei der Kohortenbildung nicht aufzuteilen, damit die Testkohorte so unabhängig wie möglich von der Trainingskohorte ist, führte zwangsläufig dazu, dass der Patient KH4 mit seinem Bestrahlungsfeld alle relevanten Felder in der Testkohorte stellt. Die Abweichungen für dieses Bestrahlungsfeld waren so deutlich, dass die vier Felder auch mit PGI als relevant klassifiziert wurden, wohingegen die vier relevanten Felder von drei unterschiedlichen Bestrahlungsfeldern in der Trainingskohorte nur zur Hälfte mit PGI erkannt worden sind.

5.5 Diskussion

In dieser Untersuchung wurden mittels eines PGI-Schlitzkamera-Systems der zweiten Generation systematisch klinische Daten über einen Zeitraum von zweieinhalb Jahren aufgenommen und ausgewertet. Anhand von in-room kCT-Scans in Behandlungsposition und der daraus folgenden anatomischen und dosimetrischen Grundwahrheit konnte erstmals die Fähigkeit von PGI zur Detektion von anatomischen Veränderungen in klinischen Protonenbestrahlungen quantifiziert und damit nachgewiesen werden.

Bedeutung der Untersuchung

Die hier systematisch durchgeführte Untersuchung anhand von 15 Prostata- und 4 Kopf-Hals-Tumor-Patienten konnte erstmals den klinischen Nutzen der PGI-Reichweiteverifikation für die Erkennung von anatomischen Veränderungen während der Behandlung nachweisen. Insgesamt konnten dadurch Untersuchungen zur klinischen Einzelfallanwendung

von PGI bestätigt und ergänzt werden (Richter et al., 2016; Xie et al., 2017, 2020). Die dafür notwendige Qualität und Konsistenz der aufgenommenen PGI-Daten über den gesamten Beobachtungszeitraum konnte durch regelmäßige Kalibrierung des Detektors und Validierung der Position des Systems zum Isozentrum gewährleistet werden (etwa aller 2–3 Monate). Das etablierte (spotbasierte) PGI-Klassifikationsmodell erreichte für die Prostata-Tumor-Patienten eine für den klinischen Einsatz vielversprechende Sensitivität und Spezifität von 74% bzw. 79% auf einer unabhängigen Testkohorte. Die vorliegenden Ergebnisse für Kopf-Hals-Tumor-Bestrahlungen weisen in die gleiche Richtung wie die Auswertung der Prostata-Tumor-Bestrahlungen mit einer Sensitivität und Spezifität von 100% bzw. 71% auf der Testkohorte, wobei diese Werte aufgrund der geringen Anzahl von untersuchten Bestrahlungsfeldern (je 11 für Trainings- und Testkohorte) mit einer hohen statistischen Unsicherheit behaftet sind. Fast zwei Jahrzehnte nach der Bekanntmachung der Idee PG für die Reichweiteverifikation in der Protonentherapie einzusetzen (Jongen & Stichelbaut, 2003), markiert diese Untersuchung einen weiteren wichtigen Meilenstein im Hinblick auf die routinemäßige klinische Anwendung von PGI zur Behandlungskontrolle.

Bereits das aktuell entwickelte (spotbasierte) Modell zur Klassifikation von PGI-Daten wäre in der Lage, routinemäßig bei klinischen Prostata-Tumor-Bestrahlungen als unabhängige Verifikation der Behandlung zu dienen, indem es z. B. die Aufnahme eines kCT triggert. Im Vergleich dazu wurden die kCT-Scans für die untersuchte Patientenkohorte nur zwei Mal pro Woche durchgeführt, in der es meist bis zu fünf Bestrahlungen gab. Für eine solche prospektive Anwendung kann davon ausgegangen werden, dass jegliche Behandlungsabweichungen mit ähnlicher Genauigkeit erkannt werden, wie die hier untersuchten anatomischen Veränderungen. Voraussetzung dafür ist jedoch, dass diese Behandlungsabweichungen, z. B. verursacht durch die Patientenpositionierung oder die Strahlanwendung, entsprechend der etablierten Grundwahrheit als relevant oder nicht relevant deklariert werden. Es ist diese erwartete Genauigkeit gegenüber jeglichen Abweichungen, weshalb die PGI-Reichweiteverifikation als Methode für die Qualitätssicherung der Bestrahlung betrachtet werden kann. PGI könnte demnach potentiell eine wichtige Rolle bei der Ermöglichung einer OAPT spielen, bei der auf Veränderungen in der Behandlung mit einer sofortigen Adaption des Bestrahlungsplans reagiert werden soll. Für die routinemäßige Anwendung des Klassifikationsmodells für Kopf-Hals-Tumore sind noch Untersuchungen an weiteren Patienten notwendig. Hauptgründe hierfür sind die geringe Anzahl der PGI-monitorierten Felder, die Bestrahlung auch unter nicht horizontalen Gantrywinkeln und die sich stärker unterscheidenden anatomischen Gegebenheiten im Bezug auf das zu durchdringende Patientengewebe.

Hintergründe für Fehlklassifikationen

Für die zukünftige Anwendung des PGI-Klassifikationsmodells in Prostata-Tumor-Bestrahlungen könnte eine höhere Spezifität oder genauer gesagt Präzision der Klassifikation hilfreich sein, um den Arbeitsaufwand von falsch positiven Ereignissen zu reduzieren. In den

hier vorgestellten Ergebnissen wurde der Schwellwert zur Klassifizierung anhand der Übereinstimmungswahrscheinlichkeit optimiert; prinzipiell kann jedoch jeder Punkt auf der ROC-Kurve zur Klassifizierung ausgewählt werden, welche den Zusammenhang zwischen Sensitivität und Spezifität bei Variation des Schwellwerts beschreibt (vgl. Abbildung 5.8). Ermöglicht durch die geringe Komplexität des auf spotweisen Reichweiteabweichungen basierenden Klassifikationsmodells wurde in diesem Zusammenhang ein genauerer Blick auf potentielle Gründe für Fehlklassifikationen geworfen. Folgendes wird vermutet: (1) Grenzfälle in der Klassifikation der Grundwahrheit (z. B. Felder mit punktuellen lokalen Veränderungen) können falsch mit PGI klassifiziert werden, da diese mit der auf der gesamten Trainingskohorte optimierten Schwelle des PGI-Klassifikationsmodells schwerer zu erkennen sind. (2) Wegen globaler Positionierungsunsicherheiten des PGI-Systems relativ zum Patienten (etwa 1,3 mm (2σ), vgl. Kapitel 4) können solche Felder gelegentlich falsch als relevant oder nicht relevant klassifiziert werden, deren Extremwerte bei den ITD-Reichweiteabweichungen nahe am Schwellwert (± 5 mm) liegen. (3) Andere Fehlklassifikationen können auf weitere Einschränkungen des PGI-Systems und der ITD-basierten Grundwahrheit zurückgeführt werden, wie im Folgenden diskutiert.

Limitationen der Studie

Eine grundlegende Limitation der Reichweiteverifikation besteht darin, dass nur solche Änderungen erkannt werden können, die Abweichungen in der Reichweite der Protonen bewirken. Jedoch gibt es durchaus anatomische Änderungen, welche die Reichweite kaum verändern, aber dennoch relevante Änderungen der Dosis im CTV und in OARs bewirken. Ein Beispiel dafür stellen kleine Verschiebungen der Prostata durch eine geänderte Blasen- oder Rektumfüllung dar (Abbildung 5.11), welche die SPR Verteilung nicht wesentlich verändern. Jedoch ist zu erwarten, dass größere Bewegungen von Weichgewebe innerhalb von dosisrelevanten Regionen im Endeffekt auch eine Veränderung des SPR herbeiführen werden. In dem speziellen Beispiel von 5.11 wurde die Bestrahlung – wie bei allen in dieser Studie mit PGI-monitorierten Fraktionen – planmäßig fortgeführt, da die implantierten Marker (PolyMark™, CIVCO Radiotherapy, U.S) gerade noch innerhalb der 3 mm Sicherheitssäume zu finden waren. Allerdings wurden im Beobachtungszeitraum dieser Studie auch Verschiebungen der Prostata registriert, bei denen die Behandlung aufgrund von Blasen- oder Rektumfüllung auf einen späteren Zeitpunkt am Behandlungstag verschoben werden musste. Typischerweise führte dies aber nicht zu einer Planadaption, da solche Veränderungen oftmals von Tag zu Tag variieren. Hieran wird deutlich, dass über die Verifikation mit PGI hinaus, die Verwendung von Tumormarkern und 3D-Bildgebung vor der Therapie weitere wichtige Säulen zur Reduktion von Sicherheitssäumen darstellen.

Die Verfügbarkeit von in-room kCT-Scans für die mit PGI monitorierten Fraktionen stellt eine wichtige Voraussetzung im Design dieser Studie dar. Im Vergleich zur Literatur über *in vivo* Reichweiteverifikation ist die am OncoRay aufgenommene Datenbasis, mit in der gleichen

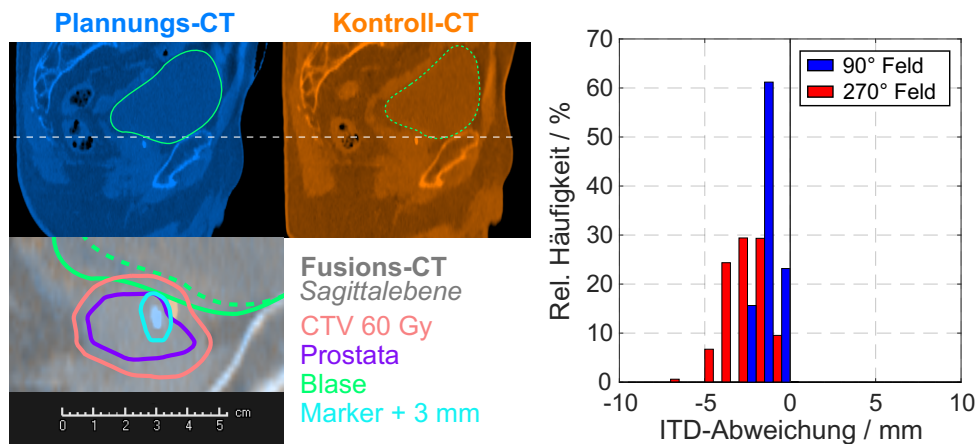


Abbildung 5.11: Beispiel für eine nach aktuellen klinischen Sicherheitssäumen nicht relevante anatomische Veränderung des Weichgewebes bei einer Prostatabestrahlung. Der Bestrahlungsplan wurde robust optimiert mit einer Reichweiteunsicherheit von 2% + 2 mm und einer durch Tumormarker überprüfte Positionierungsunsicherheit von 3 mm. Für die beiden lateralen Bestrahlungsfelder wurden die ITD-Reichweiteabweichungen der nicht aggregierten Spots in einem Histogramm dargestellt (es fließen auch Einflüsse anderer anatomischer Veränderungen ein), anhand dessen beide Felder auch nach der hier entwickelten Grundwahrheit als nicht relevant eingestuft wurden.

Fraktion aufgenommenen kCT-Scans und PGI-Aufnahmen, weltweit einzigartig. Nichtsdestotrotz stellen intrafraktionelle Veränderungen, die zufällig zwischen dem kCT-Scan und der Strahlanwendung auftreten können, eine untergeordnete Limitation der Daten dar. Für die Ergebnisse bedeuten diese Veränderungen eine potentielle Unsicherheit in der Grundwahrheit. Sie tragen somit zu der mittels Bootstrapping bestimmten Streubreite der AUC bei, wie etwa auch statistische Unsicherheiten in der Strahlanwendung oder der Patientenpositionierung.

Wie bereits erwähnt, kamen für die Bestimmung der Grundwahrheit anhand von ITD-Profilen historisch gewachsen zwei unterschiedliche Ansätze bei der Ermittlung der ITD-Reichweiteabweichungen zum Einsatz: Für die Auswertung der Prostata-Patienten wurde der *R80* verwendet, wohingegen für die Auswertung der Kopf-Hals-Tumor-Patienten ein Anpassungsverfahren Anwendung fand. Für die Bestimmung der PGI-Reichweiteabweichungen kam im Wesentlichen auch ein Anpassungsverfahren mit Fokus auf dem distalen Abfall der PGI-Profile zum Einsatz. Für die Prostata-Tumor-Patienten zeigte sich nachträglich ein Pearson-Korrelationskoeffizient von 0,93 zwischen dem *R80* und der im späteren Verlauf berechneten Reichweiteabweichung mittels Anpassung der distalen Flanke, wohingegen dieser für die Kopf-Hals-Tumor-Patienten nur 0,33 betrug. Dies zeigt, dass (1) die Verwendung des *R80* keine wesentliche Einschränkung bei den Prostata-Tumor-Patienten darstellt und (2) die Etablierung der Methode zur Anpassung der distalen Flanke notwendig für die Auswertung der Kopf-Hals-Tumor-Patienten gewesen ist. Grundlegende systematische Unterschiede zwischen den ITD- und PGI-Reichweiteabweichungen sind vorrangig auf steile Dichtegradien nahe des Bragg-Peaks und lateral zur Einstrahlrichtung zurückzuführen (Janssens et al., 2016; Priegnitz et al., 2015, 2016b; Schmid et al., 2015).

Die PGI-basierte Behandlungskontrolle ist momentan vor allem durch das eingeschränkte FOV und die statistische Unsicherheit von Spots mit kleinem Gewicht limitiert, welche die Gründe für das Filtern von Spots darstellen. Für die hypofraktionierten Prostata-Tumor-Bestrahlungen spielt dies jedoch eine untergeordnete Rolle, da das im Allgemeinen kompakte Zielvolumen gut durch das anpassbare FOV abgedeckt werden konnte, sodass im Median noch 96,2% des gesamten Spotgewichts für die PGI-Klassifikation zur Verfügung standen (vgl. Tabelle 5.1). Die Grundwahrheit hingegen wurde bewusst unter Einbeziehung aller Spots (ohne Filterung und ohne Spotaggregation) etabliert, um keine Dosisinformationen aufgrund von Einschränkungen des PGI-Systems für die Klassifikation auszuschließen, wie es für eine klinische Anwendung von PGI nötig ist. In einer zukünftigen Auswertung könnten jedoch auch die Auswirkungen der derzeitigen Einschränkungen des PGI-Systems in der Definition der Grundwahrheit berücksichtigt werden, um die prinzipielle Detektionsfähigkeit von PGI zu untersuchen. Dies wäre insbesondere für die Kopf-Hals-Tumor-Patienten interessant, da hier die Spotfilter eine größere Herausforderung darstellen: Die Bestrahlungsfelder sind in Strahlrichtung teilweise so weit ausgedehnt, dass das FOV nur einen bestimmten Teil des Bestrahlungsfeldes umfasst. Außerdem ist für die untersuchten Patienten die Dosis pro Feld deutlich geringer als im Fall der Prostata-Tumor-Bestrahlungen gewesen. Zuletzt ist auch noch ein weiterer Spotfilter zum Einsatz gekommen, um nur solche Spots auszuwerten, die vollständig im Patienten stoppen. Somit basierte die PGI-Klassifikation im Median auf nur 31,2% des gesamten Spotgewichts.

Um in der Bestimmung der PGI-Reichweiteabweichungen den Einfluss von systematischen Unsicherheiten der Reichweitevorhersage auf die Studie zu minimieren, wurde die DirectSPR-Methode in der Simulation der PGI-Signale auf den pCT-Scans verwendet (vgl. Kapitel 4). Weiterhin wurden für die Prostata-Tumor-Bestrahlungen heuristische Korrekturen auf die gemessenen PGI-Daten angewandt (vgl. Unterkapitel 3.5.4). Diese beeinflussen den Endpunkt der Studie jedoch nur geringfügig, da ansonsten während der Optimierung des PGI-Modells für systematische Fehler Rechnung getragen würde. Für Kopf-Hals-Tumor-Bestrahlungen wurde hingegen keine Korrektur angewandt, da in diesem Fall vorher noch weitere Untersuchungen in Phantomen notwendig gewesen wären.

Einordnung der etablierten PGI Klassifikationsmodelle

Wenngleich die Reichweiteabweichung mit PGI auf Spotebene vorliegt, so lässt sich doch anhand dieser Information nur schwer eine Aussage über relevante Behandlungsabweichungen treffen, da hierzu mindestens die Einordnung zum restlichen Bestrahlungsfeld notwendig ist. Hinzu kommt, dass PGI auf Spotebene eine hohe statistische Unsicherheit besitzt, welcher durch einen der beiden Ansätze zur feldweisen Klassifizierung anhand von Spots bzw. Clustern begegnet wurde. Auf Clusterebene konnte deutlich gezeigt werden, dass eine gute Korrelation zwischen den PGI und ITD-Reichweiteabweichungen besteht, was gleichzeitig eine wichtige Grundvoraussetzung für eine feldweise Klassifikation darstellt. Bei dieser

zeigte das clusterbasierte Modell jedoch keinen wesentlichen Unterschied zum spotbasierten Ansatz, da vermutlich auch beim spotbasierten Modell implizit das Spotgewicht und damit die statistische Unsicherheit der Spots in die Klassifizierung einfließen. Allerdings zeigt sich anhand des Clusterings für die Prostata-Tumor-Patienten, dass sich bereits ab *einem* relevanten Cluster (d. h. $T_{\text{rel,PGI,Cluster}} = 0\%$, vgl. Abbildung 5.8) ein sinnvolles Klassifikationsmodell ableiten lässt. Daher liegt die zukünftige Bedeutung von Clustern womöglich vor allem in der Visualisierung der PGI-Daten und ihrem Potential bei der Klassifizierung unvollständig abgestrahlter Bestrahlungsfelder für die Behandlungskontrolle in der OAPT.

In dieser Untersuchung wurde ein einfacher Ansatz zur feldweisen Beurteilung der Detektionsfähigkeit eines PGI-Systems etabliert. Es hat sich gezeigt, dass die PGI-Klassifizierungsansätze (auf Basis von Spots oder Clustern) für unterschiedliche Tumorentitäten – in diesem Fall Prostata- und Kopf-Hals-Tumore – einen sehr unterschiedlichen Schwellwert in der Optimierung ergeben haben. Dieses Verhalten zeigt, dass zukünftige PGI-Klassifikationsmodelle entitätsspezifisch trainiert werden sollten. Gründe dafür sind die unterschiedliche Definition der ITD-basierten Grundwahrheit (resultierend aus der Anpassung mit einer einheitlichen manuellen Klassifikation), der Anteil der durch PGI gefilterten Spots, die unterschiedliche Fraktionsdosis und die Aufteilung dieser Dosis auf die Bestrahlungsfelder. Außerdem spielen die unterschiedliche Eindringtiefe im Patienten und die Verwendung eines Range-Shifters als zusätzliche Quelle für den Untergrund des PGI-Signales eine Rolle. Ein entitätsspezifisches PGI-Klassifikationsmodell entspricht auch eher der Tatsache, dass Unterschiede in der robusten Optimierung und der Definition für die klinischen Sicherheitssäume zwischen den Tumorentitäten existieren. Darüber hinaus deuten die Ergebnisse dieser Arbeit darauf hin, dass der Klassifizierungsansatz (anhand des Parameters $\bar{W}_{\text{rel,PGI,Spot}}$ oder $\bar{W}_{\text{rel,PGI,Cluster}}$) für verschiedene Tumorentitäten durchaus einheitlich sein kann. Es bedarf bei den Kopf-Hals-Tumor-Patienten jedoch mehr Daten, um statistisch signifikante Schlussfolgerungen ziehen zu können.

Wie bei allen binären Klassifikationsmodellen, werden die zu Grunde liegenden Daten in zwei Klassen zusammengefasst, welche oftmals nicht die Vielschichtigkeit der Grenzfälle reflektieren. Den klinischen Anforderungen entsprechend könnten daher in zukünftigen Untersuchungen auch neue, ggf. mehrklassige Klassifikationsmodelle als Grundwahrheit etabliert werden, welche z. B. auch die Ursache der Abweichung berücksichtigen oder Abweichungen in bestimmten Geweberegionen betonen. Gleichzeitig lassen sich entsprechend im binären oder mehrklassigen Fall auch vielschichtige PGI-Klassifikationsmodelle entwickeln. Ein potentieller Ansatz zur Unterscheidung verschiedener Fehlerursachen greift hierbei auf die Etablierung von ausgewählten Merkmalen (im Engl. *features*) der Verteilung von PGI-Reichweiteabweichungen zurück (Khamfongkhrua et al., 2020). Eine potentielle Verbesserung der Klassifizierungsergebnisse könnte sich weiterhin durch die Einbeziehung von maschinellem Lernen durch merkmalsbasierte Modelle oder neuronale Netzwerke erreichen

lassen. Speziell neuronale Netzwerke extrahieren die relevanten Merkmale selbstständig aus den Daten und sind daher vielversprechend bei der Anwendung auf einer veränderten Definition der Grundwahrheit oder der Erschließung einer neuen Tumorentität. In einer ersten Studie konnte anhand von realistisch simulierten Kopf-Hals-Tumor-Patienten gezeigt werden, dass Faltungsnetzwerke, im Engl. *convolutional neural networks*, ein ähnliches Klassifikationsvermögen wie merkmalsbasierte Modelle erzielen (Pietsch et al., 2022). Zukünftige Klassifikationsmodelle führen jedoch nicht zur Revidierung der hier präsentierten Ergebnisse, sondern stellen eine generelle Möglichkeit zur Verbesserung der PGI-Klassifikation dar. Außerdem profitieren solche zukünftigen Modelle von den in dieser Arbeit geleisteten Vorarbeiten und dem Verständnis zur Grundwahrheit auf Grundlage einer manuellen und einer auf ITD-Profilen basierenden Klassifikation.

6 Genauigkeit zweier Reichweiteverifikationsmethoden – bizen-trischer Vergleich

Zur Charakterisierung von Reichweiteverifikationssystemen (Unterkapitel 2.3.2) haben verschiedene Arbeitsgruppen systematische Untersuchungen in Phantomen durchgeführt (z. B. Hueso-González et al., 2018; Nenoff et al., 2017). Um die Vergleichbarkeit von Messsystemen für die PG-basierte Reichweiteverifikation zu verbessern, wurde eine Kollaboration zwischen den sich mit PGI und PGS beschäftigenden Arbeitsgruppen am OncoRay bzw. Massachusetts General Hospital (MGH) ins Leben gerufen. Ziel war es, die Genauigkeit, Präzision und Robustheit der Messsysteme gegenüber der Absolutreichweite der Protonen und einer relativen Reichweiteverschiebung zu vergleichen, indem standardisierte PBS-Bestrahlungen eines anthropomorphen Kopfphantoms durchgeführt wurden. Zwei Aspekte sind dabei als essentiell hervorzuheben: (1) Die gemessenen Reichweiteabweichungen werden mit der auf einer exakten SPR-Kalibrierung des Kopfphantoms basierenden Grundwahrheit verglichen. (2) Die entworfenen Bestrahlungspläne basieren auf dem gleichen CT-Scan mit gleicher Dosis im Zielvolumen, Spotpositionen und Layerenergien und werden mit den klinischen Einstellungen der Bestrahlungsmaschine abgestrahlt. Dieses Vorgehen erlaubt eine standardisierte Einschätzung von unterschiedlichen PG-basierten Messsystemen, was eine wichtige Voraussetzung für die Kommissionierung und klinische Verbreitung dieser Systeme in der Protonentherapie darstellt. Die hier vorgestellten Ergebnisse wurden bereits auf Konferenzen vorgestellt (Berthold et al., 2022b; Hueso-González et al., 2022). Die Datenauswertung wurde im Rahmen der genannten Zusammenarbeit durchgeführt. Hierbei wurde die Auswertung der PGS-Daten seitens des MGH durchgeführt und für den Zweck des Vergleichs mit PGI der vorliegenden Arbeit zur Verfügung gestellt.

6.1 Material und Methoden

Der folgende Teil beschreibt den Aufbau und den Ablauf der Studie und gliedert sich in die Bildgebung und Bestrahlungsplanung für ein Kopfphantom, die experimentelle Durchführung sowie die anschließende Auswertung (Abbildung 6.1).

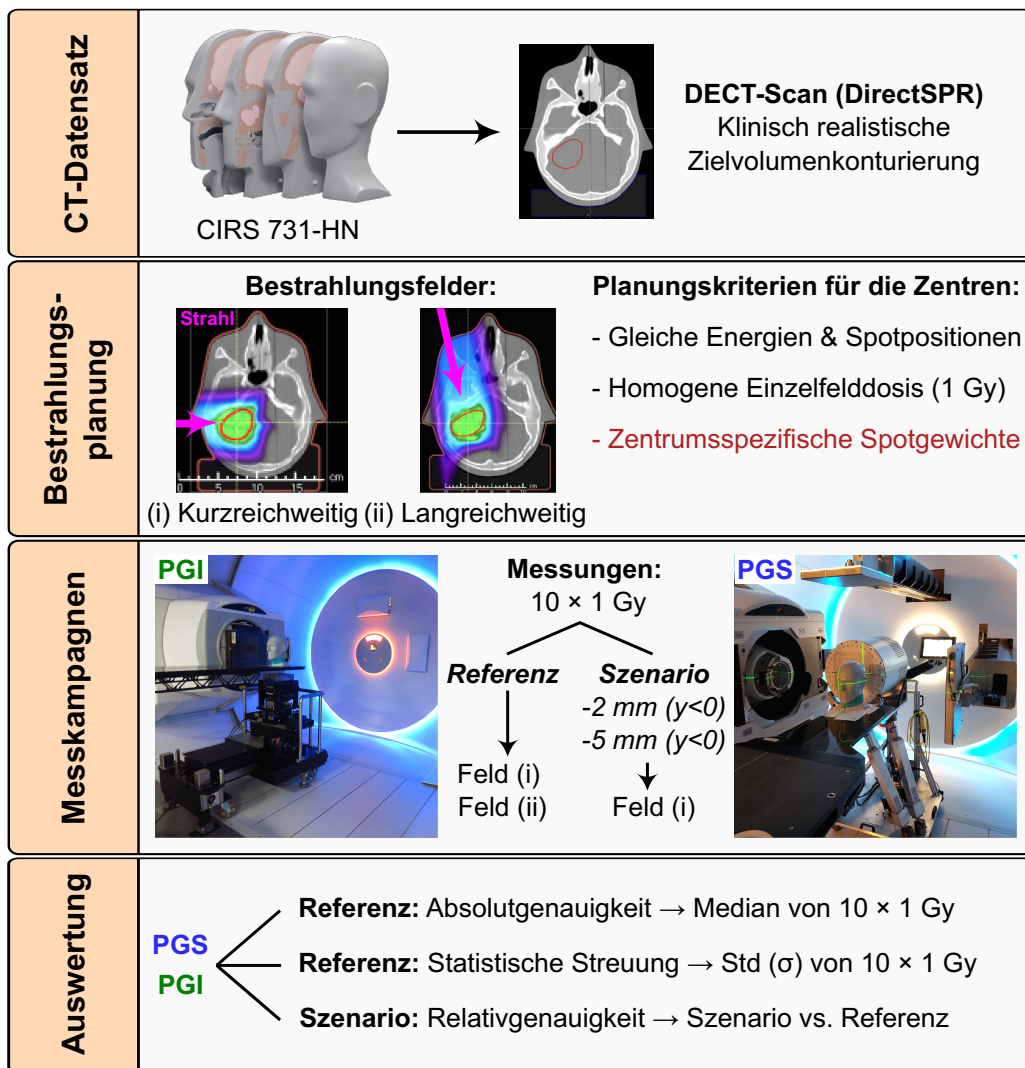


Abbildung 6.1: Design der bizen-trischen Studie für den Vergleich von PG-basierten Reichweiteverifikationssystemen anhand eines anthropomorphen Kopfphantoms.

6.1.1 Bildgebung

Das anthropomorphe Kopfphantom 731-HN des Herstellers CIRS (Norfolk, VA, USA) wurde für die Messungen zwischen den beiden Instituten ausgetauscht. Es wurde bereits zuvor grundlegend charakterisiert (Wohlfahrt et al., 2018). Dabei wurde jedem CT-Voxel eine spezifische Materialzusammensetzung, eine Elektronendichte und ein SPR-Wert zugeordnet – hier als SPR-Grundwahrheit bezeichnet. Um einen möglichen Bias zu vermeiden, wurde die SPR-Grundwahrheit jedoch bis zur vollständigen Auswertung der PGI- und PGS-Daten unter Verschluss gehalten und fließt nicht in die Bestrahlungsplanung oder die PG-basierten Arbeitsabläufe ein.

Stattdessen beruhen die Bestrahlungsplanung und die PG-basierten Arbeitsabläufe auf einem an der UPTD durchgeführten DECT-Scan mit klinischem Scanprotokoll unter der

Verwendung des DirectSPR-Verfahrens. Dieser wurde gegenüber der SPR-Grundwahrheit registriert. Der Vorteil dieses Ansatzes ist, dass der PG-basierte Arbeitsablauf für Patientenbestrahlungen (mit an den Instituten unterschiedlichen Programmen zur Bestrahlungsplanung und zur Simulation von PG) unabhängig von der SPR-Grundwahrheit validiert werden kann. Aufgrund der Einschränkung des PGI-Systems auf horizontale Bestrahlungsfelder wurde das Phantom in aufrechter Position gescannt, um reguläre Kopfbestrahlungen unter verschiedenen Gantrywinkeln nachzustellen.

6.1.2 Bestrahlungsplanung

Ein realistisches Zielgebiet in der dorsolateral rechten Region des Gehirns wurde im DECT-Scan konturiert (vgl. Wohlfahrt et al., 2018). Für die Bestrahlungsplanung wurden zwei Felder ausgewählt. Diese wurden bei horizontalem Gantrywinkel auf das aufrecht stehende und entsprechend um die (vertikale) z -Achse gedrehte Phantom appliziert:

- *Kurze Reichweite*: Laterale Einstrahlrichtung (entspricht 270° Gantrywinkel bei liegendem Phantom) mit Range-Shifter, bei welcher der Tumor auf direktem Weg erreicht wird. Dieses Feld könnte daher ein klinisches Bestrahlungsfeld darstellen.
- *Lange Reichweite*: Einstrahlrichtung von schräg vorn (entspricht 350° Gantrywinkel bei liegendem Phantom) ohne Range-Shifter, für welche viele Durchgänge durch Knochen und Hohlräume auf dem Weg zum Tumor liegen. Dieses Feld ist daher klinisch gesehen nicht realistisch, dient jedoch, aufgrund der genannten, stark inhomogenen Gewebzusammensetzung in Strahlrichtung und senkrecht dazu, als Stresstest für Methoden zur Reichweiteverifikation.

Die Spotgrößen an der UPTD sind wesentlich kleiner als an der genutzten Bestrahlungsanlage am MGH und erfordern deshalb Pläne mit kleinerem Spotabstand. Der Bestrahlungsplan wurde am MGH mit einem für die UPTD typischen Spotabstand von 6 mm erstellt. Zur Planung der Felder mit je 1 Gy homogener Einzelfelddosis wurde die Software Astroid (Kooy et al., 2010) verwendet. Sicherheitssäume um das CTV oder robuste Optimierung kamen nicht zum Einsatz. In beiden Feldern wurden 15 Energieschichten appliziert mit insgesamt 779 Spots beim kurzreichweitigen und 1117 beim langreichweitigen Feld. Der resultierende Bestrahlungsplan wurde im Bestrahlungsplanungssystem RayStation an der UPTD für die kleinere Spotgrößen nachoptimiert. Dabei wurden die gleichen Spotpositionen und Layerenergien verwendet und nur die Spotgewichte für die gleichen klinischen Ziele neu optimiert. Die unterschiedliche Spotgröße beeinflusst auch die Ausdehnung der lateralen Penumbra der Bestrahlungsfelder, sodass die Anzahl der applizierten Protonen an der UPTD bei vergleichbarer Dosisverteilung im CTV um rund 30% gegenüber des Bestrahlungsplans am MGH reduziert war.

6.1.3 Durchführung und Auswertung

Für jedes Bestrahlungsfeld wurde das jeweilige Messsystem so positioniert, dass sich mindestens ein für den klinischen Einsatz realistischer, 10 cm großer Luftspalt zwischen Kollimator und Phantom befand. Es wurde ein vergleichbarer Abstand von ca. 37 cm zwischen Isozentrum und Detektorarray für beide Messsysteme eingestellt. Das PGS-System wurde im Experiment dorsal zum Kopfphantom positioniert (Abbildung 6.1), wohingegen das PGI-System kaudal dazu positioniert wurde (Abbildung 6.2). Dennoch sollte die Fluenz der PG am Detektor in etwa gleich groß sein, da von einer ähnlich niedrigen Absorption von PG im Phantom ausgegangen werden kann. Als Teil der Unsicherheitsabschätzung wurde die Präzision der Kamera- und Phantompositionierung für mindestens zehn Positionierungen untersucht. Dafür wurden die an der Gantry befindlichen planaren Röntgensystemen genutzt.

Im folgenden Teil werden die durchgeführten Experimente beschrieben, mit denen die Fähigkeit der beiden PG-Systeme untersucht wurde, die Absolutreichweite der Protonen und Reichweiteabweichungen relativ zu einer bekannten Verschiebung zu detektieren. Für alle durchgeführten Experimente wurde die systematische Genauigkeit und statistische Streuung anhand von zehn Wiederholungsmessungen pro ausgewertetem Spot mit dem bereits in Unterkapitel 3.2.2 eingeführten Spotaggregationsradius bestimmt. Jede Gruppe wendete für ihr System entsprechende Spotfilter an; im Fall von PGI wurde das FOV auf den Bereich [-1; 5] cm auf der zentralen Strahlachse eingeschränkt (siehe Unterkapitel 3.5) und Spots mit einer aggregierten Protonenanzahl von weniger als $5 \cdot 10^7$ Protonen herausgefiltert.

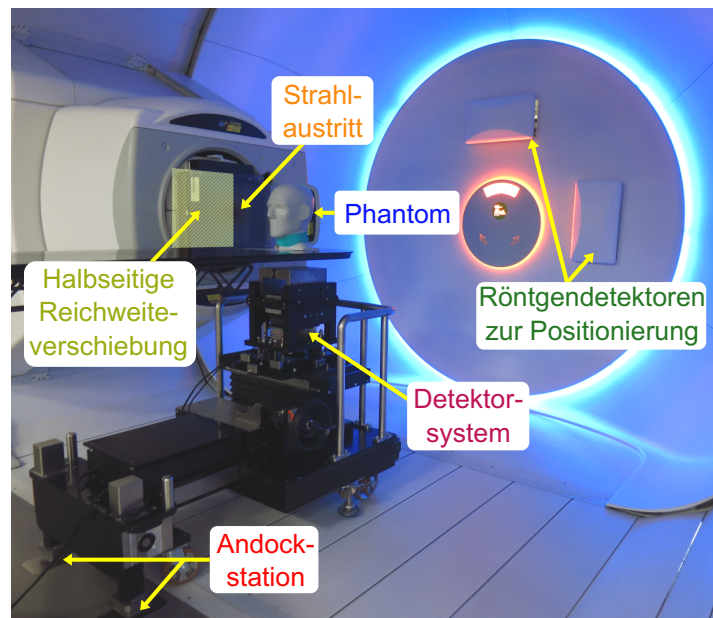


Abbildung 6.2: Experimenteller Aufbau des PGI-Systems in der UPTD für das kurzreichweitige Feld. Für die Untersuchung der relativen Genauigkeit des Detektionssystems wurde die Position der halbseitig im Bestrahlungsfeld liegenden Platten zur Reichweiteverschiebung angedeutet. Die Röntgendetektoren zur Positionierung des Phantoms befinden sich im eingefahrenen Zustand.

Auswertung zur Absolutreichweite

Zunächst wurden für die beiden Bestrahlungsfelder mit kurzer und langer Reichweite die spotweisen PGI- und PGS-Reichweiteabweichungen im Bezug auf die DirectSPR-Vorhersage ermittelt. Genau wie bei der Bestrahlungsplanung, wurde dabei die klinische Materialübersetzung verwendet, damit keine Informationen über die Materialien des Phantoms in die PG-Simulationen einfließen. Über die zehn durchgeführten Wiederholungsmessungen wurde für jeden Spot der Median (robuster gegen Ausreißer) und die Standardabweichung bestimmt.

Um die dosisbasierte Reichweiteabweichung zwischen DirectSPR und der SPR-Grundwahrheit zu ermitteln, wurde die spotweise Dosisverteilung der Bestrahlungspläne für den jeweiligen CT-Datensatz mit dem Monte-Carlo-Simulationscode gPMC (Jia et al., 2012; Qin et al., 2016) berechnet. Die ITD-Reichweiteabweichungen zwischen SPR-Grundwahrheit und dem DirectSPR-Datensatz wurden ähnlich wie in Unterkapitel 5.2.2 durch ein Anpassungsverfahren zwischen den ITD-Profilen bestimmt (Abbildung 6.3). Insgesamt zeigt sich, dass die Streuung der Abweichungen zwischen DirectSPR und der SPR-Grundwahrheit um ein Vielfaches geringer ist als die in den vorangegangenen Kapiteln beobachtete statistische Streuung der PGI-Reichweiteabweichungen in klinischen Bestrahlungen. Die Unterschiede zur Bestrahlungsanlage am MGH ergeben sich aus den verschiedenen Spotgrößen.

Als Maß für die Genauigkeit der Absolutreichweite des jeweiligen PG-Detektionssystems wird der Mittelwert über die Mediane der spotweise gemessenen Reichweiteabweichungen mit inverser Varianzgewichtung berechnet, wobei zuvor die ITD-Reichweiteabweichungen

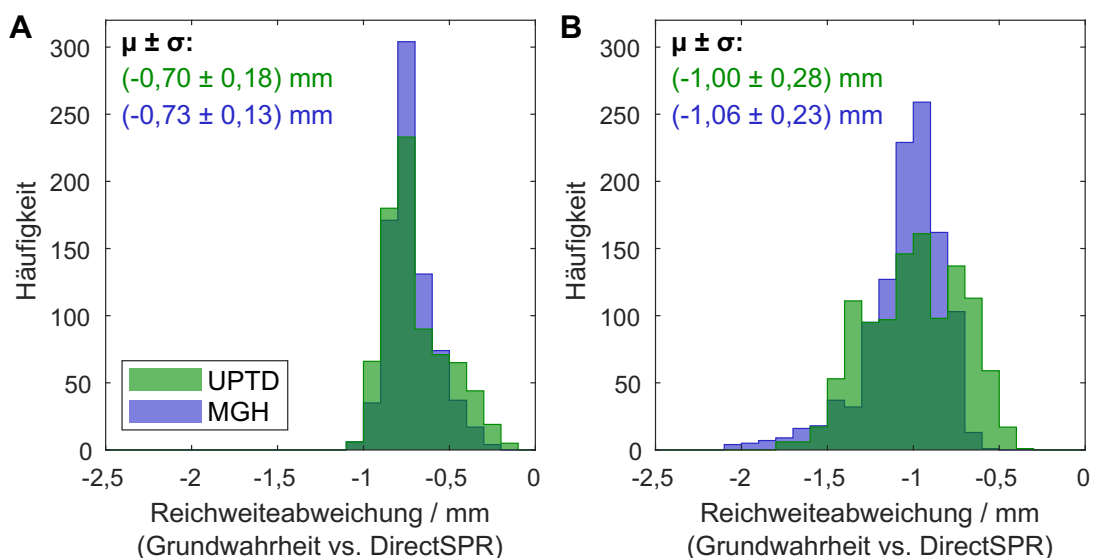


Abbildung 6.3: Histogramm der spotweisen Reichweiteabweichungen zwischen den auf der Grundwahrheit und DirectSPR basierenden ITD-Profilen für das kurzreichweitige (A) und langreichweitige (B) Bestrahlungsfeld.

zwischen DirectSPR und SPR-Grundwahrheit für jeden Spot abgezogen wurde. Als Maß für die Streuung wird die gewichtete Standardabweichung über die so korrigierten Mediane bestimmt und mittels Bootstrapping der Fehler des gewichteten Mittelwerts simuliert (vgl. Unterkapitel 4.2).

Auswertung zu relativen Reichweiteverschiebungen

Für das klinisch realistische kurzreichweitige Feld wurde zusätzlich die Detektierbarkeit von relativen Reichweiteverschiebungen untersucht. Dazu wurde in der ventralen Hälfte des Bestrahlungsplans ($y < 0$ mm) nacheinander zwei wasseräquivalente Platten (CIRS, PWDT-30-02 und PWDT-30-05) in den Dicken ($2,04 \pm 0,02$) mm bzw. ($4,98 \pm 0,02$) mm und einer quadratischen Fläche von 30×30 cm in den Strahlengang gebracht. Mit einer mehrschichtigen Ionisationskammer (Zebra, IBA Dosimetry GmbH, Schwarzenbruck, Deutschland) wurde die Reichweiteverschiebung (R_{95}) für die Energien 130, 180 und 215 MeV zu ($2,1 \pm 0,1$) mm bzw. ($5,0 \pm 0,4$) mm bestimmt. Genau genommen entspricht diese Verschiebung der Reichweiteabweichung in Wasser, die genaue Verschiebung für jeden Spot müsste simuliert werden. In erster Näherung sollte die Abweichung durch das höhere SPR der Hirnmasse etwa 4% geringer sein (Wohlfahrt et al., 2018), d. h. also ca. ($2,0 \pm 0,1$) mm und ($4,8 \pm 0,4$) mm.

Die Genauigkeit der PG-Systeme gegenüber relativen Veränderungen lässt sich aus dem Vergleich der Messung gegenüber einer Referenzmessung im Phantom bestimmen. In dieser Auswertung werden aber nicht zwei Messungen direkt miteinander verglichen, sondern indirekt über die Abweichung der jeweiligen Messung zu ein und derselben simulierten Referenz auf dem DirectSPR-Datensatz. Dadurch ist die PG-Simulation letztendlich nur ein Hilfsmittel, dass die Ergebnisse der relativen Genauigkeit nicht wesentlich beeinflusst.

Um die beiden Systeme quantitativ vergleichen zu können, wurde die mittlere Reichweiteabweichung über Spots mit und ohne Reichweiteverschiebung gebildet. Hierbei wurde ein Bereich proportional zur kleinsten Spotgröße (2σ) im Bestrahlungsplan um den Dichtegradienten ($y = 0$ mm) ausgelassen, um den Einfluss der Vermischung von unterschiedlichen Protonenreichweiten in einem Spot (aus dem Engl. *Range-Mixing*) zur reduzieren. Mit einem zweiseitigen Welch-Test ($\alpha = 0,05$) konnte überprüft werden, ob die gemessenen mittleren Reichweiteabweichungen sich im Bereich mit und ohne Reichweiteverschiebung signifikant unterscheiden.

6.2 Ergebnisse

Absolutreichweite

Die Gesamtunsicherheit zur Bestimmung der Genauigkeit der Absolutreichweite wurde wie bei der Validierung der Reichweitevorschagemethoden abgeschätzt (vgl. Unterkapitel 4.2).

Untersucht wurden dafür u. a. die Positionierungsgenauigkeit des Phantoms und der Detektionssysteme sowie weitere Unsicherheitsbeiträge, die nicht auf die PG-Systeme zurückzuführen und hier ebenso mit einem 2σ -Konfidenzniveau aufgeführt sind:

- Phantompositionierung aus Experiment: 0,4 mm (MGH) und 0,6 mm (UPTD),
- Kamerapositionierung: $0,4 \text{ mm}_{\text{stat}} + 1 \text{ mm}_{\text{sys}}$ PGS und $0,6 \text{ mm}_{\text{sys}}$ PGI (Unterkapitel 4.2),
- tägliche Variationen der wasseräquivalenten Protonenreichweite: 0,5 mm,
- Range-Shifter (falls eingesetzt): 0,5 mm,
- Genauigkeit der SPR-Grundwahrheit des Phantoms: 0,3%,
- Unsicherheit durch Materialsegmentierung in der SPR-Grundwahrheit: 1,0 mm,
- Reichweitengenauigkeit der Strahlmodelle im Monte-Carlo-Modell: 0,2 mm,
- Bildregistrierung zwischen SPR-Grundwahrheit und DirectSPR-Datensatz: 0,2 mm,
- Bestimmung der dosimetrischen Reichweiteabweichungen: 0,2 mm.

In Abbildung 6.4 ist die spotweise Genauigkeit der Absolutreichweite für das PGI- und PGS-System mit der jeweils abgeschätzten Gesamtunsicherheit gezeigt. Als gewichteter Mittelwert und Standardabweichung ergaben sich im kurzreichweitigen Feld ($-0,49 \pm 0,04$) mm bzw. 0,82 mm (σ) für das PGI-System und ($2,40 \pm 0,04$) mm bzw. 0,77 mm (σ) für das PGS-System. Im langreichweitigen Feld ergaben sich gewichteter Mittelwert und Standardabweichung als ($2,35 \pm 0,09$) mm bzw. 1,91 mm (σ) für das PGI-System und ($1,30 \pm 0,09$) mm bzw. 1,54 mm (σ) für das PGS-System. Damit stimmt die mit den beiden unterschiedlichen PG-Systemen gemessene Absolutreichweite bis auf wenige Millimeter mit der SPR-Grundwahrheit überein, wodurch sich das klinische Anwendungspotential zur Behandlungsverifikation in der Protonentherapie zeigt. Insbesondere erreicht das PGI-System für das klinisch realistische Feld eine bessere Genauigkeit als das PGS-System.

Wie erwartet nimmt die spotweise statistische Unsicherheit mit steigender Anzahl an Protonen ab, wobei die Unsicherheit beim PGS-System etwa um einen Faktor zwei geringer ist (Abbildung 6.5 A). Dies ist nicht zuletzt auf die um einen Faktor vier größere Anzahl an detektierten Ereignissen pro Proton in diesem Setup zurückzuführen (Abbildung 6.5 B). Als mittlere Standardabweichung der Absolutreichweite (über die zehn Wiederholungsmessungen) aller ausgewerteten Spots ergab sich für PGI ein Wert von 1,44 mm im kurzreichweitigen Feld und 1,61 mm im langreichweitigen Feld. Analog dazu ergaben sich für PGS die Werte 0,71 mm im kurzreichweitigen Feld und 1,02 mm im langreichweitigen Feld.

Während also für das kurzreichweitige Feld die mittlere spotweise Standardabweichung der Absolutreichweite sich zwischen den beiden Systemen um einen Faktor zwei unterscheidet (PGI: 1,44 mm, PGS: 0,71 mm), ist die Standardabweichung über die spotweisen Mediane aus zehn Messungen nahezu gleich für beide Systeme (PGI: 0,82 mm, PGS: 0,77 mm). Dies deutet darauf hin, dass systematische Unsicherheiten beider PG-Systeme einen wesentlichen Einfluss auf die ermittelte Streuung der Mediane haben.

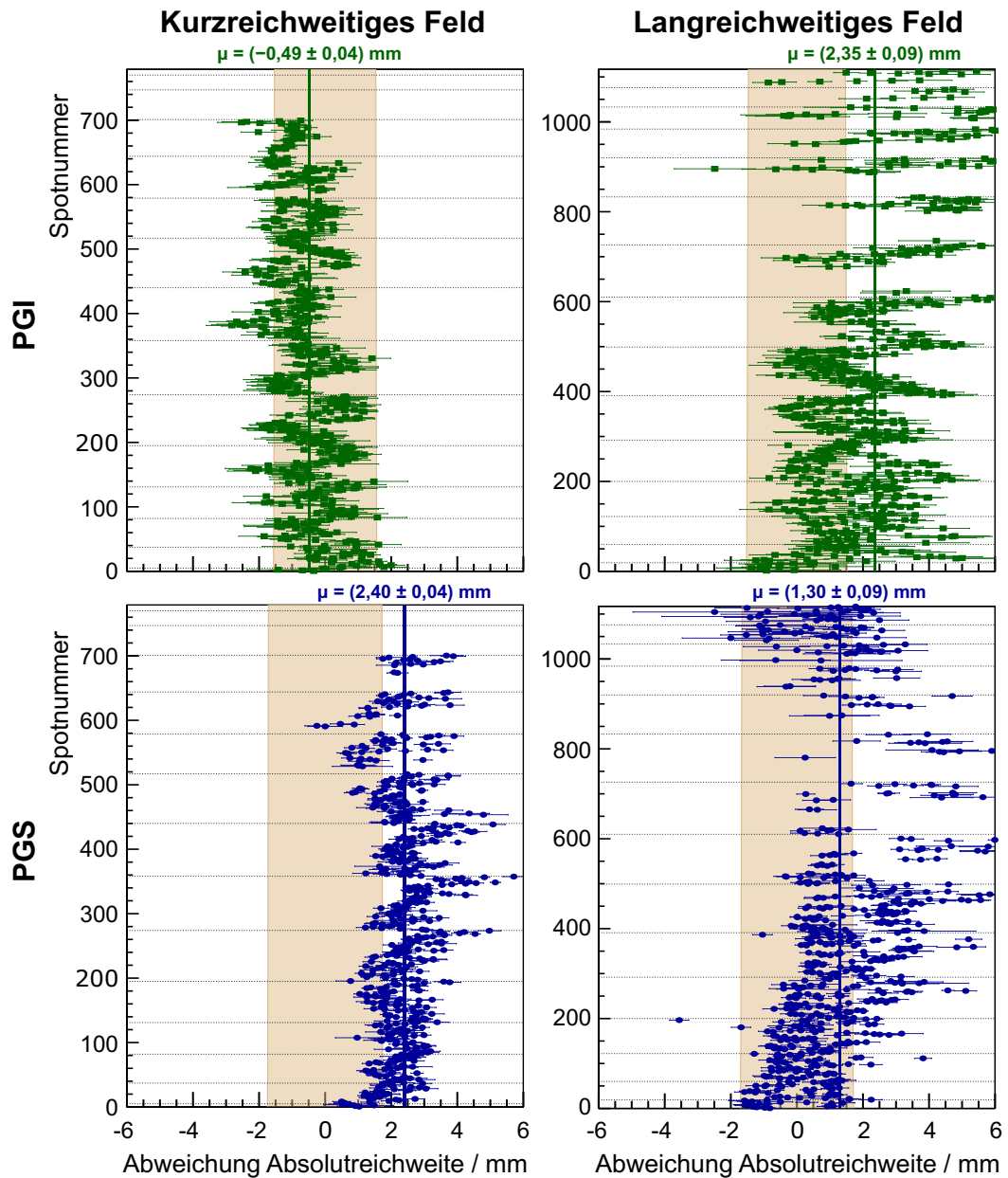


Abbildung 6.4: Spotweise Genauigkeit gegenüber der Absolutreichweite für das PGI-System (oben) und das PGS-System (unten), aufgeteilt in kurzreichweitiges (links) und langreichweitiges (rechts) Bestrahlungsfeld. Die Punkte entsprechen dem Median über zehn Wiederholungsmessungen und die Fehlerbalken dem Standardfehler des Median (nicht die Standardabweichung der Einzelmessung). Die vertikale Linie repräsentiert den gewichteten Mittelwert über alle dargestellten Spots und die horizontalen gestrichelten Linien trennen die Energieschichten voneinander. Der braune Bereich markiert das 2σ -Konfidenzintervall entsprechend der Gesamtunsicherheit des jeweiligen Systems.

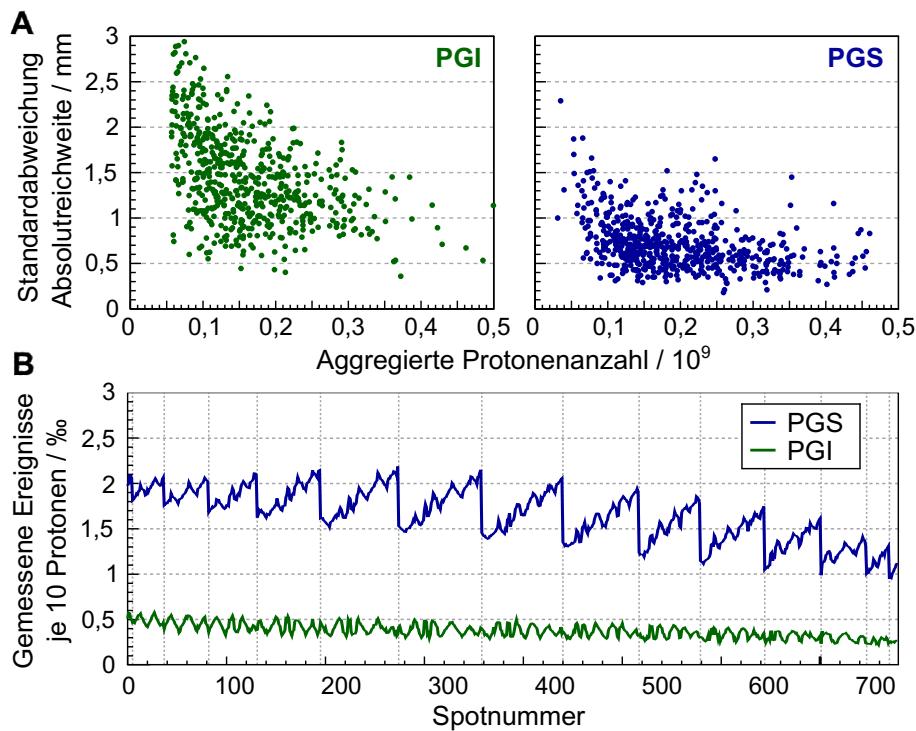


Abbildung 6.5: Spotweise Präzision (A) und detektierte Ereignisanzahl je zehn Protonen (B) – ca. jedes zehnte Proton emittiert PG – für das PGI- und PGS-System im kurzreichweitigen Feld. Die vertikal gestrichelten Linien im unteren Plotbereich (B) trennen die Energieschichten voneinander.

Relative Reichweiteverschiebungen

Der Einfluss eines steilen, lateralen Dichtegradienten auf die mit den PG-Systemen gemessenen Reichweiteabweichungen zeigt sich in Abbildung 6.6. Man erkennt deutlich die Abhängigkeit von der Spotgröße, die sich für die genutzten Bestrahlungsanlagen an der UPTD und am MGH um fast einen Faktor zwei unterscheiden: je größer die Spotgröße, desto größer das Range-Mixing und desto flacher der Verlauf von Reichweiteabweichung entlang von steilen, lateralen Dichtegradienten. Für beide PG-Systeme zeigt sich ein nahezu symmetrischer und stetiger Verlauf der Reichweiteabweichungen am Dichtegradienten, wie es für die Erkennung von relativen Reichweiteänderungen in Patientenbestrahlungen von Vorteil ist. Beim quantitativen Vergleich der mittleren Reichweiteabweichung im Bereich mit ($y \leq -2\sigma$) und ohne ($y \geq 2\sigma$) Reichweiteverschiebung (Tabelle 6.1) zeigte sich für alle Fälle ein signifikanter Unterschied. Somit können beide PG-Systeme Reichweiteabweichungen von etwa 2 mm und mehr erkennen. Hierbei wurden die im Wasser vorhergesagten Reichweiteverschiebungen im Mittel um bis zu 0,8 mm oder 16% unterschätzt. Diese Abweichungen können teilweise auf noch verbleibende Einflüsse des Range-Mixing und die um etwa 4% geringere erwartete Reichweiteverschiebung im Phantom zurückzuführen sein.

Tabelle 6.1: Überblick zur Detektierbarkeit relativer Reichweiteabweichungen mit einem Dichtegradienten entlang der y -Achse. Gezeigt ist der Mittelwert und dessen mittels Bootstrapping bestimmte Unsicherheit für jeden der in Abbildung 6.6 dargestellten Bereiche unter Berücksichtigung der Spotgröße (σ).

	2 mm Verschiebung		5 mm Verschiebung	
	$y \leq -2\sigma$	$y \geq 2\sigma$	$y \leq -2\sigma$	$y \geq 2\sigma$
PGI	$(-2,01 \pm 0,09)$ mm	$(-0,17 \pm 0,05)$ mm	$(-4,18 \pm 0,07)$ mm	$(-0,26 \pm 0,05)$ mm
PGS	$(-1,79 \pm 0,05)$ mm	$(-0,13 \pm 0,07)$ mm	$(-4,84 \pm 0,04)$ mm	$(0,42 \pm 0,09)$ mm

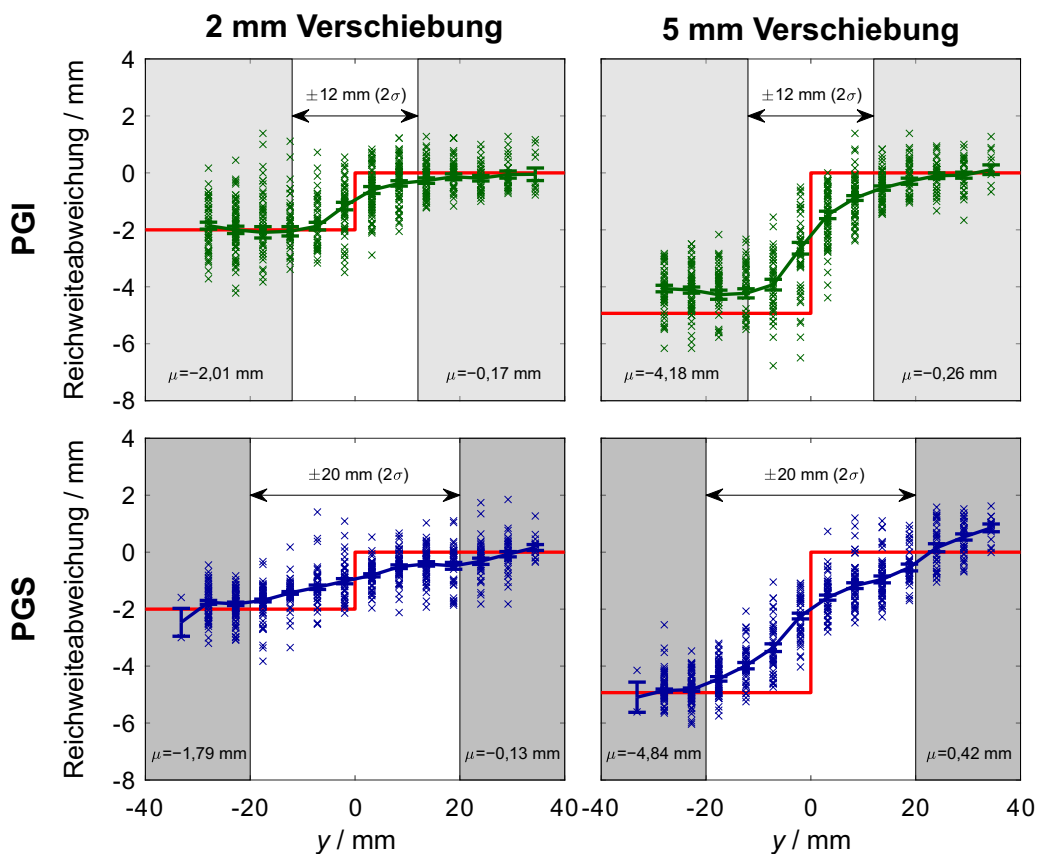


Abbildung 6.6: Detektierbarkeit relativer Reichweiteabweichungen. Jeder Datenpunkt entspricht einem Spot. Die rote Stufenfunktion zeigt die zu erwartende Reichweiteabweichung im Gehirn (2,0 mm und 4,8 mm für $y < 0$ mm) durch die in den Strahlengang eingebrachten Platten an. Die Fehlerbalken markieren den Mittelwert und dessen Unsicherheit für alle Spots mit der gleichen y -Koordinate; die Verbindungslinie zwischen den Mittelwerten dient der Visualisierung. Am unteren Rand der grauen Kästen ist der gewichtete Mittelwert aller Spots in diesem Bereich angegeben. Der Bereich um die Plattenkante wurde je nach anlagenspezifischer Spotgröße (σ) nicht in der Mittelwertbildung berücksichtigt.

6.3 Diskussion

In Rahmen dieser bizen-trischen Studie ist es erstmals gelungen, zwei bereits klinisch eingesetzte Prototypen für PG-basierte Reichweiteverifikation miteinander im Bezug auf die

Detektion der Absolutreichweite und hinsichtlich relativer Reichweiteabweichungen in einem anthropomorphen Kopfphantom zu vergleichen. In diesem Zusammenhang wurden Experimente für den standardisierten und von der Einrichtung unabhängigen Vergleich von Methoden zur Reichweiteverifikation etabliert.

Vergleich der Detektionssysteme

Die beiden untersuchten Prototypen der PGI- und PGS-Methode zeigten insgesamt eine hohe Genauigkeit für die Absolutreichweite mit einer systematischen Abweichung kleiner als 2,4 mm für beide Behandlungsfelder. Durch die zehnmahlige Wiederholung der Messungen konnte gleichzeitig die spotweise statistische Unsicherheit in Abhängigkeit von der Protonenanzahl ermittelt werden. Außerdem konnten für das kurzreichweitige Feld die erwarteten relativen Reichweiteabweichungen im Gehirn innerhalb von 0,6 mm bzw. 13% detektiert werden. Diese Ergebnisse unterstreichen das Potential der weltweit ersten klinisch eingesetzten Prototypen für PG-basierte Reichweiteverifikation mitsamt der dahinterstehenden Simulations- und Auswertungssoftware. Aufgrund der unterschiedlichen Spotgrößen an der UPTD und an der genutzten Bestrahlungsanlage am MGH ist ein direkter Vergleich der spotweisen Reichweiteabweichungen von PGI und PGS begrenzt, obwohl die gleichen Spotpositionen genutzt wurden. Dennoch lassen sich folgende Schlussfolgerungen ableiten: Das PGS-System besitzt unter den in dieser Studie vergleichbaren Bedingungen der Reichweiteverifikation eine geringe statistische Unsicherheit, wohingegen die Genauigkeit in der Bestimmung der Absolutreichweite durch das PGI-System für das klinisch realistische kurzreichweitige Feld höher ist. Die leicht bessere statistische Unsicherheit der PGS-Methode ist u. a. auf das geringere Detektionsvolumen des PGI-Systems (LYSO, 504 cm^3) gegenüber des PGS-Systems (Lanthanbromid (LaBr_3), 1236 cm^3) zurückzuführen. Die höhere Genauigkeit des PGI-Systems hingegen ist in Teilen auch dem genaueren Positionierungsverfahren zuzuschreiben (siehe Beitrag Gesamtunsicherheit), welches sich am Strahlisozentrum (Berthold et al., 2021) anstatt eines Referenzpunkts am Phantom oder Patienten (Hueso-González et al., 2018) orientiert. Bei der Erkennung von relativen Reichweiteverschiebungen sind beide Systeme in etwa gleich gut. Hervorzuheben ist auch, dass ein ähnlich großer Anteil an Spots oder Protonen herausgefiltert wurde, sodass das für die Auswertung nutzbare FOV unter den Bedingungen in dieser Studie gleichwertig ist. Schlussendlich konnte in dieser Studie gezeigt werden, dass beide Systeme für den klinischen Einsatz geeignet sind, wobei beide Arbeitsgruppen noch weitere Verbesserungen ihrer Systeme für möglich halten.

Herausforderungen und Bedeutung der Studie

Anhand des langreichweitigen Feldes zeigen sich die Herausforderungen für die Reichweiteverifikation in stark inhomogenem Gewebe (vgl. Janssens et al., 2016; Priegnitz et al., 2015, 2016b): Durch die lateralen Dichtegradienten kommt es u. a. zum Range-Mixing, sodass die abfallende Flanke des Bragg-Peaks verzerrt wird. Dies stellt sowohl für die Bestimmung der

Reichweiteabweichungen auf den ITD-Profilen aber auch für die PG-basierten Methoden eine Herausforderung dar. Außerdem könnten bereits kleine Abweichungen der simulierten von der applizierten Spotgröße große Auswirkungen auf die ermittelten Reichweiteabweichungen haben, was zur Erklärung der systematischen Abweichungen beider PG-Systeme für das langreichweitige Feld beiträgt. Ähnlich problematisch können sich auch nahe am Bragg-Peak gelegene Inhomogenitäten in Strahlrichtung auswirken.

Im Vergleich zu einer Messung am Patienten wird in dem verwendeten Kopfphantom durch die Protonenbestrahlung weniger PG induziert (Faktor $\approx 2/3$), was auf die unterschiedliche Materialzusammensetzung zurückzuführen ist. Im menschlichen Gewebe ist der Sauerstoffanteil größer und der Kohlenstoffanteil geringer als im Kopfphantom, während die PG-Emission doppelt so effektiv für ^{16}O als Targetkern wie für ^{12}C ist (Verburg & Seco, 2014). Folglich ist die hier gemessene statistische Unsicherheit im Phantom um etwa einen Faktor $\sqrt{1,5} \approx 1,2$ größer als im Patienten.

Anwendungsbereiche für Reichweiteverifikation in der Protonentherapie sind primär die Detektion von Behandlungsabweichungen, z. B. durch anatomische Veränderungen, wie im vorangegangenen Kapitel 5 untersucht. Aber auch die Validierung der SPR-Kalibrierung, wie in Kapitel 4 beschrieben, stellt eine bedeutende Anwendung dieser Systeme zur Verbesserung der Therapie dar. Dadurch könnten diese Technologien eine Schlüsselrolle bei der Entwicklung neuer Arbeitsabläufe, wie z. B. der OAPT, einnehmen. Der hier durchgeführte standardisierte Vergleich ermöglicht eine objektive Einordnung der Ergebnisse dieser Arbeit zu bestehenden und zukünftigen Messgeräten und -methoden für die Reichweiteverifikation.

7 Zusammenfassung

Die wachsende Anzahl an Protonentherapieanlagen weltweit eröffnet immer mehr Patienten die Möglichkeit, von einer präzisen und schonenden Strahlentherapie zu profitieren. Um die Bestrahlung noch weiter zu verbessern ist es notwendig, Unsicherheiten in der Bestrahlung weiter zu reduzieren und Arbeitsabläufe für die online-adaptive Protonentherapie (OAPT) zu entwickeln. Hierbei spielen Methoden zur Reichweiteverifikation eine bedeutende Rolle, indem sie dazu beitragen, Abweichungen in der Bestrahlung zu erkennen und somit als Trigger- oder Qualitätssicherungsinstrument für die Anpassung der Bestrahlungsplanung zu dienen.

In dieser Arbeit wurde erstmals systematisch das Potential eines Reichweiteverifikationssystems auf Grundlage der Prompt-Gamma-Bildgebung, engl. *Prompt-Gamma-Imaging* (PGI), unter klinischen Bestrahlungen von Prostata- und Kopf-Hals-Tumor-Patienten evaluiert. Folgende Beiträge wurden hierbei im Rahmen dieser Arbeit geleistet:

- die Kalibrierung und Charakterisierung des PGI-Systems an der Bestrahlungsanlage der Universitäts Protonen Therapie Dresden, die Durchführung von Patientenaufnahmen mittels PGI sowie regelmäßige Messungen zur Qualitätssicherung,
- die Bestimmung der Genauigkeit verschiedener Methoden zur Reichweitevorschau der Protonen im Patienten, darunter das DirectSPR-Verfahren (Peters et al., 2022), anhand von PGI-Messungen bei Prostata-Tumor-Bestrahlungen,
- die Quantifizierung des Klassifikationspotentials von relevanten und nicht relevanten anatomischen Veränderungen bei Bestrahlungsfeldern von Prostata- und Kopf-Hals-Tumor-Patienten anhand von aufgenommenen PGI-Daten,
- die erstmalige Durchführung eines institutsübergreifenden Vergleichs des untersuchten PGI-Systems mit einem System auf Basis der Prompt-Gamma-Spektroskopie (PGS) im Rahmen einer Kollaboration zwischen dem Massachusetts General Hospital (MGH) und OncoRay – Nationales Zentrum für Strahlenforschung in der Onkologie.

Die auf der Computertomographie (CT) basierende Reichweitevorschau in der Protonentherapie konnte zuletzt aufgrund des DirectSPR-Verfahrens wesentlich verbessert werden (Peters et al., 2022). Jedoch ist es erst mithilfe von PGI überhaupt gelungen, Reichweitevorschauverfahren unmittelbar im Patienten zu validieren, wodurch im Fall von DirectSPR

die umfassende Validierung der Methode im anthropomorphen Kopfphantom (Wohlfahrt et al., 2018) und biologischem Gewebe (Möhler et al., 2018) komplementiert wurde. Gleichzeitig konnte mittels pseudo-monoenergetischen CT-Scans die Genauigkeit von mehreren auf Hounsfield Lookup-Tabelle (HLUT) basierenden Verfahren bestimmt werden. Diese pseudo-monoenergetischen CT-Scans wurden ebenso wie DirectSPR aus den mittels zweier Röntgenspektren aufgenommenen CT-Scans (DECT) berechnet. Dadurch war es möglich, mehrere klinisch genutzte Methoden zur Reichweitevorschätzung in der Validierung gegenüberzustellen. Hierbei stellte sich heraus, dass das DirectSPR- und ein darauf basierendes HLUT-Verfahren innerhalb der Gesamtunsicherheit (2σ) im Mittel keine Abweichung zwischen vorhergesagter und mit PGI gemessener Protonenreichweite zeigten, wohingegen beim Standard-HLUT-Verfahren eine relevante Reichweiteabweichung von $(1,8 \pm 0,4)\%$ ermittelt wurde.

Relevante anatomische Veränderungen im Verlauf der Therapie zu erkennen ist eine aktuelle Herausforderung für die Protonentherapie und der Hintergrund für die Entwicklung einer OAPT. In dieser Studie konnte erstmals nachgewiesen werden, dass ein PGI-basiertes Messsystem interfractionelle anatomische Veränderungen detektieren kann: Relevante Behandlungsabweichungen pro Bestrahlungsfeld konnten bei Prostata-Tumor-Patienten durch ein auf PGI-Daten basierendes Klassifikationsmodell mit einer Sensitivität und Spezifität von 74% ($KI_{95\%}$: [58; 89]%) bzw. 79% ($KI_{95\%}$: [66; 92]%) erkannt werden. Damit übereinstimmend sind die ersten Ergebnisse für eine deutlich kleinere Kohorte von Kopf-Hals-Tumor-Patienten, deren systematische Auswertung im Rahmen dieser Arbeit bereits eine Errungenschaft darstellt. Somit konnte zum einen gezeigt werden, dass der entwickelte PGI-Klassifikationsansatz potentiell auf andere Tumorentitäten übertragbar ist, was eine Vereinfachung für die klinische Implementierung der PGI-Behandlungsverifikation bedeuten würde. Zum anderen zeigte sich, dass eine automatische Klassifikation der PGI-Daten grundsätzlich möglich ist, was eine wichtige Voraussetzung für den Einsatz des PGI-System im Kontext der OAPT darstellt.

Über die Auswertung der Patientendaten hinaus wurde in einer gemeinsamen Kollaboration mit dem MGH ein Vergleich zwischen dem dortigen PGS-System und dem in dieser Arbeit untersuchten PGI-System durchgeführt. Dafür wurde zunächst ein standardisiertes Studienprotokoll mit Experimenten an einem anthropomorphen Kopfphantom entwickelt. Als Prämisse für diese Experimente galt, dass diese an einer beliebigen Bestrahlungsanlage für Protonen durchgeführt werden können, um dadurch direkt mit dieser Studie vergleichbare Resultate zu erhalten. Im Ergebnis zeigten beide Detektionssysteme eine hohe Genauigkeit für die Detektion der Absolutreichweite ($\leq 2,4$ mm) und bei der Erkennung von relativen Reichweiteschiebungen ($\leq 13\%$). Im Vergleich zu den mehreren Millimeter großen Unsicherheiten aus der Reichweitevorschätzung und bei anatomischen Veränderungen kann geschlussfolgert werden, dass beide Systeme bereits jetzt für den klinischen Einsatz geeignet sind. Bei der

direkten Gegenüberstellung war die Genauigkeit der Absolutreichweite für das klinisch realistische Bestrahlungsfeld mit dem PGI-System besser als für das PGS-System (-0,5 mm vs. 2,4 mm). Dafür erreichte das PGS-System unter den hier ähnlichen experimentellen Bedingungen eine um den Faktor zwei bessere Präzision im Vergleich zum PGI-System. Diese Unterschiede stellen jedoch keine generellen Beschränkungen der Reichweiteverifikationsmethoden dar, denn eine Verbesserung der untersuchten Detektionssysteme ist prinzipiell möglich.

Zusammengefasst konnte in dieser Arbeit gezeigt werden, dass die Behandlungskontrolle mit dem untersuchten PGI-Reichweiteverifikationssystem deutlich zur Verbesserung der Protonentherapie beitragen kann. Zum einen ermöglicht PGI die systematische Validierung der CT-basierten Reichweitevorhersage im Patienten, woraufhin potentiell die Reichweitevorhersage oder die Reichweiteunsicherheiten bei der Bestrahlungsplanung angepasst werden könnten. Zum anderen kann PGI relevante Behandlungsabweichungen erkennen, hier anhand der Detektionsfähigkeit gegenüber anatomischen Veränderungen demonstriert, sodass PGI prinzipiell in der Lage ist – nicht zuletzt im Kontext einer OAPT – als Trigger- oder Qualitätssicherungsinstrument eingesetzt zu werden. Als weiterer Meilenstein für eine prospektive klinische Anwendung von PGI konnten in Zusammenarbeit mit dem MGH standardisierte Experimente zum Vergleich verschiedener Messsysteme und -methoden für die Reichweiteverifikation entwickelt und durchgeführt werden. Dadurch wurde gleichzeitig die objektive Einordnung der hier mit dem PGI-System ermittelten Ergebnisse gegenüber dem PGS-System und zukünftigen Systemen ermöglicht.

8 Summary

The worldwide growing number of proton therapy facilities enables more and more patients to benefit from a precise and sparing radiation therapy. In order to further improve the irradiation, it is necessary to reduce uncertainties in irradiation and develop workflows for an online-adaptive proton therapy (OAPT). Here, range verification methods could play an important role by helping to detect deviations in the irradiation and thus serving as a trigger or quality assurance tool for adapting the treatment plan. This work was the first to systematically evaluate the potential of a prompt-gamma-imaging (PGI) based range verification system during clinical irradiations of prostate and head and neck tumor patients. The following contributions were made in the course of this work:

- the calibration and characterization of the PGI system at the irradiation facility of the Universitäts Protonen Therapie Dresden, as well as performing patient acquisitions with PGI and regular quality assurance measurements,
- the quantification of the accuracy of various methods for proton range prediction in patients, including the DirectSPR method (Peters et al., 2022), based on PGI measurements during prostate cancer treatments,
- the quantification of the classification potential of relevant and non-relevant anatomical changes for irradiation fields of prostate and head and neck tumor patients based on acquired PGI data,
- the performance of a first-time inter-institutional comparison of the investigated PGI system with a system based on prompt-gamma-spectroscopy (PGS) as part of a collaboration between Massachusetts General Hospital (MGH) and OncoRay – National Center for Radiation Research in Oncology.

Computer tomography (CT) based range prediction in proton therapy has recently been substantially improved due to DirectSPR (Peters et al., 2022). However, it was only with the help of PGI that range prediction methods could be validated directly in patients at all, in the case of DirectSPR complementing the comprehensive validation of the method in an antropomorphic head phantom (Wohlfahrt et al., 2018) and biological tissue (Möhler et al., 2018). At the same time, the accuracy of several methods based on a Hounsfield lookup table (HLUT) could be determined on the pseudo-monoenergetic CT (monoCT) scans which,

like DirectSPR, were determined from dual-energy CT (DECT) scans. This enabled a comparison of several clinically used proton range prediction methods within the validation. It was found that within the total uncertainty (2σ), DirectSPR and a threonon based HLUT method showed on average no deviation between predicted and PGI-measured proton range, whereas a relevant range deviation of $(1.8 \pm 0.4)\%$ was determined for the standard HLUT method.

Detecting relevant anatomical changes during the course of therapy is currently a challenge for proton therapy and the main reason for the development of an OAPT. In this study, it was proven for the first time that a PGI-based measurement system can detect interfractional anatomical changes: Relevant deviations per irradiated treatment field could be detected in prostate tumor patients by a PGI-based classification model with a sensitivity and specificity of 74% ($KI_{95\%}$: [58; 89]%) and 79% ($KI_{95\%}$: [66; 92]%), respectively. Consistent with this are the first results for a much smaller cohort of head and neck tumor patients, the systematic evaluation of which is already an achievement in the context of this work. Thus, on the one hand, it could be shown that the developed PGI-classification approach is potentially transferable to other tumor entities, which would mean a simplification for the clinical implementation of PGI treatment verification. On the other hand, it was shown that automatic classification of PGI data is in principle possible, which is an important prerequisite for the use of a PGI system in the context of OAPT.

Beyond the analysis of the patient data, in a joint collaboration with the MGH a comparison was performed between their PGS system and the PGI system that was investigated in this work. At first, a standardized study protocol was developed with experiments based on an anthropomorphic head phantom. The premise for these experiments was that they could be performed at any proton facility, thereby providing results directly comparable to this study. In conclusion, both detection systems revealed high accuracy for absolute range detection (≤ 2.4 mm) and relative range deviation detection ($\leq 13\%$). Compared to the several millimeters of uncertainties from range prediction and anatomical changes, this points out that both systems are already suitable for clinical use. In direct comparison, the accuracy of the absolute range for the clinically realistic treatment field was better for the PGI system than for the PGS system (-0.5 mm vs. 2.4 mm). In return, the PGS system showed a factor of two better precision compared to the PGI system under the here similar experimental conditions. However, these differences do not represent a general limitation of the range verification methods because improvements of the investigated detection systems are in principle possible.

In summary, this work was able to show that treatment monitoring with the investigated PGI range verification system can distinctly contribute to the improvement of proton therapy. On the one hand, PGI enables a systematic validation of CT-based range prediction in patients, upon which range prediction or range uncertainties in treatment planning could potentially be adjusted. On the other hand, PGI can detect relevant treatment deviations, shown here

by the detection capability of anatomical changes, so that PGI is in principle capable – not least in the context of OAPT – of being used as a trigger or quality assurance tool. As a further milestone for a prospective clinical application of PGI, standardized experiments for the comparison of different measuring systems and methods for range verification could be developed and performed in cooperation with MGH. At the same time this allows the objective comparison of the here obtained results for the PGI system with the PGS system and future systems.

Literatur

- Albertini, F., Matter, M., Nenoff, L., Zhang, Y., & Lomax, A. (2020). Online daily adaptive proton therapy. *British Journal of Radiology*, 93(1107), 20190594. <https://doi.org/10.1259/bjr.20190594>
- Assmann, W., Kellnberger, S., Reinhardt, S., Lehrack, S., Edlich, A., Thirolf, P. G., Moser, M., Dollinger, G., Omar, M., Ntziachristos, V., & Parodi, K. (2015). Ionoacoustic characterization of the proton Bragg peak with submillimeter accuracy. *Medical Physics*, 42(2), 567–574. <https://doi.org/10.1118/1.4905047>
- Assmann, W., & Parodi, K. (2020). The Sound of Ions: Acoustic Detection of High-Energy Beams. *Nuclear Physics News*, 30(1), 20–24. <https://doi.org/10.1080/10619127.2020.1717269>
- Bär, E., Lalonde, A., Zhang, R., Jee, K.-W., Yang, K., Sharp, G., Liu, B., Royle, G., Bouchard, H., & Lu, H.-M. (2018). Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples. *Medical Physics*, 45(1), 48–59. <https://doi.org/https://doi.org/10.1002/mp.12666>
- Baumann, M., Krause, M., Overgaard, J., Debus, J., Bentzen, S. M., Daartz, J., Richter, C., Zips, D., & Bortfeld, T. (2016). Radiation oncology in the era of precision medicine. *Nature Reviews Cancer*, 16(4), 234–249. <https://doi.org/10.1038/NRC.2016.18>
- Berger, M., Inokuti, M., Andersen, H., Bichsel, H., Powers, D., Seltzer, S., Thwaites, D., & Watt, D. (1993). ICRU Report 49: Stopping Powers and Ranges for Protons and Alpha Particles. *Journal of the International Commission on Radiation Units and Measurements*, os25(2). <https://journals.sagepub.com/toc/crub/os-25/2>
- Berthold, J., Hübinger, L., Piplack, N., Pietsch, J., Khamfongkhrua, C., Thiele, J., Appold, S., Traneus, E., Janssens, G., Smeets, J., Stützer, K., & Richter, C. (2022a). OC-0945 Treatment verification with prompt-gamma imaging: Detection sensitivity of anatomical changes in HNC. *Radiotherapy and Oncology*, 170, S837–S838. [https://doi.org/10.1016/s0167-8140\(22\)02725-6](https://doi.org/10.1016/s0167-8140(22)02725-6)
- Berthold, J., Hueso-González, F., Wohlfahrt, P., Bortfeld, T., Khamfongkhrua, C., Tattenberg, S., Zarifi, M., Richter, C., & Verburg, J. M. (2022b, 27. Juni–2. Juli). *Multi-institutional benchmark experiment of prompt-gamma imaging and spectroscopy for treatment verification using an anthropomorphic head phantom* [Konferenzvortrag]. 60th Annual Conference of the Particle Therapy Cooperative Group, Miami, FL, U.S.

- Berthold, J., Khamfongkhrua, C., Petzoldt, J., Thiele, J., Hölscher, T., Wohlfahrt, P., Peters, N., Jost, A., Hofmann, C., Janssens, G., Smeets, J., & Richter, C. (2021). First-In-Human Validation of CT-Based Proton Range Prediction Using Prompt Gamma Imaging in Prostate Cancer Treatments. *International Journal of Radiation Oncology Biology Physics*, *111*(4), 1033–1043. <https://doi.org/10.1016/J.IJROBP.2021.06.036>
- Berthold, J., Piplack, N., Traneus, E., Pietsch, J., Khamfongkhrua, C., Thiele, J., Hölscher, T., Janssens, G., Smeets, J., Stützer, K., & Richter, C. (2023). Proceedings to the 59th Annual Conference of the Particle Therapy Cooperative Group (PTCOG59 2021 Online): O068 - Treatment verification with prompt-gamma-imaging: Detection of anatomical changes in prostate-cancer proton therapy. *International Journal of Particle Therapy*, *9*(3), 58–242. <https://doi.org/10.14338/IJPT-22-PTCOG59-9.3>
- Bethe, H. (1930). Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Annalen der Physik*, *397*(3), 325–400. <https://doi.org/10.1002/andp.19303970303>
- Bom, V., Joulaeizadeh, L., & Beekman, F. (2012). Real-time prompt gamma monitoring in spot-scanning proton therapy using imaging through a knife-edge-shaped slit. *Physics in Medicine & Biology*, *57*(2), 297–308. <https://doi.org/10.1088/0031-9155/57/2/297>
- Bray, F., Laversanne, M., Weiderpass, E., & Soerjomataram, I. (2021). The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer*, *127*(16), 3029–3030. <https://doi.org/10.1002/cncr.33587>
- Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*, 6. <https://doi.org/10.1186/S12864-019-6413-7>
- De Ruyscher, D., Sterpin, E., Haustermans, K., & Depuydt, T. (2015). Tumour movement in proton therapy: Solutions and remaining questions: A review. *Cancers*, *7*(3), 1143–1153. <https://www.mdpi.com/2072-6694/7/3/829>
- Dedes, G., Dickmann, J., Niepel, K., Wesp, P., Johnson, R. P., Pankuch, M., Bashkirov, V., Rit, S., Volz, L., Schulte, R. W., Landry, G., & Parodi, K. (2019). Experimental comparison of proton CT and dual energy x-ray CT for relative stopping power estimation in proton therapy. *Physics in Medicine & Biology*, *64*(16), 165002. <https://doi.org/10.1088/1361-6560/ab2b72>
- DeLuca, P. M. (2007). ICRU Report 78: Prescribing, Recording, and Reporting Proton-Beam Therapy. *Journal of the International Commission on Radiation Units and Measurements*, *7*(2). <https://journals.sagepub.com/toc/crua/7/2>
- Echner, G. (2018). Patientenlagerung und -positionierung. In W. Schlegel, C. P. Karger & O. Jäkel (Hrsg.), *Medizinische Physik: Grundlagen – Bildgebung – Therapie – Technik* (S. 515–524). Springer Spektrum. https://doi.org/10.1007/978-3-662-54801-1_25

- Enghardt, W., Parodi, K., Crespo, P., Fiedler, F., Pawelke, J., & Pönisch, F. (2004). Dose quantification from in-beam positron emission tomography. *Radiotherapy and Oncology*, 73(S2), S96–S98. [https://doi.org/10.1016/S0167-8140\(04\)80024-0](https://doi.org/10.1016/S0167-8140(04)80024-0)
- Eulitz, J., Troost, E. G. C., Klünder, L., Raschke, F., Hahn, C., Schulz, E., Seidlitz, A., Thiem, J., Karpowitz, C., Hahlbohm, P., Grey, A., Engelland, K., Löck, S., Krause, M., & Lühr, A. (2023). Increased relative biological effectiveness and periventricular radiosensitivity in proton therapy of glioma patients. *Radiotherapy and Oncology*, 178, 109422. <https://doi.org/10.1016/j.radonc.2022.11.011>
- Everett, D. B., Fleming, J. S., Todd, R. W., & Nightingale, J. M. (1977). Gamma-radiation imaging system based on the Compton effect. *Proceedings of the Institution of Electrical Engineers*, 124(11), 995–1000. <https://doi.org/10.1049/piee.1977.0203>
- Fani, F., Schena, E., Saccomandi, P., & Silvestri, S. (2014). CT-based thermometry: An overview. *International Journal of Hyperthermia*, 30(4), 219–227. <https://doi.org/10.3109/02656736.2014.922221>
- Ferrero, V., Fiorina, E., Morrocchi, M., Pennazio, F., Baroni, G., Battistoni, G., Belcari, N., Camarlinghi, N., Ciocca, M., Guerra, A. D., Donetti, M., Giordanengo, S., Giraudo, G., Patera, V., Peroni, C., Rivetti, A., Rolo, M. D. D. R., Rossi, S., Rosso, V., ... Bisogni, M. G. (2018). Online proton therapy monitoring: clinical test of a Silicon-photodetector-based in-beam PET. *Scientific Reports*, 8(1), 4100. <https://doi.org/10.1038/s41598-018-22325-6>
- Fiorina, E., Ferrero, V., Baroni, G., Battistoni, G., Belcari, N., Camarlinghi, N., Cerello, P., Ciocca, M., Simoni, M. D., Donetti, M., Dong, Y., Embriaco, A., Fischetti, M., Franciosini, G., Giraudo, G., Kraan, A., Laruina, F., Luongo, C., Maestri, D., ... Bisogni, M. G. (2021). Detection of Interfractional Morphological Changes in Proton Therapy: A Simulation and In Vivo Study With the INSIDE In-Beam PET. *Frontiers in Physics*, 8, 660. <https://doi.org/10.3389/fphy.2020.578388>
- Fracchiolla, F., Engwall, E., Janson, M., Tamm, F., Lorentini, S., Fellin, F., Bertolini, M., Algranati, C., Righetto, R., Farace, P., Amichetti, M., & Schwarz, M. (2021). Clinical validation of a GPU-based Monte Carlo dose engine of a commercial treatment planning system for pencil beam scanning proton therapy. *Physica Medica*, 88, 226–234. <https://doi.org/10.1016/j.ejmp.2021.07.012>
- Golnik, C., Hueso-González, F., Müller, A., Dendooven, P., Enghardt, W., Fiedler, F., Kormoll, T., Roemer, K., Petzoldt, J., Wagner, A., & Pausch, G. (2014). Range assessment in particle therapy based on prompt γ -ray timing measurements. *Physics in Medicine & Biology*, 59(18), 5399–5422. <https://doi.org/10.1088/0031-9155/59/18/5399>
- Hayakawa, Y., Tada, J., Arai, N., Hosono, K., Sato, M., Wagai, T., Tsuji, H., & Tsujii, H. (1995). Acoustic pulse generated in a patient during treatment by pulsed proton radiation beam. *Radiation Oncology Investigations*, 3(1), 42–45. <https://doi.org/10.1002/roi.2970030107>

- Hubbell, J. H., & Seltzer, S. M. (2004). *Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients* (version 1.4). National Institute of Standards and Technology, Gaithersburg, MD. Verfügbar 7. Februar 2022 unter <http://physics.nist.gov/xaamdi>
- Hübinger, L. (2021, Oktober). *Range verification of proton therapy for head and neck cancer patients with prompt-gamma imaging: sensitivity of the detection of anatomical changes* [Masterarbeit]. Technische Universität Dresden.
- Hueso-González, F., Berthold, J., Wohlfahrt, P., Bortfeld, T., Khamfongkhrua, C., Tattenberg, S., Zarifi, M., Verburg, J. M., & Richter, C. (2022, 5.–12. November). *Experimental comparison of Prompt Gamma-Ray Imaging and Spectroscopy under Equalized Conditions Using a Calibrated Head Phantom* [Konferenzvortrag]. 2022 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room Temperature Semiconductor Detector Conference, Milano, Italia.
- Hueso-González, F., Enghardt, W., Fiedler, F., Golnik, C., Janssens, G., Petzoldt, J., Prieels, D., Priegnitz, M., Römer, K. E., Smeets, J., Stappen, F. V., Wagner, A., & Pausch, G. (2015). First test of the prompt gamma ray timing method with heterogeneous targets at a clinical proton therapy facility. *Physics in Medicine & Biology*, *60*(16), 6247–6272. <https://doi.org/10.1088/0031-9155/60/16/6247>
- Hueso-González, F., Rabe, M., Ruggieri, T. A., Bortfeld, T., & Verburg, J. M. (2018). A full-scale clinical prototype for proton range verification using prompt gamma-ray spectroscopy. *Physics in Medicine & Biology*, *63*(18), 185019. <https://doi.org/10.1088/1361-6560/aad513>
- Hünemohr, N., Krauss, B., Tremmel, C., Ackermann, B., Jäkel, O., & Greilich, S. (2013). Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates. *Physics in Medicine & Biology*, *59*(1), 83–96. <https://doi.org/10.1088/0031-9155/59/1/83>
- Jäkel, O. (2018a). Bildgebung für die Strahlentherapie. In W. Schlegel, C. P. Karger & O. Jäkel (Hrsg.), *Medizinische Physik: Grundlagen – Bildgebung – Therapie – Technik* (S. 473–484). Springer Spektrum. https://doi.org/10.1007/978-3-662-54801-1_23
- Jäkel, O. (2018b). Qualität und Sicherheit in der Strahlentherapie. In W. Schlegel, C. P. Karger & O. Jäkel (Hrsg.), *Medizinische Physik: Grundlagen – Bildgebung – Therapie – Technik* (S. 623–633). Springer Spektrum. https://doi.org/10.1007/978-3-662-54801-1_28
- Jäkel, O., Jacob, C., Schardt, D., Karger, C. P., & Hartmann, G. H. (2001). Relation between carbon ion ranges and x-ray CT numbers. *Medical Physics*, *28*(4), 701–703. <https://doi.org/10.1118/1.1357455>
- Janssens, G., Smeets, J., Vander Stappen, F., Prieels, D., Clementel, E., Hotoiu, E. L., & Sterpin, E. (2016). Sensitivity study of prompt gamma imaging of scanned beam

- proton therapy in heterogeneous anatomies. *Radiotherapy and Oncology*, 118(3), 562–567. <https://doi.org/10.1016/J.RADONC.2015.11.002>
- Jia, X., Schümann, J., Paganetti, H., & Jiang, S. B. (2012). GPU-based fast Monte Carlo dose calculation for proton therapy. *Physics in Medicine and Biology*, 57(23), 7783–7797. <https://doi.org/10.1088/0031-9155/57/23/7783>
- Johnson, R. P. (2017). Review of medical radiography and tomography with proton beams. *Reports on Progress in Physics*, 81(1), 16701. <https://doi.org/10.1088/1361-6633/aa8b1d>
- Joint Committee for Guides in Metrology / Working Group 1. (2008). *JCGM 100:2008 GUM 1995 with minor corrections: Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. Joint Committee for Guides in Metrology. https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf
- Jones, K. C., Nie, W., Chu, J. C., Turian, J. V., Kassae, A., Sehgal, C. M., & Avery, S. (2018). Acoustic-based proton range verification in heterogeneous tissue: Simulation studies. *Physics in Medicine & Biology*, 63(2), 025018. <https://doi.org/10.1088/1361-6560/aa9d16>
- Jongen, Y., & Stichelbaut, F. (2003, 26.–29. Oktober). *Verification of the proton beams position in the patient by the detection of prompt gamma rays emission* [Konferenzvortrag Abstrakt]. 39th Meeting of the Particle Therapy Co-Operative Group, San Francisco, CA, U.S. https://www.ptcog.ch/archive/steering_committee_docs/ptcog39-2003/2003_PTCOG_39_Abstracts.pdf
- Kabuki, S., Ueno, K., Kurosawa, S., Iwaki, S., Kubo, H., Miuchi, K., Fujii, Y., Kim, D., Kim, J., Kohara, R., Miyazaki, O., Sakae, T., Shirahata, T., Takayanagi, T., Terunuma, T., Tsukahara, Y., Yamamoto, E., Yasuoka, K., & Tanimori, T. (2009). Study on the use of electron-tracking compton gamma-ray camera to monitor the therapeutic proton dose distribution in real time. *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2437–2440. <https://doi.org/10.1109/NSSMIC.2009.5402130>
- Karger, C. P. (2018). Klinische Strahlenbiologie. In W. Schlegel, C. P. Karger & O. Jäkel (Hrsg.), *Medizinische Physik: Grundlagen – Bildgebung – Therapie – Technik* (S. 451–472). Springer Spektrum. https://doi.org/10.1007/978-3-662-54801-1_22
- Kellnberger, S., Assmann, W., Lehrack, S., Reinhardt, S., Thirolf, P., Queirós, D., Sergiadis, G., Dollinger, G., Parodi, K., & Ntziachristos, V. (2016). Ionoacoustic tomography of the proton Bragg peak in combination with ultrasound and optoacoustic imaging. *Scientific Reports*, 6, 29305. <https://doi.org/10.1038/srep29305>
- Khamfongkhrua, C. (2021, Februar). *Automatic classification of treatment-deviation sources in proton therapy using prompt-gamma-imaging information* [Disseration, Technische Universität Dresden]. Qucosa. <https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-760581>

- Khamfongkhruea, C., Berthold, J., Janssens, G., Petzoldt, J., Smeets, J., Pausch, G., & Richter, C. (2020). Classification of the source of treatment deviation in proton therapy using prompt-gamma imaging information. *Medical Physics*, *47*(10), 5102–5111. <https://doi.org/10.1002/MP.14393>
- Kim, D., Yim, H., & Kim, J. W. (2009). Pinhole camera measurements of prompt gamma-rays for detection of beam range variation in proton therapy. *Journal of the Korean Physical Society*, *55*(4), 1673–1676. <https://doi.org/10.3938/jkps.55.1673>
- Knopf, A.-C., & Lomax, A. (2013). In vivo proton range verification: a review. *Physics in Medicine & Biology*, *58*(15), R131–R160. <https://doi.org/10.1088/0031-9155/58/15/r131>
- Kooy, H. M., Clasio, B. M., Lu, H. M., Madden, T. M., Bentefour, H., Depauw, N., Adams, J. A., Trofimov, A. V., Demaret, D., Delaney, T. F., & Flanz, J. B. (2010). A Case Study in Proton Pencil-Beam Scanning Delivery. *International Journal of Radiation Oncology Biology Physics*, *76*(2), 624–630. <https://doi.org/10.1016/J.IJROBP.2009.06.065>
- Kraan, A. C., Berti, A., Retico, A., Baroni, G., Battistoni, G., Belcari, N., Cerello, P., Ciocca, M., Simoni, M. D., Sarto, D. D., Donetti, M., Dong, Y., Embriaco, A., Ferrero, V., Fiorina, E., Fischetti, M., Franciosini, G., Giraud, G., Laruina, F., ... Bisogni, M. G. (2022). Localization of anatomical changes in patients during proton therapy with in-beam PET monitoring: A voxel-based morphometry approach exploiting Monte Carlo simulations. *Medical Physics*, *49*(1), 23–40. <https://doi.org/10.1002/mp.15336>
- Krieger, H. (2019). Wechselwirkungen geladener Teilchen. In *Grundlagen der Strahlungsphysik und des Strahlenschutzes* (S. 257–305). Springer Spektrum. https://doi.org/10.1007/978-3-662-60584-4_9
- Krimmer, J., Angellier, G., Balleyguier, L., Dauvergne, D., Freud, N., Hérault, J., Létang, J. M., Mathez, H., Pinto, M., Testa, E., & Zoccarato, Y. (2017). A cost-effective monitoring technique in particle therapy via uncollimated prompt gamma peak integration. *Applied Physics Letters*, *110*(15), 154102. <https://doi.org/10.1063/1.4980103>
- Krimmer, J., Dauvergne, D., Létang, J. M., & Testa, E. (2018). Prompt-gamma monitoring in hadrontherapy: A review. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *878*, 58–73. <https://doi.org/10.1016/J.NIMA.2017.07.063>
- Landberg, T., Chavaudra, J., Dobbs, J., Gerard, J.-P., Hanks, G., Horiot, J.-C., Johansson, K.-A., Möller, T., Purdy, J., Suntharalingam, N., & Svensson, H. (1999). ICRU Report 62: Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50). *Journal of the International Commission on Radiation Units and Measurements*, *os32*(1). <https://journals.sagepub.com/toc/crub/os-32/1>
- Landberg, T., Chavaudra, J., Dobbs, J., Hanks, G., Johansson, K.-A., Möller, T., & Purdy, J. (1993). ICRU Report 50: Prescribing, Recording and Reporting Photon Beam

- Therapy. *Journal of the International Commission on Radiation Units and Measurements*, os26(1). <https://journals.sagepub.com/toc/crub/os-26/1>
- Langendijk, J. A., Boersma, L. J., Rasch, C. R., van Vulpen, M., Reitsma, J. B., van der Schaaf, A., & Schuit, E. (2018). Clinical Trial Strategies to Compare Protons With Photons. *Seminars in Radiation Oncology*, 28(2), 79–87. <https://doi.org/10.1016/j.semradonc.2017.11.008>
- Lehrack, S., Assmann, W., Bertrand, D., Henrotin, S., Herault, J., Heymans, V., Stappen, F. V., Thirolf, P. G., Vidal, M., Walle, J. V. D., & Parodi, K. (2017). Submillimeter ionoacoustic range determination for protons in water at a clinical synchrocyclotron. *Physics in Medicine & Biology*, 62(17), L20–L30. <https://doi.org/10.1088/1361-6560/aa81f8>
- Liu, X. (2012). Classification accuracy and cut point selection. *Statistics in Medicine*, 31(23), 2676–2686. <https://doi.org/10.1002/SIM.4509>
- Low, D. A., Harms, W. B., Mutic, S., & Purdy, J. A. (1998). A technique for the quantitative evaluation of dose distributions. *Medical Physics*, 25(5), 656–661. <https://doi.org/10.1118/1.598248>
- Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 405(2), 442–451. [https://doi.org/10.1016/0005-2795\(75\)90109-9](https://doi.org/10.1016/0005-2795(75)90109-9)
- Meijers, A., Free, J., Wagenaar, D., Deffet, S., Knopf, A. C., Langendijk, J. A., & Both, S. (2020). Validation of the proton range accuracy and optimization of CT calibration curves utilizing range probing. *Physics in Medicine & Biology*, 65(3), 03NT02. <https://doi.org/10.1088/1361-6560/ab66e1>
- Meijers, A., Oria, C. S., Free, J., Langendijk, J. A., Knopf, A. C., & Both, S. (2021). Technical Note: First report on an in vivo range probing quality control procedure for scanned proton beam therapy in head and neck cancer patients. *Medical Physics*, 48(3), 1372–1380. <https://doi.org/10.1002/mp.14713>
- Min, C. H., Kim, C. H., Youn, M. Y., & Kim, J. W. (2006). Prompt gamma measurements for locating the dose falloff region in the proton therapy. *Applied Physics Letters*, 89(18), 183517. <https://doi.org/10.1063/1.2378561>
- Min, C. H., Lee, H. R., Kim, C. H., & Lee, S. B. (2012). Development of array-type prompt gamma measurement system for in vivo range verification in proton therapy. *Medical Physics*, 39(4), 2100–2107. <https://doi.org/10.1118/1.3694098>
- Mishra, M. V., Aggarwal, S., Bentzen, S. M., Knight, N., Mehta, M. P., & Regine, W. F. (2017). Establishing Evidence-Based Indications for Proton Therapy: An Overview of Current Clinical Trials. *International Journal of Radiation Oncology Biology Physics*, 97(2), 228–235. <https://doi.org/10.1016/j.ijrobp.2016.10.045>
- Möhler, C., Russ, T., Wohlfahrt, P., Elter, A., Runz, A., Richter, C., & Greilich, S. (2018). Experimental verification of stopping-power prediction from single- and dual-energy

- computed tomography in biological tissues. *Physics in Medicine & Biology*, 63(2), 025001. <https://doi.org/10.1088/1361-6560/aaa1c9>
- Moteabbed, M., Trofimov, A., Sharp, G. C., Wang, Y., Zietman, A. L., Efstathiou, J. A., & Lu, H. M. (2016). A Prospective Comparison of the Effects of Interfractional Variations on Proton Therapy and Intensity Modulated Radiation Therapy for Prostate Cancer. *International Journal of Radiation Oncology Biology Physics*, 95(1), 444–453. <https://doi.org/10.1016/J.IJROBP.2015.12.366>
- Mumot, M., Algranati, C., Hartmann, M., Schippers, J. M., Hug, E., & Lomax, A. J. (2010). Proton range verification using a range probe: Definition of concept and initial analysis. *Physics in Medicine & Biology*, 55(16), 4771–4782. <https://doi.org/10.1088/0031-9155/55/16/010>
- Nenoff, L., Priegnitz, M., Janssens, G., Petzoldt, J., Wohlfahrt, P., Trezza, A., Smeets, J., Pausch, G., & Richter, C. (2017). Sensitivity of a prompt-gamma slit-camera to detect range shifts for proton treatment verification. *Radiotherapy and Oncology*, 125(3), 534–540. <https://doi.org/10.1016/j.radonc.2017.10.013>
- Paganetti, H. (2012). Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Physics in Medicine & Biology*, 57(11), R99–R117. <https://doi.org/10.1088/0031-9155/57/11/r99>
- Paganetti, H. (2014). Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. *Physics in Medicine & Biology*, 59(22), R419–R472. <https://doi.org/10.1088/0031-9155/59/22/R419>
- Paganetti, H., Beltran, C., Both, S., Dong, L., Flanz, J., Furutani, K., Grassberger, C., Grosshans, D. R., Knopf, A. C., Langendijk, J. A., Nystrom, H., Parodi, K., Raaymakers, B. W., Richter, C., Sawakuchi, G. O., Schippers, M., Shaitelman, S. F., Teo, B. K., Unkelbach, J., ... Lomax, T. (2021a). Roadmap: proton therapy physics and biology. *Physics in Medicine & Biology*, 66(5), 05RM01. <https://doi.org/10.1088/1361-6560/ABCD16>
- Paganetti, H., Botas, P., Sharp, G. C., & Winey, B. (2021b). Adaptive proton therapy. *Physics in Medicine & Biology*, 66(22), 22TR01. <https://doi.org/10.1088/1361-6560/AC344F>
- Park, J. H., Kim, S. H., Ku, Y., Kim, C. H., Lee, H. R., Jeong, J. H., Lee, S. B., & Shin, D. H. (2019). Multi-slit prompt-gamma camera for locating of distal dose falloff in proton therapy. *Nuclear Engineering and Technology*, 51(5), 1406–1416. <https://doi.org/10.1016/j.net.2019.03.008>
- Particle Therapy Co-Operative Group. (2022). *Statistics of patients treated in particle therapy facilities worldwide*. Verfügbar 26. Juni 2022 unter <https://www.ptcog.ch/index.php/patient-statistics-2>
- Patch, S. K., Covo, M. K., Jackson, A., Qadadha, Y. M., Campbell, K. S., Albright, R. A., Bloemhard, P., Donoghue, A. P., Siero, C. R., Gimpel, T. L., Small, S. M., Ninemire,

- B. F., Johnson, M. B., & Phair, L. (2016). Thermoacoustic range verification using a clinical ultrasound array provides perfectly co-registered overlay of the Bragg peak onto an ultrasound image. *Physics in Medicine & Biology*, *61*(15), 5621–5638. <https://doi.org/10.1088/0031-9155/61/15/5621>
- Pausch, G., Berthold, J., Enghardt, W., Römer, K., Straessner, A., Wagner, A., Werner, T., & Kögler, T. (2020). Detection systems for range monitoring in proton therapy: Needs and challenges. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *954*, 161227. <https://doi.org/10.1016/J.NIMA.2018.09.062>
- Perali, I., Celani, A., Bombelli, L., Fiorini, C., Camera, F., Clementel, E., Henrotin, S., Janssens, G., Prieels, D., Roellinghoff, F., Smeets, J., Stichelbaut, F., & Stappen, F. V. (2014). Prompt gamma imaging of proton pencil beams at clinical dose rate. *Physics in Medicine & Biology*, *59*(19), 5849–5871. <https://doi.org/10.1088/0031-9155/59/19/5849>
- Permatasari, F. F., Eulitz, J., Richter, C., Wohlfahrt, P., & Lühr, A. (2020). Material assignment for proton range prediction in Monte Carlo patient simulations using stopping-power datasets. *Physics in medicine and biology*, *65*(18), 185004. <https://doi.org/10.1088/1361-6560/AB9702>
- Persson, G. F., Nygaard, D. E., Hollensen, C., af Rosenschöld, P. M., Mouritsen, L. S., Due, A. K., Berthelsen, A. K., Nyman, J., Markova, E., Roed, A. P., Roed, H., Korreman, S., & Specht, L. (2012). Interobserver delineation variation in lung tumour stereotactic body radiotherapy. *The British Journal of Radiology*, *85*(1017), 1211–1328. <https://doi.org/10.1259/bjr/76424694>
- Peters, N., Wohlfahrt, P., Dahlgren, C. V., de Marzi, L., Ellerbrock, M., Fracchiolla, F., Free, J., Gomà, C., Góra, J., Jensen, M. F., Kajdrowicz, T., Mackay, R., Molinelli, S., Rinaldi, I., Rompokos, V., Siewert, D., van der Tol, P., Vermeren, X., Nyström, H., ... Richter, C. (2021). Experimental assessment of inter-centre variation in stopping-power and range prediction in particle therapy. *Radiotherapy and Oncology*, *163*, 7–13. <https://doi.org/10.1016/J.RADONC.2021.07.019>
- Peters, N., Wohlfahrt, P., Hofmann, C., Möhler, C., Menkel, S., Tschiche, M., Krause, M., Troost, E. G., Enghardt, W., & Richter, C. (2022). Reduction of clinical safety margins in proton therapy enabled by the clinical implementation of dual-energy CT for direct stopping-power prediction. *Radiotherapy and Oncology*, *166*, 71–78. <https://doi.org/10.1016/J.RADONC.2021.11.002>
- Petzoldt, J., Janssens, G., Nenoff, L., Richter, C., & Smeets, J. (2018). Correction of Geometrical Effects of a Knife-Edge Slit Camera for Prompt Gamma-Based Range Verification in Proton Therapy. *Instruments*, *2*(4), 25. <https://doi.org/10.3390/instruments2040025>

- Pietsch, J., Khamfongkhrua, C., Berthold, J., Janssens, G., Stützer, K., Löck, S., & Richter, C. (2022). Automatic detection and classification of treatment deviations in proton therapy from realistically simulated prompt gamma imaging data. *Medical Physics*, *50*(1), 506–517. <https://doi.org/10.1002/mp.15975>
- Piplack, N. (2021, November). *Evaluierung eines prompt-gamma-basierten Systems zur automatisierten Erkennung anatomischer Veränderungen bei Prostata-Bestrahlungen in der Protonentherapie* [Diplomarbeit]. Technische Universität Dresden.
- Priegnitz, M., Barczyk, S., Nenoff, L., Golnik, C., Keitz, I., Werner, T., Mein, S., Smeets, J., Vander Stappen, F., Janssens, G., Hotoiu, L., Fiedler, F., Prieels, D., Enghardt, W., Pausch, G., & Richter, C. (2016a). Towards clinical application: prompt gamma imaging of passively scattered proton fields with a knife-edge slit camera. *Physics in Medicine & Biology*, *61*(22), 7881–7905. <https://doi.org/10.1088/0031-9155/61/22/7881>
- Priegnitz, M., Helmbrecht, S., Janssens, G., Perali, I., Smeets, J., Stappen, F. V., Sterpin, E., & Fiedler, F. (2015). Measurement of prompt gamma profiles in inhomogeneous targets with a knife-edge slit camera during proton irradiation. *Physics in Medicine & Biology*, *60*(12), 4849–4871. <https://doi.org/10.1088/0031-9155/60/12/4849>
- Priegnitz, M., Helmbrecht, S., Janssens, G., Perali, I., Smeets, J., Stappen, F. V., Sterpin, E., & Fiedler, F. (2016b). Detection of mixed-range proton pencil beams with a prompt gamma slit camera. *Physics in Medicine & Biology*, *61*(2), 855–871. <https://doi.org/10.1088/0031-9155/61/2/855>
- Qin, N., Botas, P., Giantsoudi, D., Schuemann, J., Tian, Z., Jiang, S. B., Paganetti, H., & Jia, X. (2016). Recent developments and comprehensive evaluations of a GPU-based Monte Carlo package for proton therapy. *Physics in Medicine and Biology*, *61*(20), 7347–7362. <https://doi.org/10.1088/0031-9155/61/20/7347>
- Richter, C., Pausch, G., Barczyk, S., Priegnitz, M., Keitz, I., Thiele, J., Smeets, J., Stappen, F. V., Bombelli, L., Fiorini, C., Hotoiu, L., Perali, I., Prieels, D., Enghardt, W., & Baumann, M. (2016). First clinical application of a prompt gamma based in vivo proton range verification system. *Radiotherapy and Oncology*, *118*(2), 232–237. <https://doi.org/10.1016/j.radonc.2016.01.004>
- Sarosiek, C., DeJongh, E. A., Coutrakon, G., DeJongh, D. F., Duffin, K. L., Karonis, N. T., Ordoñez, C. E., Pankuch, M., Rykalin, V., Winans, J. R., & Welsh, J. S. (2021). Analysis of characteristics of images acquired with a prototype clinical proton radiography system. *Medical Physics*, *48*(5), 2271–2278. <https://doi.org/10.1002/mp.14801>
- Schippers, M. (2012). Proton Accelerators. In H. Paganetti (Hrsg.), *Proton therapy physics* (S. 61–102). CRC press.
- Schlegel, W., Thieke, C., Jäkel, O., Fast, M., & Knopf, A.-C. (2018). Bestrahlungsverfahren. In W. Schlegel, C. P. Karger & O. Jäkel (Hrsg.), *Medizinische Physik: Grundlagen –*

- Bildgebung – Therapie – Technik* (S. 525–577). Springer Spektrum. https://doi.org/10.1007/978-3-662-54801-1_26
- Schmid, S., Landry, G., Thieke, C., Verhaegen, F., Ganswindt, U., Belka, C., Parodi, K., & Dedes, G. (2015). Monte Carlo study on the sensitivity of prompt gamma imaging to proton range variations due to interfractional changes in prostate cancer patients. *Physics in Medicine & Biology*, *60*(24), 9329–9347. <https://doi.org/10.1088/0031-9155/60/24/9329>
- Shakirin, G., Braess, H., Fiedler, F., Kunath, D., Laube, K., Parodi, K., Priegnitz, M., & Enghardt, W. (2011). Implementation and workflow for PET monitoring of therapeutic ion irradiation: A comparison of in-beam, in-room, and off-line techniques. *Physics in Medicine & Biology*, *56*(5), 1281–1298. <https://doi.org/10.1088/0031-9155/56/5/004>
- Smeets, J., Roellinghoff, F., Prieels, D., Stichelbaut, F., Benilov, A., Busca, P., Fiorini, C., Peloso, R., Basilavecchia, M., Frizzi, T., Dehaes, J. C., & Dubus, A. (2012). Prompt gamma imaging with a slit camera for real-time range control in proton therapy. *Physics in Medicine & Biology*, *57*(11), 3371–3405. <https://doi.org/10.1088/0031-9155/57/11/3371>
- Sørensen, B. S., Pawelke, J., Bauer, J., Burnet, N. G., Dasu, A., Høyer, M., Karger, C. P., Krause, M., Schwarz, M., Underwood, T. S., Wagenaar, D., Whitfield, G. A., & Lühr, A. (2021). Does the uncertainty in relative biological effectiveness affect patient treatment in proton therapy? *Radiotherapy and Oncology*, *163*, 177–184. <https://doi.org/10.1016/j.radonc.2021.08.016>
- Sterpin, E., Janssens, G., Smeets, J., Stappen, F. V., Prieels, D., Priegnitz, M., Perali, I., & Vynckier, S. (2015). Analytical computation of prompt gamma ray emission and detection for proton range verification. *Physics in Medicine & Biology*, *60*(12), 4915–4946. <https://doi.org/10.1088/0031-9155/60/12/4915>
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209–249. <https://doi.org/10.3322/caac.21660>
- Taasti, V. T., Michalak, G. J., Hansen, D. C., Deisher, A. J., Kruse, J. J., Krauss, B., Muren, L. P., Petersen, J. B. B., & McCollough, C. H. (2017). Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples. *Physics in Medicine & Biology*, *63*(1), 015012. <https://doi.org/10.1088/1361-6560/aa952f>
- Trofimov, A., Nguyen, P. L., Efstathiou, J. A., Wang, Y., Lu, H. M., Engelsman, M., Merrick, S., Cheng, C. W., Wong, J. R., & Zietman, A. L. (2011). Interfractional variations in the set-up of pelvic bony anatomy and soft tissue, and their implications on the delivery of proton therapy for localized prostate cancer. *International Journal of Radiation Oncology Biology Physics*, *80*(3), 928–937. <https://doi.org/10.1016/J.IJROBP.2010.08.006>

- Troost, E. G., Thorwarth, D., & Oyen, W. J. (2015). Imaging-based treatment adaptation in radiation oncology. *Journal of Nuclear Medicine*, 56(12), 1922–1929. <https://doi.org/10.2967/jnumed.115.162529>
- Unkelbach, J., Alber, M., Bangert, M., Bokrantz, R., Chan, T. C., Deasy, J. O., Fredriksson, A., Gorissen, B. L., Herk, M. V., Liu, W., Mahmoudzadeh, H., Nohadani, O., Siebers, J. V., Witte, M., & Xu, H. (2018). Robust radiotherapy planning. *Physics in Medicine & Biology*, 63(22), 22TR02. <https://doi.org/10.1088/1361-6560/aae659>
- Unkelbach, J., & Paganetti, H. (2018). Robust Proton Treatment Planning: Physical and Biological Optimization. *Seminars in Radiation Oncology*, 28(2), 88–96. <https://doi.org/10.1016/J.SEMRADONC.2017.11.005>
- Verburg, J. M., Hueso-Gonzalez, F., Tattenberg, S., Wohlfahrt, P., Ruggieri, T., & Bortfeld, T. (2020, 12.–14. Juli). *First-In-Human Use of Prompt Gamma-Ray Spectroscopy for Proton Range Verification* [Konferenzvortrag Abstrakt]. 2020 Joint AAPM | COMP Virtual Meeting. <https://w3.aapm.org/meetings/2020AM/programInfo/programAbs.php?sid=8807&aid=53508>
- Verburg, J. M., & Seco, J. (2014). Proton range verification through prompt gamma-ray spectroscopy. *Physics in Medicine & Biology*, 59(23), 7089–7106. <https://doi.org/10.1088/0031-9155/59/23/7089>
- Wang, Y., Efstathiou, J. A., Sharp, G. C., Lu, H. M., Frank Ciernik, I., & Trofimov, A. V. (2011). Evaluation of the dosimetric impact of interfractional anatomical variations on prostate proton therapy using daily in-room CT images. *Medical Physics*, 38(8), 4623–4633. <https://doi.org/10.1118/1.3604152>
- Water, T. A., Bijl, H. P., Schilstra, C., Pijls-Johannesma, M., & Langendijk, J. A. (2011). The Potential Benefit of Radiotherapy with Protons in Head and Neck Cancer with Respect to Normal Tissue Sparing: A Systematic Review of Literature. *The Oncologist*, 16(3), 366–377. <https://doi.org/10.1634/theoncologist.2010-0171>
- Weber, D. C., Langendijk, J. A., Grau, C., & Thariat, J. (2020). Proton therapy and the European Particle Therapy Network: The past, present and future. *Cancer/Radiothérapie*, 24(6–7), 687–690. <https://doi.org/10.1016/j.canrad.2020.05.002>
- Werner, T., Berthold, J., Hueso-González, F., Koegler, T., Petzoldt, J., Roemer, K., Richter, C., Rinscheid, A., Straessner, A., Enghardt, W., & Pausch, G. (2019). Processing of prompt gamma-ray timing data for proton range measurements at a clinical beam delivery. *Physics in Medicine & Biology*, 64(10), 105023. <https://doi.org/10.1088/1361-6560/AB176D>
- White, M. C., Holman, D. M., Boehm, J. E., Peipins, L. A., Grossman, M., & Henley, S. J. (2014). Age and cancer risk: A potentially modifiable relationship. *American Journal of Preventive Medicine*, 46(3S1), S7–S15. <https://doi.org/10.1016/j.amepre.2013.10.029>

- Wiegand, S. (2022, 27. Juni). *Gebündeltes Know-how zur Entwicklung der Strahlentherapie der Zukunft*. Verfügbar 10. Januar 2023 unter <https://tu-dresden.de/tu-dresden/newsportal/news/gebundeltes-know-how-zur-entwicklung-der-strahlentherapie-der-zukunft>
- Wilson, R. R. (1946). Radiological Use of Fast Protons. *Radiology*, 47, 487–491. <https://doi.org/10.1148/47.5.487>
- Wohlfahrt, P. (2018, Oktober). *Dual-Energy Computed Tomography for Accurate Stopping-Power Prediction in Proton Treatment Planning* [Disseration, Technische Universität Dresden]. Qucosa. <https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-317554>
- Wohlfahrt, P., Möhler, C., Enghardt, W., Krause, M., Kunath, D., Menkel, S., Troost, E. G. C., Greilich, S., & Richter, C. (2020). Refinement of the Hounsfield look-up table by retrospective application of patient-specific direct proton stopping-power prediction from dual-energy CT. *Medical Physics*, 47(4), 1796–1806. <https://doi.org/10.1002/mp.14085>
- Wohlfahrt, P., Möhler, C., Hietschold, V., Menkel, S., Greilich, S., Krause, M., Baumann, M., Enghardt, W., & Richter, C. (2017a). Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans. *International Journal of Radiation Oncology Biology Physics*, 97(2), 427–434. <https://doi.org/10.1016/j.ijrobp.2016.10.022>
- Wohlfahrt, P., Möhler, C., Richter, C., & Greilich, S. (2018). Evaluation of Stopping-Power Prediction by Dual- and Single-Energy Computed Tomography in an Anthropomorphic Ground-Truth Phantom. *International Journal of Radiation Oncology Biology Physics*, 100(1), 244–253. <https://doi.org/10.1016/j.ijrobp.2017.09.025>
- Wohlfahrt, P., Möhler, C., Stützer, K., Greilich, S., & Richter, C. (2017b). Dual-energy CT based proton range prediction in head and pelvic tumor patients. *Radiotherapy and Oncology*, 125(3), 526–533. <https://doi.org/10.1016/j.radonc.2017.09.042>
- Wohlfahrt, P., Möhler, C., Troost, E. G., Greilich, S., & Richter, C. (2019). Dual-Energy Computed Tomography to Assess Intra- and Inter-Patient Tissue Variability for Proton Treatment Planning of Patients With Brain Tumor. *International Journal of Radiation Oncology Biology Physics*, 105(3), 504–513. <https://doi.org/10.1016/j.ijrobp.2019.06.2529>
- Wohlfahrt, P., & Richter, C. (2020). Status and innovations in pre-treatment CT imaging for proton therapy. *The British Journal of Radiology*, 93(1107), 20190590. <https://doi.org/10.1259/bjr.20190590>
- Woodard, H. Q., & White, D. R. (1986). The composition of body tissues. *The British Journal of Radiology*, 59(708), 1209–1218. <https://doi.org/10.1259/0007-1285-59-708-1209>
- Xie, Y., Ainsley, C., Yin, L., Zou, W., McDonough, J., Solberg, T. D., Lin, A., & Teo, B. K. K. (2018). Ex vivo validation of a stoichiometric dual energy CT proton stopping power

- ratio calibration. *Physics in Medicine & Biology*, 63(5), 055016. <https://doi.org/10.1088/1361-6560/aaae91>
- Xie, Y., Bentefour, E. H., Janssens, G., Smeets, J., Stappen, F. V., Hotoiu, L., Yin, L., Dolney, D., Avery, S., O'Grady, F., Prieels, D., McDonough, J., Solberg, T. D., Lustig, R. A., Lin, A., & Teo, B.-K. K. (2017). Prompt Gamma Imaging for In Vivo Range Verification of Pencil Beam Scanning Proton Therapy. *International Journal of Radiation Oncology Biology Physics*, 99(1), 210–218. <https://doi.org/10.1016/j.ijrobp.2017.04.027>
- Xie, Y., Petzoldt, J., Janssens, G., O'Grady, F., Yin, L., Bentefour, E. H., Smeets, J., Prieels, D., Lustig, R. A., Lin, A., & Teo, B. K. (2020). Prompt gamma imaging for the identification of regional proton range deviations due to anatomic change in a heterogeneous region. *British Journal of Radiology*, 93(1116), 20190619. <https://doi.org/10.1259/BJR.20190619>
- Yepes, P., Adair, A., Grosshans, D., Mirkovic, D., Poenisch, F., Titt, U., Wang, Q., & Mohan, R. (2018). Comparison of Monte Carlo and analytical dose computations for intensity modulated proton therapy. *Physics in Medicine & Biology*, 63(4), 045003. <https://doi.org/10.1088/1361-6560/aaa845>
- Yuan, T. Z., Zhan, Z. J., & Qian, C. N. (2019). New frontiers in proton therapy: Applications in cancers. *Cancer Communications*, 39(1), 61. <https://doi.org/10.1186/s40880-019-0407-3>
- Yule, G. U. (1912). On the Methods of Measuring Association Between Two Attributes. *Journal of the Royal Statistical Society*, 75(6), 579–652. <https://doi.org/10.2307/2340126>
- Zhu, X., & Fakhri, G. E. (2013). Proton Therapy Verification with PET Imaging. *Theranostics*, 3(10), 731–740. <https://doi.org/10.7150/thno.5162>

Danksagungen

Dankbar schaue ich auf die wertvollen Begegnungen, die hilfsbereite Unterstützung und die fachlichen Ratschläge während meiner Zeit als Doktorand am OncoRay zurück. Allen voran möchte ich Prof. Dr. Wolfgang Enghardt danken, der mir als Doktorvater die Möglichkeit gegeben hat, in diesem interdisziplinären Forschungsumfeld zu promovieren und immer als Ansprechpartner mit einem offenen Ohr zur Verfügung stand. Weiterhin möchte ich Prof. Dr. Dietmar Georg dafür danken, dass er als Gutachter für diese Arbeit zur Verfügung stand.

Ein besonderer Dank gilt auch meinem Betreuer Prof. Dr. Christian Richter für die vielen Stunden der Diskussion, die Lenkung der Arbeit sowie seine fachliche Anleitung und Förderung. Durch ihn lernte ich eine gute Balance zwischen der Bearbeitung von Detailfragen und dem Voranbringen des großen Ganzen zu finden.

Die Unterstützung vieler Kollegen aus dem OncoRay hat diese Arbeit auf vielfältige Weise bereichert. Zunächst möchte ich mich ganz herzlich bei meiner Arbeitsgruppe *Hochpräzisions-Protonentherapie* für den regen wissenschaftlichen Austausch und eine tolle Arbeitsatmosphäre bedanken. Gemeinsam mit Dr. Chirasak Khamfongkhurea und Julian Pietsch konnte überhaupt erst die gewaltige Datengrundlage von Patientenbestrahlungen für diese Arbeit geschaffen werden. Außerdem bereicherten sie die zahlreichen Nachtschichten, in denen wir sehr produktiv mit Experimenten zu Gange waren. Ein spezieller Dank gilt Dr. Kristin Stützer, die in Diskussionen oftmals neue Ideen und Perspektiven eingebracht hat, und Dr. Steffen Löck, der mir bei vielen statistischen Fragen weiterhelfen konnte. Dr. Nils Peters und Dr. Patrick Wohlfahrt waren in allen Belangen der CT-Bildgebung kompetente Ansprechpartner und darüber hinaus auch wertvolle Wegbegleiter. Die von mir betreuten Bachelor-, Diplom- und Masterstudenten Angelina Joost, Nick Piplack bzw. Lisa Hübinger haben mich sehr bei der Auswertung der Patientendaten unterstützt und wesentlich dazu beigetragen, die manuelle Klassifikation zu etablieren. Für seine technische Unterstützung bei der Herstellung der winkelverstellbaren Halterung zur Validierung der erweiterten Simulationssoftware möchte ich Michael Reiche danken.

Überaus dankbar blicke ich auch auf die große Unterstützung aus der Klinik zurück. Hervorheben möchte ich insbesondere Dr. Tobias Hölscher, der die Aufnahme von Prostata-Tumorpatienten in der Studie mit großem Engagement ermöglicht hat. Eine große Unterstützung

Danksagungen

bei der Studiendurchführung erhielt ich auch von Julia Thiele und ihrem Team von MTRAs sowie dem Team der medizinischen Physik, dem ich auch herzlich für die Durchführung der Zebra-Messung für das gemeinsam mit MGH durchgeführte Experiment danke. Ich danke der Klinikleitung Prof. Dr. Mechthild Krause und Prof. Dr. Dr. Esther Troost für die Unterstützung der PRIMA-Studie sowie die Diskussion der Ergebnisse und möglicher nächster Schritte.

Ein wichtiger Teil dieser Arbeit bezieht sich auf die Zusammenarbeit mit dem MGH in Boston, die im Wesentlichen von Dr. Joost Verburg, Dr. Fernando Hueso-González und Dr. Patrick Wohlfahrt begleitet wurde. Für die Unterstützung und Zurverfügungstellung der gemeinsamen Datenauswertung möchte ich mich herzlich bedanken.

Weiterhin ist die Firma IBA während der Erstellung dieser Arbeit ein verlässlicher Kooperationspartner gewesen. Dazu gehörte der technische Support des zur Verfügung gestellten PGI-Systems sowie die Wartung und Erweiterung der Simulationssoftware durch Dr. Julien Smeets, Dr. Guillaume Janssens und vormals Dr. Johannes Petzoldt. Auch das IBA Team vor Ort unter der Leitung von Enrico Weser stand jederzeit hilfsbereit zur Seite, wenn es um die Planung von Nachtschichten und die Belange der Protonenbestrahlungsanlage ging. Anerkennen möchte ich auch die Leistung von Erik Traneus der Firma RaySearch, welcher ein für diese Arbeit wichtiges Auswertungstool für die RayStation zu Verfügung gestellt hat.

Zum Schluss möchte ich mich ganz besonders bei meiner Frau Anne-Sophie bedanken, die mir immer den Rücken freigehalten und mich mit ihrem wertvollen Rat stets unterstützt hat. Genauso danke ich meiner Familie und meinen Freunden für ihren Rückhalt sowie viele motivierende Gespräche. Ein herzliches Dankeschön gilt zuletzt auch allen, die geholfen haben diese Arbeit Korrektur zu lesen und mir wertvolle Hinweise zur Verbesserung gegeben haben.

Erklärungen zur Eröffnung des Promotionsverfahren

1. Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.
2. Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich Unterstützungsleistungen von folgenden Personen erhalten:
Dipl.-Ing. Nick Piplack,
M.Sc. Lisa Hübinger,
Dr. Nils Peters,
Dr. Patrick Wohlfahrt,
Dr. Fernando Hueso-González,
Dr. Kristin Stützer,
Prof. Dr. Christian Richter und
Prof. Dr. Wolfgang Enghardt.
3. Weitere Personen waren an der geistigen Herstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe eines kommerziellen Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.
4. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.
5. Die Inhalte dieser Dissertation wurden in folgender Form veröffentlicht:
 - Berthold, J., Khamfongkhrua, C., Petzoldt, J., Thiele, J., Hölscher, T., Wohlfahrt, P., Peters, N., Jost, A., Hofmann, C., Janssens, G., Smeets, J., & Richter, C. (2021). First-In-Human Validation of CT-Based Proton Range Prediction Using Prompt Gamma Imaging in Prostate Cancer Treatments. *International Journal of Radiation Oncology Biology Physics*, 111(4), 1033–1043. <https://doi.org/10.1016/j.IJROBP.2021.06.036>
6. Ich bestätige, dass es keine zurückliegenden erfolglosen Promotionsverfahren gab.
7. Ich bestätige, dass ich die Promotionsordnung der Medizinischen Fakultät Carl Gustav Carus der Technischen Universität Dresden anerkenne.
8. Ich habe die Zitierrichtlinien für Dissertationen an der Medizinischen Fakultät der Technischen Universität Dresden zur Kenntnis genommen und befolgt.

9. Ich bin mit den an der Technischen Universität Dresden geltenden „Richtlinien zur Sicherung guter wissenschaftlicher Praxis, zur Vermeidung wissenschaftlichen Fehlverhaltens und für den Umgang mit Verstößen“ einverstanden.

Dresden, 28. Februar 2023

Erklärung über die Einhaltung der gesetzlichen Bestimmungen

Hiermit bestätige ich die Einhaltung der folgenden aktuellen gesetzlichen Vorgaben im Rahmen meiner Dissertation

- das zustimmende Votum der Ethikkommission bei Klinischen Studien, epidemiologischen Untersuchungen mit Personenbezug oder Sachverhalten, die das Medizinproduktegesetz betreffen. Ethikvotum-Nr.: EK181042015
- ~~Die Einhaltung der Bestimmungen des Tierschutzgesetzes~~
nicht zutreffend
- ~~Die Einhaltung des Gentechnikgesetzes~~
nicht zutreffend
- die Einhaltung von Datenschutzbestimmungen der Medizinischen Fakultät und des Universitätsklinikums Carl Gustav Carus.

Dresden, 28. Februar 2023