

Aus der Klinik und Poliklinik für Radiologie  
der Ludwig-Maximilians-Universität München

Direktor: Prof. Dr. Jens Ricke

**Detektion von Imaging Biomarkern zur Erstellung eines prädiktiven  
Radiomics-Profiles zur Charakterisierung des Prostatakarzinoms**

Dissertation

zum Erwerb des Doktorgrades der Medizin

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität München

vorgelegt von

Andrea Gisela Maria Schreier

aus

Linz

2022

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

Berichterstatter: Prof. Dr. med. Jens Ricke

Mitberichterstatter: Prof. Dr. med. Karl-Hans Englmeier  
Prof. Dr. med. Michael Ewers

Mitbetreuung durch den  
promovierten Mitarbeiter: Prof. Dr. Michael Ingrisch  
PD Dr. med. Dominik Nörenberg

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 28.04.2022

# Inhaltsverzeichnis

|  |              |
|--|--------------|
| Abkürzungsverzeichnis .....  | VI           |
| Abbildungsverzeichnis.....   | IX           |
| Tabellenverzeichnis.....   | XI           |
| Motivation der Arbeit.....   | XII          |
| <b>1 Einleitung.....</b>   | <b>- 1 -</b> |
| 1.1 <i>Das Prostatakarzinom: relevant und wichtig zu erforschen?</i> .....   | - 1 -        |
| 1.1.1 Ätiologie, Klinik und Pathologie des Prostatakarzinoms .....   | - 2 -        |
| 1.2 <i>Konventionelle Prostatakarzinomdiagnostik</i> .....   | - 4 -        |
| 1.2.1 Basisdiagnostik und Diagnosesicherung.....   | - 4 -        |
| 1.2.2 Risikoklassifizierung und aktive Überwachung.....  | - 5 -        |
| 1.2.3 Probleme des konventionellen Diagnose- und Risikostratifizierungsprozesses .....   | - 6 -        |
| 1.3 <i>Die MRT in der Prostatakarzinomdiagnostik</i> .....   | - 8 -        |
| 1.3.1 Grundlagen.....  | - 8 -        |
| Die Prostata in der MRT .....  | - 10 -       |
| 1.3.2 Die Rolle der mpMRT in der Prostatakarzinomdiagnostik .....  | - 11 -       |
| Die MRT-TRUS-Fusionsbiopsie .....  | - 11 -       |
| Karzinomdetektion und lokales Staging.....   | - 13 -       |
| Therapieplanung und aktive Überwachung .....   | - 14 -       |
| 1.3.3 Qualitative, strukturierte Befundung: PI-RADS-Score .....  | - 15 -       |
| Definition .....   | - 15 -       |
| Beispielfälle.....   | - 17 -       |
| Verbesserungen in der Detektion klinisch signifikanter Prostatakarzinome durch die mpMRT<br>sowie PI-RADS 2.0 und Limitationen ..... | - 19 -       |
| 1.3.4 Quantitative Beurteilung mit dem ADC-Wert .....  | - 20 -       |
| Definition .....   | - 20 -       |
| Verbesserungen in der Detektion klinisch signifikanter Prostatakarzinome durch den ADC-Wert<br>und Limitationen .....                | - 21 -       |
| 1.3.5 Radiomics: Automatisierte, quantitative Bildanalyse und Machine-learning Verfahren. - 21                                       | - 21 -       |
| -  | -            |
| Definition .....   | - 21 -       |
| Radiomics-Workflow .....   | - 22 -       |
| Radiomics-Ansätze zur Charakterisierung des Prostatakarzinoms aus der aktuellen Literatur - 24                                       | - 24 -       |
| -  | -            |

|          |   |               |
|----------|---|---------------|
| <b>2</b> | <b>Zielsetzung</b> .....  | <b>- 27 -</b> |
| 2.1      | <i>Arbeitshypothese</i> .....   | - 27 -        |
| <b>3</b> | <b>Material und Methoden</b> .....  | <b>- 29 -</b> |
| 3.1      | <i>Studiendesign</i> .....  | - 29 -        |
| 3.1.1    | Studientyp.....   | - 29 -        |
| 3.1.2    | Studienpopulation .....   | - 29 -        |
|          | Einschlusskriterien.....  | - 29 -        |
| 3.2      | <i>Bildgebung: MRT-Untersuchung</i> .....   | - 31 -        |
| 3.3      | <i>Strukturierte qualitative Befundung (PI-RADS)</i> .....  | - 31 -        |
| 3.4      | <i>Quantitative Analyse der DW-MRT</i> .....  | - 31 -        |
| 3.5      | <i>Radiomics-Analyse</i> .....  | - 32 -        |
| 3.5.1    | Segmentierung.....  | - 32 -        |
| 3.5.2    | Bildnachbearbeitung.....  | - 33 -        |
| 3.5.3    | Feature-Berechnung .....  | - 34 -        |
| 3.5.4    | Training der Klassifikatoren und Feature-Auswahl.....   | - 35 -        |
| 3.6      | <i>Statistische Analysen</i> .....  | - 37 -        |
| <b>4</b> | <b>Ergebnisse</b> .....   | <b>- 38 -</b> |
| 4.1      | <i>Demographische und klinische Merkmale der Patientenstichprobe</i> .....  | - 38 -        |
| 4.2      | <i>Beispielfälle</i> .....  | - 40 -        |
| 4.2.1    | Bioptisch gesichertes klinisch signifikantes Prostatakarzinom.....  | - 40 -        |
|          | Fall 1.....   | - 40 -        |
|          | Fall 2.....   | - 41 -        |
| 4.2.2    | Bioptisch gesichertes klinisch nicht signifikantes Prostatakarzinom beziehungsweise kein Karzinomnachweis in der Histologie ..... | - 42 -        |
|          | Fall 3.....   | - 42 -        |
|          | Fall 4.....   | - 43 -        |
| 4.3      | <i>ROC-Analysen</i> .....   | - 44 -        |
| 4.3.1    | Strukturierte qualitative Befundung (PI-RADS) .....   | - 44 -        |
| 4.3.2    | Quantitative Analyse der DW-MRT (mADC) .....  | - 45 -        |
| 4.3.3    | Radiomics-Analyse .....   | - 46 -        |
| 4.3.4    | Vergleich PI-RADS, mADC und Radiomics.....  | - 50 -        |
| 4.4      | <i>Zusammenhänge zwischen verschiedenen Variablen und klinisch signifikanten Prostatakarzinomen</i> .....                         | - 53 -        |
| 4.4.1    | PSA-Wert und klinisch signifikante Prostatakarzinome.....   | - 53 -        |
| 4.4.2    | PSA-Dichte und klinisch signifikante Prostatakarzinome .....  | - 54 -        |

|          |   |               |
|----------|---|---------------|
| 4.4.3    | Prostatavolumen und klinisch signifikante Prostatakarzinome .....   | - 55 -        |
| 4.4.4    | PI-RADS-Score und klinisch signifikante Prostatakarzinome .....   | - 56 -        |
| 4.4.5    | Mittlerer ADC-Wert und klinisch signifikante Prostatakarzinome .....  | - 57 -        |
| 4.5      | <i>Genauigkeit des Histologie-Ergebnisses der Biopsie im Vergleich zum<br/>Prostatektomiepräparat .....</i> | <i>- 58 -</i> |
| <b>5</b> | <b>Diskussion .....</b>   | <b>- 59 -</b> |
| 5.1      | <i>Interpretation der Ergebnisse .....</i>  | <i>- 59 -</i> |
| 5.1.1    | PI-RADS .....   | - 59 -        |
| 5.1.2    | ADC.....  | - 60 -        |
| 5.1.3    | Radiomics .....   | - 60 -        |
| 5.2      | <i>Einordnung der Ergebnisse in den aktuellen Stand der Forschung .....</i>                                 | <i>- 61 -</i> |
| 5.3      | <i>Besondere Stärken und Limitationen.....</i>  | <i>- 64 -</i> |
| 5.4      | <i>Konsequenzen .....</i>   | <i>- 67 -</i> |
| 5.4.1    | Zukünftige Rolle der mpMRT in der Prostatadiagnostik.....   | - 67 -        |
|          | Biopsienaive Patienten .....  | - 67 -        |
|          | MRT-TRUS-Fusionsbiopsie.....  | - 69 -        |
|          | Aktive Überwachung .....  | - 70 -        |
|          | Biparametrische MRT (bpMRT) .....   | - 71 -        |
| 5.4.2    | Zukünftige Rolle von Radiomics-Modellen .....   | - 72 -        |
| 5.5      | <i>Herausforderungen und nächste Schritte in Radiomics.....</i>   | <i>- 74 -</i> |
| 5.6      | <i>Ausblick.....</i>  | <i>- 75 -</i> |
| <b>6</b> | <b>Zusammenfassung.....</b>   | <b>- 77 -</b> |
|          | <b>Literaturverzeichnis.....</b>  | <b>- 79 -</b> |
|          | <b>Danksagung .....</b>   | <b>- 85 -</b> |
|          | <b>Eidesstattliche Versicherung .....</b>   | <b>- 87 -</b> |

# Abkürzungsverzeichnis

|              |  |
|--------------|--|
| <b>ACR</b>   | American College of Radiology                            |
| <b>ADC</b>   | Apparent Diffusion Coefficient                           |
| <b>mADC</b>  | Mean Apparent Diffusion Coefficient                      |
| <b>AFMS</b>  | Anteriores Fibromuskuläres Stroma                        |
| <b>AUC</b>   | Area Under Curve   |
| <b>BPH</b>   | Benigne Prostatahyperplasie                              |
| <b>CAD</b>   | Computer-Aided Diagnosis                                 |
| <b>CDR</b>   | Cancer Detection Rate                                    |
| <b>CT</b>    | Computertomographie                                      |
| <b>CUP</b>   | Cancer of Unknown Primary                                |
| <b>CZ</b>    | Zentrale Zone  |
| <b>DCE</b>   | Dynamic Contrast Enhancement                             |
| <b>DKFZ</b>  | Deutsches Krebsforschungszentrum                         |
| <b>DWI</b>   | Diffusion Weighted Imaging                               |
| <b>EAU</b>   | European Association of Urology                          |
| <b>EPE</b>   | Extraprostatic Extension (extraprostatische Ausbreitung) |
| <b>ERSPC</b> | Europ. Randomized Study for Screening of Prostate-Cancer |
| <b>ESUR</b>  | European Society of Urogenital Radiology                 |
| <b>GLCM</b>  | Gray Level Cooccurrence Matrix                           |
| <b>GLRLM</b> | Gray Level Run Length Matrix                             |
| <b>GLSZM</b> | Gray Level Size Zone Matrix                              |

|                      |  |
|----------------------|--|
| <b>HE-Färbung</b>    | Hämatoxylin-Eosin-Färbung  |
| <b>HIFU-Therapie</b> | Hochintensive fokussierte Ultraschall-Therapie                         |
| <b>IBSI</b>          | Image Biomarker Standardization Initiative                             |
| <b>MITK</b>          | Medical Imaging Interaction Toolkit                                    |
| <b>mpMRT</b>         | Multiparametrische MRT   |
| <b>bpMRT</b>         | Biparametrische MRT  |
| <b>NVB</b>           | Neurovascular Bundle (neurovaskuläres Bündel)                          |
| <b>PI-RADS</b>       | Prostate Imaging Reporting and Data System                             |
| <b>PRECISION</b>     | Prostate Evaluation for Clinically Important Disease (Studie)          |
| <b>PROMIS</b>        | Prostate MRI Imaging Study   |
| <b>PSA</b>           | Prostata-spezifisches Antigen  |
| <b>PSMA-PET</b>      | Prostataspezifisches Membranantigen – Positronen-Emissions-Tomographie |
| <b>PZ</b>            | Periphere Zone   |
| <b>mRMR</b>          | Minimum Redundancy Maximum Relevance                                   |
| <b>ROC</b>           | Receiver Operating Characteristic                                      |
| <b>ROI</b>           | Region of Interest   |
| <b>R-Status</b>      | Residualtumor-Status   |
| <b>SMOTE</b>         | Synthetic Minority Oversampling Technique                              |
| <b>SVI</b>           | Seminal Vesicle Invasion (Samenblaseninfiltration)                     |
| <b>SVM</b>           | Support Vector Machines  |
| <b>T1w, T2w</b>      | T1 weighted (T1-gewichtet), T2 weighted (T2-gewichtet)                 |
| <b>TNM-Stadium</b>   | Tumor-Nodus-Metastasen-Stadium   |

|                     |                                      |
|---------------------|--------------------------------------|
| <b>TRUS-Biopsie</b> | Transrektale-Ultraschall-Biopsie     |
| <b>TZ</b>           | Transitionalzone                     |
| <b>UICC</b>         | Union International Contre le Cancer |
| <b>USPSTF</b>       | US Preventive Services Task Force    |
| <b>VOI</b>          | Volume of Interest                   |

# Abbildungsverzeichnis

|   |      |
|---|------|
| Abb. 1: Skizze der zonalen Anatomie der Prostata nach McNeal (13). .....  | 3 -  |
| Abb. 2: Axiales Schnittbild einer Prostata in der T2-Wichtung. ....   | 10 - |
| Abb. 3: 79-jähriger Patient mit einem PSA-Wert von 7,5 ng/ml.....   | 17 - |
| Abb. 4: 61-jähriger Patient mit einem PSA-Wert von 25 ng/ml.....  | 17 - |
| Abb. 5: 75-jähriger Patient mit einem PSA-Wert von 20,7 ng/ml.....  | 18 - |
| Abb. 6: 55-jähriger Patient mit einem PSA-Wert von 3,9 ng/ml.....   | 18 - |
| Abb. 7: 63-jähriger Patient mit einem PSA-Wert von 9,9 ng/ml.....   | 19 - |
| Abb. 8: Radiomischer Workflow.....  | 24 - |
| Abb. 9: Patienteneinschluss.....  | 30 - |
| Abb. 10: Beispiel der Segmentierung eines Prostatakarzinoms bei einem 73-jährigen Patienten. PSA-Wert 12,0 ng/ml. Verdächtiger Herdbefund in der posteromedialen peripheren Zone links in der Drüsenmitte (PI-RADS 4, T3a). ... | 33 - |
| Abb. 11: Prinzip der Texturparameter.....   | 35 - |
| Abb. 12: 60-jähriger Patient mit einem PSA-Wert von 5,9 ng/ml.....  | 40 - |
| Abb. 13: 73-jähriger Patient mit einem PSA-Wert von 4,6 ng/ml.....  | 41 - |
| Abb. 14: 60-jähriger Patient mit einem PSA-Wert von 3,6 ng/ml.....  | 42 - |
| Abb. 15: 66-jähriger Patient mit einem PSA-Wert von 17,1 ng/ml und klinischem Verdacht auf ein Prostatakarzinom.....  | 43 - |
| Abb. 16: ROC-Kurve der Befundung nach PI-RADS. ....   | 44 - |
| Abb. 17: ROC-Kurve der quantitativen Analyse anhand des mittleren ADC-Wertes.....   | 45 - |
| Abb. 18: ROC-Kurven der Radiomics-Analyse: logistische Regression. ....   | 47 - |
| Abb. 19: ROC-Kurven der Radiomics-Analyse: „Basis“-Random-Forest-Algorithmus („class weights“, mRMR-Feature-Auswahl mit 25 Features).....   | 47 - |
| Abb. 20: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus mit „class weights“ und mRMR-Feature-Auswahl mit 50 Features. ....   | 48 - |
| Abb. 21: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus mit „class weights“ und mRMR-Feature-Auswahl mit 100 Features. ....  | 48 - |
| Abb. 22: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus ohne „class weights“ und mRMR-Feature-Auswahl mit 25 Features. ....  | 49 - |
| Abb. 23: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus mit „class weights“ ohne mRMR-Feature-Auswahl. ....  | 49 - |

Abb. 24: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus mit SMOTE statt „class weights“ und mRMR-Feature-Auswahl mit 25 Features.- 50 -

Abb. 25: Scatterplot zur Übersicht über die AUC-Werte aller Methoden hinsichtlich ihrer Differenzierungsfähigkeit von klinisch signifikanten versus klinisch nicht signifikanten Prostatakarzinomen. .... - 51 -

Abb. 26: Boxplot der PSA-Werte bei klinisch signifikanten und nicht signifikanten Prostatakarzinomen..... - 53 -

Abb. 27: Boxplot der PSA-Dichte bei klinisch signifikanten und klinisch nicht signifikanten Prostatakarzinomen. .... - 54 -

Abb. 28: Boxplot der Prostatavolumina bei klinisch signifikanten und nicht signifikanten Prostatakarzinomen. .... - 55 -

Abb. 29: Punktplot der PI-RADS-Scores bei klinisch signifikanten und klinisch nicht signifikanten Prostatakarzinomen. n = Anzahl der Fälle. .... - 56 -

Abb. 30: Boxplot der mittleren ADC-Werte bei klinisch signifikanten und klinisch nicht signifikanten Prostatakarzinomen. .... - 57 -

# Tabellenverzeichnis

|   |        |
|---|--------|
| Tabelle 1: Risikoklassifizierung des lokal begrenzten Prostatakarzinoms nach Thompson et al.....  | - 5 -  |
| Tabelle 2: Übersicht über die Sequenzen der mp-Prostata-MRT.....  | - 11 - |
| Tabelle 3: Beurteilung nach PI-RADS 2.0 (52).....   | - 16 - |
| Tabelle 4: MRT-Protokoll des klinikinternen 3-Tesla-Scanners (Magnetom Skyra, Siemens).....   | - 31 - |
| Tabelle 5: Merkmale der eingeschlossenen Patienten (n = 140). ....  | - 39 - |
| Tabelle 6: Übersicht über die AUC-Werte aller Methoden hinsichtlich ihrer Differenzierungsfähigkeit von klinisch signifikanten versus klinisch nicht signifikanten Prostatakarzinomen. .... | - 52 - |
| Tabelle 7: Kontingenztafel zur Genauigkeit des Biopsieergebnisses (mit der Histologie aus dem Prostatektomiepräparat als Goldstandard). ....  | - 58 - |

## Motivation der Arbeit

Das Prostatakarzinom ist weltweit die zweithäufigste Krebserkrankung (1) und in Deutschland die zweithäufigste Krebstodesursache des Mannes (2). Es kann in seiner Histologie, Tumoraggressivität und lokalen Ausbreitung sehr heterogen sein, weswegen eine exakte Risikostratifizierung für ein effizientes Krankheitsmanagement unumgänglich ist. Die Beurteilung der Größe und Lage des Tumors, der Tumoraggressivität sowie eines möglichen organüberschreitenden Wachstums ist daher vor einer Therapie entscheidend, um die optimale Therapiestrategie festzulegen. Diese kann dann, je nach Krankheitsstadium, von einer reinen Beobachtung mit regelmäßigen Kontrollen über medikamentöse oder minimal-invasive Therapieverfahren bis hin zu einer radikalen operativen Entfernung der gesamten Drüse reichen (3).

Ziel der vorliegenden Arbeit ist es, durch die automatisierte, quantitative Bildanalyse zusammen mit Verfahren des maschinellen Lernens (sog. Radiomics-Techniken) die in den Bilddaten enthaltenen Informationen optimal zu nutzen und so durch eine Magnetresonanztomografie-basierte Charakterisierung des Prostatakarzinoms die Risikostratifizierung im prätherapeutischen Diagnoseprozess zu unterstützen. In Zeiten der Präzisionsmedizin soll damit prospektiv in erster Linie ein Beitrag zur Verbesserung der persönlichen Therapieplanung und des individuellen Patientenergebnisses geleistet werden. Die Magnetresonanztomografie könnte somit zukünftig im Krankheitsmanagement des Prostatakarzinoms eine weitaus größere Rolle einnehmen als bisher und hier vor allem zu einer exakteren Risikostratifizierung beitragen.

# 1 Einleitung

## 1.1 Das Prostatakarzinom: relevant und wichtig zu erforschen?

Als das Prostatakarzinom 1853 zum ersten Mal identifiziert wurde, ging man damals noch von einer „sehr seltenen Krankheit“ aus (4). Diese Annahme hat sich zunehmend als gegenteilig erwiesen, so wird mittlerweile von einem „führenden Gesundheitsproblem“ gesprochen (5). Tatsächlich ist das Prostatakarzinom aktuell weltweit die zweithäufigste Krebserkrankung des Mannes (1) und die zweithäufigste Krebstodesursache bei Männern in Deutschland (2). Männer in westlichen Industrieländern haben ein 40-prozentiges Lebenszeitrisiko an einem Prostatakarzinom zu erkranken, jedoch nur ein 10-prozentiges Risiko Symptome zu entwickeln und nur ein 3-prozentiges Risiko zu versterben (6). Allerdings sinkt die Sterblichkeit nur in Ländern mit hohem Lebensstandard, in Ländern mit niedrigerem Lebensstandard steigt sie im Laufe der letzten Jahre leicht an (7). Das mittlere Erkrankungsalter liegt bei 69 Jahren, wobei relevant ist, dass sich, ausgehend vom Stand 2014, bis 2050 der Anteil an Männern über 60 Jahren in der deutschen Bevölkerung verdoppeln soll (8). Demnach ist das Prostatakarzinom nicht nur aktuell ein sehr häufiges und somit sehr relevantes Krankheitsbild, sondern wird in Zukunft sogar weiter an Bedeutung gewinnen.

Die sinkende Mortalität erkrankter Männer in Ländern mit hohem Lebensstandard erfordert ein neues Krankheitsmanagement, da nicht alle detektierten Prostatakarzinome zwingend einer definitiven Therapie unterzogen werden müssen. Die Entscheidung zu einer Therapie sowie der definitiven Wahl der Therapie hängt maßgeblich von der Tumoraggressivität und der Risikokonstellation des Patienten ab. Genauso müssen in Ländern mit geringerem Lebensstandard Maßnahmen zur Verbesserung der Versorgung von Patienten mit einem Prostatakarzinom getroffen werden.

Auch das Gesundheitsbudget würde äußerst von einem reformierten Krankheitsmanagement des Prostatakarzinoms profitieren. Durch eine präzisere Risikostratifizierung, reduzierte Biopsieraten sowie der Auswahl adäquater

Therapiemethoden für den einzelnen Patienten könnte sich eine drastischere Kostenreduktion ergeben: Eine Ende 2018 in JAMA publizierte Studie von Trogon et al. untersuchte die Kosten, die durch fast 50 000 an Prostatakarzinom erkrankten Männern  $\geq 70$  Jahren mit lokalisiertem Prostatakarzinom im Zeitraum zwischen 2004 bis 2007 entstanden (9). Die Median-Kosten lagen drei Jahre nach Diagnose bei Patienten mit definitiver Therapie durchschnittlich bei 14.000 USD, bei Patienten mit risikoarmen Prostatakarzinomen (Gleason-Score  $\leq 6$ , s. 1.2.2) und reiner Verlaufsbeobachtung mit regelmäßigen Kontrollen nur bei knapp 2.000 USD. Hätte man letzteres Vorgehen bei entsprechenden Patienten mit risikoarmen Prostatakarzinomen konsequent angewandt, hätten so 320 Millionen USD eingespart und die Ausgaben von insgesamt 1,2 Milliarden USD für alle betrachteten Patienten in diesen drei Jahren deutlich reduziert werden können.

Ein an der Maximierung der Lebensqualität orientiertes Krankheitsmanagement des Prostatakarzinoms ist daher aufgrund der sehr hohen Prävalenz und steigenden Inzidenz bei niedriger krankheitsspezifischer Mortalität äußerst relevant und muss tiefgründiger erforscht werden. Diese Arbeit möchte hier bei der Verbesserung der prätherapeutischen Risikostratifizierung zur Entscheidungsunterstützung und individuellen Therapieplanung mittels auf MRT-Daten basierten maschinellen Lernmodellen mitwirken.

### **1.1.1 Ätiologie, Klinik und Pathologie des Prostatakarzinoms**

Auch wenn das Prostatakarzinom eine weit verbreitete Erkrankung ist, ist seine Ätiologie nicht genau geklärt. Die Leitlinie der European Association of Urology (10) nennt verschiedene Risikofaktoren, wobei das Alter als der Hauptrisikofaktor gilt. Des Weiteren spielt sowohl die genetische Prädisposition eine Rolle, als auch ernährungsbedingte und sozioökonomische Faktoren, welche die geografische Abhängigkeit der Inzidenz (erhöht in Nordamerika und -europa, erniedrigt im asiatischen und südeuropäischen Raum) erklären könnten. Studien konnten auch einen Zusammenhang zwischen chronischen Entzündungen sowie sexuell übertragbaren Krankheiten und der Entstehung von Prostatakarzinomen zeigen (2).

Durch die heutigen Früherkennungsuntersuchungen werden Prostatakarzinome meist in frühen Stadien detektiert, in denen sie sich häufig klinisch unauffällig

präsentieren. Fortgeschrittene Prostatakarzinome zeigen vermehrt Symptome, wie Dysurie, Nykturie, Pollakisurie, Hämaturie, Hämatospermie, erektile Dysfunktion, Gewichtsverlust oder Knochenschmerzen durch ossäre Metastasen. (11)

Pathomorphologisch sind Prostatakarzinome in erster Linie epitheliale Adenokarzinome. Ferner erwähnt die aktuelle WHO-Klassifikation weitere seltene und besonders differenzierte Prostatakarzinome sowie von Prostatakarzinomen zu differenzierende gutartige und atypische Läsionen (3). Die Mehrzahl der Adenokarzinome (75%) geht von der peripheren Zone der Prostata aus, 20% entwickeln sich in der Transitionalzone und nur 5% in der zentralen Zone (12). Die zonale Anatomie der Prostata ist zur Veranschaulichung in Abb. 1 dargestellt.

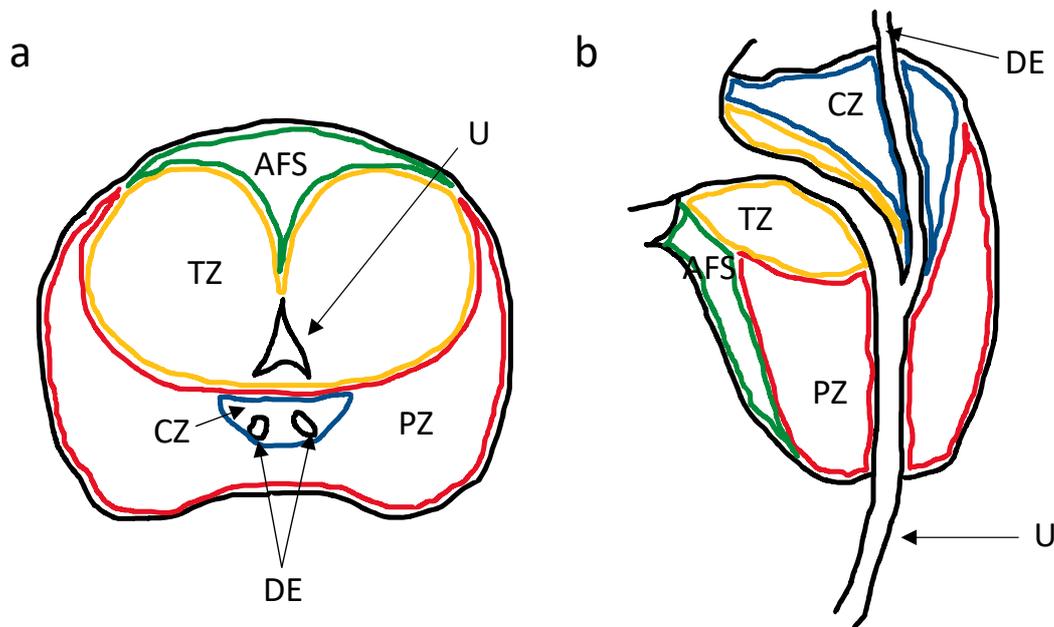


Abb. 1: Skizze der zonalen Anatomie der Prostata nach McNeal (13).

(a) axialer Schnitt. (b) sagittaler Schnitt. AFS = anteriores fibromuskuläres Stroma, CZ = zentrale Zone, PZ = periphäre Zone, TZ = Transitionalzone, DE = Ductus ejaculatorius, U = Urethra.

Neben der zonenabhängigen Entwicklung ist die Multifokalität des Prostatakarzinoms als wichtiges Tumorcharakteristikum zu nennen. Es wurde gezeigt, dass sich in 60 – 90% der Prostatakarzinome mehrere, klonal unabhängige Tumorfokalisationen in der Prostata gleichzeitig oder zeitlich versetzt entwickeln (14). Interessant ist, dass eine Studie von Liu et al. auf einen

monoklonalen Ursprung von Prostatakrebismetastasen hindeutet (15). Dies würde die Indexläsionshypothese bestärken, welche davon ausgeht, dass die pathologischen Merkmale der größten Tumorkläsion, die sogenannte Indexläsion, die Prognose bestimmen ((16), (17)). Als grundlegendes Merkmal etablierte sich primär der Gleason-Score, der laut Leitlinie neben dem TNM-Stadium und dem R-Status für die prognostische Evaluation des Prostatakarzinoms herangezogen werden soll. Dieser bereits in den 1970er Jahren entwickelte Score basiert rein auf der histologischen Morphologie des Drüsenmusters in der HE-Färbung und reicht von 2 bis 10, wobei eine höhere Zahl für ein schlechter differenziertes Karzinom steht. Der Gleason-Score korreliert mit Tumorgröße, R-Status, klinischen Verlauf und Überleben (18).

## **1.2 Konventionelle Prostatakarzinomdiagnostik**

### **1.2.1 Basisdiagnostik und Diagnosesicherung**

Laut aktueller S3-Leitlinie Prostatakarzinom besteht die Basisdiagnostik bei Verdacht auf ein Prostatakarzinom beziehungsweise im Rahmen des gesetzlichen Früherkennungsprogrammes in Deutschland für Männer ab dem 45. Lebensjahr aus einer digital-rektalen Untersuchung (DRU) der Prostata und einer nicht im Screening enthaltenen PSA-Wert-Bestimmung. Eine transrektale Ultraschalluntersuchung (TRUS) kann als ergänzende Diagnostik eingesetzt werden, weist allerdings mit der DRU vergleichbare, limitierte Testgüteparameter auf (Sensitivität etwa 59%, Spezifität etwa 94%). Unabhängig davon kann die TRUS für die Bestimmung des Prostatavolumens genutzt werden und somit hilfreich bei der Bewertung des PSA-Wertes sein. (3)

Zur Diagnosesicherung empfiehlt die S3-Leitlinie Prostatakarzinom bei einem suspekten Befund in der Basisdiagnostik eine Prostatabiopsie. Diese sollte ab einem kontrollierten PSA-Wert von 4 ng/ml (unter Einbezug der den PSA-Wert verändernden Einflussfaktoren wie beispielsweise das Prostatavolumen oder Harnwegsinfekte), einem auffälligen PSA-Anstieg (ohne Wechsel des Bestimmungsverfahrens) oder einem karzinomverdächtigen Ergebnis bei der digital-rektalen Untersuchung durchgeführt werden. Bei der herkömmlichen systematischen transrektalen Ultraschall-geführten Biopsie (TRUS-Biopsie)

werden hierbei nach einem festen Schema 10 – 12 Stenzen medial und lateral aus dem Apex, der Mitte und der Basis der Prostata entnommen und der Gleason-Score bestimmt. (3)

## 1.2.2 Risikoklassifizierung und aktive Überwachung

Für das weitere Staging und die Therapieplanung erfolgt dann eine Risikoklassifizierung in drei Kategorien nach der S3-Leitlinie (3): Das lokal begrenzte (T1 – 2 N0 M0), das lokal fortgeschrittene (T3 – 4 N0 M0) und das fortgeschrittene/metastasierte Prostatakarzinom (N1 und/oder M1). Jede Kategorie lässt sich noch untergliedern, wobei im Rahmen dieser Dissertation nur die genauere Risikoklassifizierung des lokal begrenzten Prostatakarzinoms von Interesse ist (19), welche in Tabelle 1 veranschaulicht ist.

| Risikokategorie | Gleason-Score | PSA (ng/ml)   | cT-Status  |
|-----------------|---------------|---------------|------------|
| niedrig         | 6             | ≤ 10          | cT1c, cT2a |
| intermediär     | 7             | > 10 und ≤ 20 | cT2b       |
| hoch            | ≥ 8           | > 20          | cT2c       |

Tabelle 1: Risikoklassifizierung des lokal begrenzten Prostatakarzinoms nach Thompson et al.

Die Unterscheidung in „klinisch signifikante“ und „klinisch nicht signifikante“ Prostatakarzinome wird ebenfalls zur Risikoklassifizierung in der Literatur angewendet, ist aber nicht einheitlich definiert (20). In dieser Arbeit wurde entsprechend der Leitlinie und vorheriger Studien ein klinisch signifikantes Prostatakarzinom als ein Prostatakarzinom mit einem Gleason-Score von ≥ 7 (3+4) in mindestens einem Gewebezylinder definiert (chirurgische Klassifikation).

Relevant ist diese Klassifizierung in erster Linie für die Entscheidung über mögliche Therapieoptionen. Klinisch nicht signifikante Prostatakarzinome, also nicht sehr aggressive, in einem frühen Stadium detektierte Prostatakarzinome, wachsen oft nur langsam und rufen bei älteren Patienten womöglich nie Symptome hervor, weswegen sofortige definitive Therapieoptionen, wie beispielsweise die radikale Prostatektomie, nicht immer sinnvoll sind. Wie sehr die Diagnose eines Prostatakarzinoms die Lebensqualität beeinflusst, hat eine große im Jahr 2019 in „The Lancet“ publizierte Studie von Downing et al.

untersucht (21). Dabei wurden über 30 000 Männern 18 – 42 Monate nach Diagnose zu funktionellem Befinden und allgemeiner gesundheitsbezogener Lebensqualität befragt. Ergebnis war, dass sexuelle Dysfunktion, Schwierigkeiten beim Wasserlassen und hormonelle Beschwerden eine Beeinträchtigung der Lebensqualität darstellen, ansonsten jedoch das subjektiv beurteilte Gesundheitsgefühl vergleichbar zur normalen Population ist. Nicht aggressive Prostatakarzinome können daher „aktiv überwacht“ werden („Active Surveillance“). Im Rahmen einer aktiven Überwachung soll das Prostatakarzinom in den ersten beiden Jahren durch PSA-Bestimmung und DRU alle drei Monate kontrolliert werden. Bleibt der PSA-Wert stabil, ist 6-monatlich zu untersuchen. Eine Rebiopsie sollte nach 6 Monaten durchgeführt werden und danach in den ersten drei Jahren alle 12 – 18 Monate erfolgen, später bei stabilem Befund alle drei Jahre. Bei Karzinomprogress wird eine definitive, spezifische Therapie eingeleitet. Unter folgenden Voraussetzungen ist die aktive Überwachung laut S3-Leitlinie Prostatakarzinom eine valide Option: Gleason-Score  $\leq 6$ , PSA  $\leq 10$  ng/ml, Stadium cT1 oder cT2a, Karzinomnachweis in  $\leq 2/10$  – 12 Stanzen und  $\leq 50\%$  Tumor/ Stanze. Die aktive Überwachung soll somit unter kontrollierten Bedingungen eine Übertherapie mit den verbundenen Risiken und Nebenwirkungen verhindern aber gleichzeitig ein rechtzeitiges Eingreifen ermöglichen. (3)

Demnach ist es umso wichtiger, ein exzellentes Krankheitsmanagement mit einer exakten Risikostratifizierung für die Entscheidungsunterstützung und Therapieplanung zu etablieren, um den Patienten längst möglich eine hohe Lebensqualität zu ermöglichen und die potentielle therapieassoziierte Morbidität zu minimieren. Eine MRT-basierte Herangehensweise zur bildbasierten Tumorcharakterisierung und Risikostratifizierung von karzinomsuspekten Prostataläsionen wird in dieser Arbeit untersucht.

### **1.2.3 Probleme des konventionellen Diagnose- und Risikostratifizierungsprozesses**

Zunächst lässt sich das kontrovers diskutierte Prostatakrebs-Screening (22) anführen. In Deutschland sieht das gesetzliche Früherkennungsprogramm für Männer ab dem 45. Lebensjahr einmal pro Jahr eine digital-rektale Untersuchung der Prostata vor. Eine PSA-Messung ist darin nicht enthalten, wird aber je nach

Quelle empfohlen. So sprechen sich die US Preventive Services Task Force (USPSTF) 2018 (23) und die European Association of Urology (EAU) 2019 (24) für ein PSA-Screening aus, obwohl dadurch nur weniger als ein Todesfall durch Prostatakrebs pro 1 000 Männer über eine Screeningdauer von 10 Jahren verhindert werden konnte (25). In der ERSPC (The European Randomized Study of Screening for Prostate Cancer)-Studie mit über 60 000 Männern fielen 76% der aufgrund eines erhöhten PSA-Wertes durchgeführten Biopsien negativ aus. Bei den positiv ausgefallenen Biopsien lag die Rate der Überdiagnose bei 50%, da ein klinisch nicht signifikantes Prostatakarzinom diagnostiziert wurde, das zur Lebenszeit nie zu Symptomen oder Tod geführt hätte (26). Folgen davon sind eine mögliche biopsieassoziierte Morbidität, potentiell die Lebensqualität einschränkende Auswirkungen einer aktiven Therapie, die psychische Belastung der Diagnose sowie gesellschaftlich betrachtet gesamtwirtschaftliche Nachteile. Auch ist unklar, inwiefern das Screening eine Senkung der krankheitsspezifischen Mortalität bedingt, oder ob dieser Effekt durch eine Verbesserung der Therapien verursacht wird (23). Allerdings „ist es weniger die Frage, ob Screening wirklich die Prostatakrebs-Mortalität senken kann, sondern, ob mögliche negative Folgen (Überdiagnose, Übertherapie, Beeinträchtigung der Lebensqualität) in sinnvollem Verhältnis zum vermuteten therapeutischen Effekt führen“ (Wirth et al., 2016, Die Urologie 2. Prostatakarzinom: Epidemiologie, Ätiologie, Prävention und Früherkennung/Screening, p. 1121) (27).

Zudem ist die begrenzte Aussagekraft der herkömmlichen systematischen TRUS-Biopsie problematisch. Studien ergaben, dass 10 – 30% der Prostatakarzinome der TRUS-Biopsie entgehen können (28). Es zeigte sich auch, dass nach wiederholter TRUS-Biopsie bei ungefähr einem Drittel der Patienten (27% (29), 34% (30)) ein zuvor diagnostiziertes, klinisch nicht signifikantes Prostatakarzinom zu einem klinisch signifikanten Prostatakarzinom hochgestuft wurde. Ebenso wurde bei circa einem Drittel der Patienten (36% (31), 30% (32), 25% (33)) ein klinisch nicht signifikantes Prostatakarzinom nach erneuter Bestimmung des Gleason-Scores aus dem Prostatektomiepräparat hochgestuft. Signifikante Prädiktoren für ein Upgrading waren laut beiden Studien ein höheres Alter, ein höherer PSA-Wert und ein geringeres Gewicht der Prostata. Außerdem wird bei der traditionellen TRUS-Biopsie nur äußerst selten

die extrakapsuläre Ausbreitung erfasst, welche jedoch für ein exaktes Staging und die weitere Therapieplanung essentiell ist (34).

Zusammenfassend lässt sich sagen, dass das traditionelle Vorgehen in der Prostatakarzinomdiagnostik (digital-rektale Untersuchung, PSA-Wert, systematische TRUS-Biopsie) von zusätzlichen klinischen und bildgebenden Verfahren, wie der MRT, profitieren könnte, um eine genauere Risikostratifizierung und eine bessere Identifikation von Patienten für die aktive Überwachung zu gewährleisten. Dies ist von größter Relevanz, da sich die Therapiemethoden mitsamt deren Nebenwirkungen für die unterschiedlichen Risikokategorien grundlegend unterscheiden. Ebenso hat es volkswirtschaftliche Auswirkungen, sodass eine Optimierung des Diagnoseprozesses mit weniger Krankenständen und Krankenhausaufenthalten eine deutliche Kostenreduktion im Gesundheitsbudget bewirken könnte. Diese Arbeit möchte mit einem MRT-basierten Ansatz den Mehrwert für die Risikostratifizierung im aktuellen Diagnose- und Therapieentscheidungsprozess liefern.

### **1.3 Die MRT in der Prostatakarzinomdiagnostik**

Die ersten Versuche, Prostatakarzinome mit der MRT zu lokalisieren, fanden bereits in den 1980er Jahren anhand T1- und T2-gewichteter Bilder statt (35), aber scheiterten an zu geringer Sensitivität und Spezifität. Durch stetige Verbesserungen etablierte sich die multiparametrische Magnetresonanztomografie (mpMRT), bestehend aus morphologischer und funktioneller Bildgebung, in den letzten Jahren vor allem als wichtiges Werkzeug zur Verbesserung zielgerichteter Biopsien sowie im Rahmen von Studienprotokollen auch zur Therapieplanung (3).

#### **1.3.1 Grundlagen**

Das Prinzip der MRT (36) basiert auf Signalen, die Wasserstoffatome in einem starken Magnetfeld (meist 1,5 oder 3,0 Tesla) nach Anregung mit einem Hochfrequenzimpuls aussenden. Diese Signale klingen über zwei verschiedene Prozesse ab: Die longitudinale und transversale Relaxation, die jeweils nach zwei charakteristischen Zeitspannen T1 und T2 stattfinden. Diese Zeiten sind von der Umgebung der Wasserstoffatome abhängig und charakteristisch für die

verschiedenen Gewebe. So stellt sich Fett in T1w hyperintens dar, Wasser in T1w hypo- und in T2w hyperintens, Luft in T2w hypointens. Diese beiden Gewichtungen eignen sich daher gut für die Beurteilung der Morphologie.

Die in der Prostata-Bildgebung heutzutage standardmäßig verwendete multiparametrische MRT (mpMRT) vereint Sequenzen zur morphologischen und funktionellen MRT-Bildgebung, da sie neben den morphologischen T1w- und T2w-Sequenzen auch funktionelle Sequenzen, wie die Diffusions- (diffusion weighted imaging, DWI) und Perfusionsbildgebung (dynamic contrast enhancement, DCE) beinhaltet. Die DWI-Sequenz basiert auf der Abschwächung zweier Impulse in einer T2w-Sequenz (37): Je einfacher in einem Gewebe Wassermoleküle diffundieren können, desto weniger T2-Signal senden sie beim zweiten im Vergleich zum ersten Impuls zurück, da sich zum Zeitpunkt des Eintreffens des zweiten Impulses die Wassermoleküle bereits um einiges weiterbewegt haben. Können sie wenig diffundieren, senden sie ein höheres T2-Signal zurück. Wie stark das Signal zwischen den beiden Impulsen durch die Diffusion abgeschwächt wird, bestimmt der b-Wert: Je höher der b-Wert desto stärker die Signalabschwächung. Basierend auf den DWI-Daten können mit mindestens zwei b-Werten, je mehr desto genauer, ADC-Karten erstellt werden, indem für jedes Pixel der ADC (apparent diffusion coefficient)-Wert errechnet wird. Zur Läsionscharakterisierung wird meist der Mittelwert der ADC-Werte (mean ADC, mADC) aus einer definierten ROI (Region of interest) bestimmt. Der ADC-Wert oder scheinbare Diffusionskoeffizient korreliert signifikant mit der Zelldichte (38). Somit zeigt sich ein Prostatakarzinom signalreich auf DWI-Bildern mit hohen b-Werten und signalarm auf dazugehörigen errechneten ADC-Karten. Bei der dynamischen kontrastmittelverstärkten Sequenz (DCE) werden vor, während und nach Kontrastmittelinjektion T1w-Bilder aufgenommen, um das Enhancement, die Dynamik der Signalverstärkung von Strukturen nach Kontrastmittelgabe, zu beurteilen. Die mpMRT ist bis heute die einzige Modalität, die sowohl die nötige räumliche Auflösung als auch den nötigen Weichteilkontrast hat, um Prostatakarzinome morphologisch näher charakterisieren zu können (39).

## Die Prostata in der MRT

Die Zonen der Prostata lassen sich in der MRT wie bereits erwähnt am besten in der morphologischen T2w-Sequenz beurteilen (s. Abb. 2): Die periphere Zone (PZ) liegt im axialen Schnittbild hufeisenförmig dorsal und zeigt sich hyperintens aufgrund des flüssigkeitsreichen Drüsengewebes. Ventral davon liegen die zentrale Zone (CZ) und die Transitionalzone (TZ), die beide hypointens sind und sich schwierig unterscheiden lassen. Noch einmal ventral davon liegt das anteriore fibromuskuläre Stroma (AFMS), das sich ebenfalls hypointens darstellt. Außen um die Prostata herum zeigt sich die Kapsel als ein schmales, hypointenses Band, die wichtig für die Beurteilung der extraprostatatischen Ausbreitung ist.

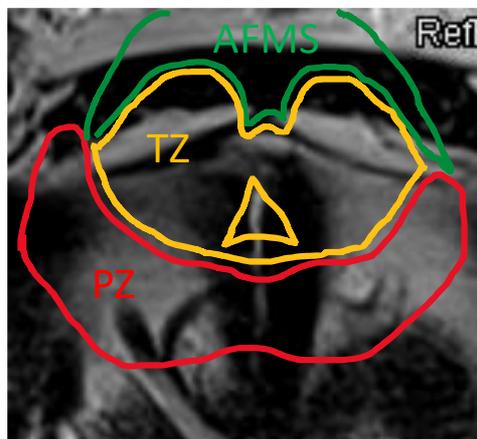


Abb. 2: Axiales Schnittbild einer Prostata in der T2-Wichtung.

PZ = periphere Zone, TZ = Transitionalzone, AFMS = anteriores fibromuskuläres Stroma.

Eine Übersicht über die für die mp-Prostata-MRT relevanten Sequenzen gibt Tabelle 2 (40).

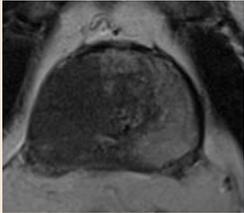
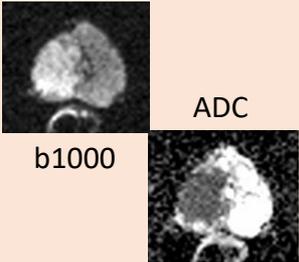
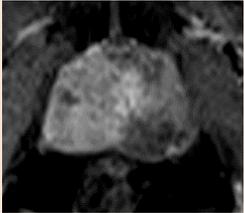
| Sequenz                                | T2w   | DWI   | DCE   |
|--|---|---|---|
| <b>Prinzip</b>                         | Gute morphologische Beurteilung von benignen und malignen Prozessen: Hochauflösende und kontrastreiche Darstellung der zonalen Prostata- und periprostatichen Anatomie (Samenbläschen, neurovaskuläre Bündel, Harnblase, Rektum und M. levator ani) | Visualisierung von zeldichteren Arealen:<br>Diffusionsrestriktion von Wassermolekülen als Marker für eine erhöhte Zelldichte und neoplastische Veränderungen von normalem Drüsengewebe  | Früheres und intensiveres Enhancement von Tumorgewebe durch tumorinduzierte Neoangiogenese (dichtere, schlechter ausgebildete Blutgefäße mit erhöhter Perfusion)  |
| <b>Rolle bei der Prostatabefundung</b> | <ul style="list-style-type: none"> <li>• Detektion und Charakterisierung von Läsionen (nach PI-RADS führende Sequenz für die TZ)</li> <li>• Lokoregionäres Staging (extraprostatistische Extension)</li> </ul>                                      | <ul style="list-style-type: none"> <li>• Detektion und Charakterisierung von Läsionen (nach PI-RADS führende Sequenz für die PZ)</li> </ul>   | <ul style="list-style-type: none"> <li>• Sekundäre Detektion und Charakterisierung von Läsionen bei Unklarheiten</li> </ul>   |
| <b>Limitationen</b>                    | <ul style="list-style-type: none"> <li>• Anfällig für Bewegungsartefakte aufgrund der langen Aufnahmezeit</li> <li>• Keine quantitativen Informationen</li> </ul>   | <ul style="list-style-type: none"> <li>• Anfällig für Artefakte durch Luft im Rektum</li> <li>• Bewegungsartefakte</li> <li>• Distorsionen</li> <li>• Nicht standardisiert, limitierte Reproduzierbarkeit der quantitativen ADC-Analyse (keine definierten Cut-off-Werte)</li> <li>• Keine sichere Differenzierung durch ADC-Werte von benignen und malignen Prozessen</li> </ul> | <ul style="list-style-type: none"> <li>• Keine sichere Differenzierung von variablem Enhancement maligner und benigner Prozesse (z.B. Prostatitis, BPH)</li> <li>• Längere Aufnahmezeit (&gt;2 min) um die Permeabilität zu beurteilen</li> </ul> |
| <b>Beispielbild</b>                    |   |    |   |

Tabelle 2: Übersicht über die Sequenzen der mp-Prostata-MRT.

Die Beispielaufnahme zeigt ein kapselüberschreitendes Prostatakarzinom in der peripheren Zone rechts mit Infiltrationen des rechten neurovaskulären Bündels. Der Tumor zeigt sich in T2w stark hypointens mit kräftiger Diffusionsstörung (hypointens in der ADC-, hyperintens in der DWI-Aufnahme) und frühem, fokalem Kontrastmittelenhancement im Vergleich zum umliegenden Gewebe in der DCE-Aufnahme. PZ = periphere Zone, TZ = Transitionalzone.

### 1.3.2 Die Rolle der mpMRT in der Prostatakarzinomdiagnostik

#### Die MRT-TRUS-Fusionsbiopsie

Die klinische Hauptanwendung der Prostata-MRT in Deutschland ist aktuell die Detektion von Prostatakarzinomen vor allem bei Patienten mit negativer TRUS-Biopsie und fortbestehendem Karzinomverdacht (41). Hier soll eine gezielte Biopsie suspekter Areale in der MRT durch Fusion dieser zuvor gewonnenen

Bilddaten mit Echtzeit-Ultraschallbildern erfolgen (MRT-TRUS-Fusionsbiopsie) (3). Der Stellenwert der Fusionsbiopsie unterscheidet sich je nach Szenario.

Im Falle einer Rebiopsie nach negativer systematischer TRUS-Erstbiopsie besteht Konsensus, dass durch eine gezielte Fusionsbiopsie mehr klinisch signifikante Prostatakarzinome detektiert werden können als durch eine erneut systematische Biopsie (42). Grund dafür ist dabei eine erhöhte Prävalenz von anterioren Prostatakarzinomen, die oft mit der systematischen Biopsie nicht erfasst, aber gut durch die mpMRT detektiert werden (43).

Im Falle einer sogenannten „Bestätigungsbiopsie“ trägt eine gezielte Fusionsbiopsie zusätzlich zur systematischen Biopsie zu einer höheren Detektionsrate von klinisch signifikanten Prostatakarzinomen bei (44). Die Durchführung einer „Bestätigungsbiopsie“ (confirmatory biopsy) wird von einigen Studien vor Einschluss in die aktive Überwachung empfohlen, um bei der Primärbioptie potentiell nicht erfasste, hochgradige Prostatakarzinome zu identifizieren (42), was eine Änderung der Therapiestrategie zur Folge hätte.

Im Falle einer Biopsie bei biopsienaiven Patienten ist noch kein genaues Vorgehen etabliert. Drei große, multizentrische Studien beschäftigten sich mit diesem Szenario. In der PROMIS-Studie (45), 2017 in „The Lancet“ publiziert, wurde die Detektionsrate für klinisch signifikante Prostatakarzinome der mpMRT und der systematischen TRUS-Biopsie verglichen: Die mpMRT erzielte eine deutlich höhere Sensitivität. Daraus wurde gefolgert, dass ein Einsatz der mpMRT als Triage vor Durchführung einer Biopsie fast einem Drittel der Patienten die Biopsie ersparen könnte, zudem würden weniger klinisch insignifikante und mehr klinisch signifikante Prostatakarzinome detektiert werden. Die PRECISION-Studie (46), 2018 im „New England Journal of Medicine“ publiziert, bestärkte diesen Gedanken weiter und zeigte, dass die gezielte Fusionsbiopsie, die nur bei Verdacht auf ein Prostatakarzinom in der mpMRT durchgeführt wurde, höhere Detektionsraten von klinisch signifikanten Prostatakarzinomen aufwies als die standardmäßig durchgeführte systematische TRUS-Biopsie. Daraus kann jedoch nicht geschlossen werden, dass die Fusions- der systematischen Biopsie überlegen ist, was sich in Zusammenschau mit der MRT-FIRST-Studie (47), 2019 in „The Lancet“ publiziert, zeigt: Hier stellte sich heraus, dass sich die gezielte Fusionsbiopsie und die systematische TRUS-

Biopsie bezüglich ihrer Detektionsrate für klinisch signifikante Prostatakarzinome nicht signifikant unterschieden und in Kombination das beste Ergebnis erzielen. Somit ersetzt die Fusionsbiopsie wahrscheinlich nicht die systematische Biopsie, die, nicht außer Acht zu lassen, in der Studie 14% der detektierten, klinisch signifikanten Prostatakarzinome erkannte, welche der Fusionsbiopsie entgangen sind. Aufgefallen ist hier auch, dass der Großteil der durch die Fusionsbiopsie identifizierten klinischen signifikanten Prostatakarzinome mit einer Läsionsgröße > 15 mm darstellten.

### **Karzinomdetektion und lokales Staging**

Die traditionelle Prostatadiagnostik sah lange bei klinischem Verdacht auf ein Prostatakarzinom (auffällige digital-rektale Untersuchung, erhöhter PSA-Wert) primär eine systematische Biopsie und erst sekundär eine mpMRT für das lokale Staging vor. Dieser Ablauf ist in der allgemeinen Karzinomdiagnostik unüblich, zumal die bioptische Diagnosesicherung konventionell erst nach einer entsprechenden Bildgebung erfolgt (42). Mit verbesserter Genauigkeit spielt die mpMRT mittlerweile auch bei der Karzinomdetektion vor einer Biopsie eine immer größere Rolle. In der aktuellen europäischen Leitlinie der EAU (European Association of Urology) (10) wird die Durchführung einer Prostata-MRT bereits vor der Primärbioptie empfohlen, in der deutschen Leitlinie (3) wird sie bisher nur als Kann-Option aufgeführt. Vor allem bei der Detektion von Prostatakarzinomen mit einem Gleason-Score  $\geq 7$  und einem Volumen > 0,5 ml erreicht sie eine hohe Sensitivität (88 – 92% bei Prostatakarzinomen mit einem Gleason-Score von 7, 96% bei Prostatakarzinomen mit einem Gleason-Score  $\geq 8$ ) (48) und entdeckt somit vor allem die relevanten, klinisch signifikanten Prostatakarzinome. Es konnte gezeigt werden, dass sich die diagnostische Genauigkeit durch zusätzliche Durchführung einer Prostata-mpMRT verglichen zur konventionellen Prostatakarzinomdiagnostik (DRU, PSA-Wert und Prostatavolumen) von 81% auf 91% anheben lässt (49). Neben der reinen Detektion von Prostatakarzinomen erlaubt die mpMRT anhand des PI-RADS-Scores und des ADC-Wertes gleichzeitig eine Einschätzung ihrer Aggressivität (s. 1.3.3 und 1.3.4). Auch wenn laut PROMIS-Studie (45) die mpMRT vor Erstbiopsie eine deutlich höhere Sensitivität und einen höheren negativen prädiktiven Wert für klinisch signifikante Prostatakarzinome als eine erste systematische TRUS-Biopsie hat und so

unnötige Biopsien reduzieren könnte, wird die genaue Rolle der mpMRT bei biopsienaiven Patienten nach wie vor kontrovers diskutiert (42).

Für das lokale Staging von Prostatakarzinomen ist die mpMRT von großer Bedeutung: Die extrakapsuläre Ausbreitung (extent of extracapsular extension, ECE), die Invasion des neurovaskulären Bündels (neurovascular bundle, NVB), der Samenblasen (seminal vesicle invasion, SVI) oder angrenzender Strukturen sowie die lokalen Lymphknoten können beurteilt werden, und beeinflussen maßgeblich das weitere Vorgehen (43).

### **Therapieplanung und aktive Überwachung**

Die mpMRT spielt eine tragende Rolle bei der Gestaltung des Therapiekonzeptes (43), vor allem bei der Planung gezielter Biopsien sowie gezielter Therapien (Strahlentherapie, Brachytherapie, fokale Therapien). Bei Letzteren ist die mpMRT zur Selektion geeigneter Patienten und genauen Lokalisierung des Karzinoms von großer Bedeutung, um eine bildbasierte, das gesunde Prostatagewebe maximal schonende Therapie zu ermöglichen. Ferner kann die Durchführung einer mpMRT vor einer Operation eine bessere Einschätzung der Tumorausbreitung ermöglichen und für ein das neurovaskuläre Bündel schonendes Operationsvorgehen sprechen, welches eine deutlich bessere postoperative Funktionalität verspricht. Es wurde gezeigt, dass nach Einbezug des MRT-Befundes der Prostata das zuvor anhand der klinischen Information geplante präoperative Resektionsausmaß in 27% abgeändert wurde und so in 61% davon das neurovaskuläre Bündel erhalten werden konnte (50).

Im Rahmen der aktiven Überwachung empfiehlt die S3-Leitlinie Folgendes: Wurde vor Einschluss in die aktive Überwachung eine initiale mpMRT mit systematischer und gegebenenfalls gezielter Biopsie durchgeführt, sollte eine Rebiopsie mit erneuter mpMRT sowie systematischer Biopsie nach zwölf Monaten erfolgen. Patienten ohne initiale mpMRT vor Einschluss in die aktive Überwachung sollten eine mpMRT mit systematischer und gegebenenfalls gezielter Biopsie innerhalb von sechs Monaten erhalten. Im weiteren Verlauf der aktiven Überwachung ist die Evidenz zur Rolle der mpMRT noch unzulänglich. Bei vorheriger Identifikation einer Läsion in der mpMRT wird vorgeschlagen, sofern die Mittel vorhanden sind, diese jährlich mit einer erneuten mpMRT zu kontrollieren und bei klinischem oder bildgebendem Befundprogress zu

rebiopsieren. Bei stabilem klinischem und bildgebendem Befund raten die Autoren zu einem 3-jährigen Zeitintervall für erneute mpMRT-Kontrollen und Rebiopsien. (3)

### **1.3.3 Qualitative, strukturierte Befundung: PI-RADS-Score**

#### **Definition**

PI-RADS (Prostate Imaging Reporting and Data System) ist ein strukturiertes Befundungssystem für die mpMRT der Prostata, mit dem Ziel, die Erstellung des Befundes sowie auch die Untersuchung zu objektivieren und deren Qualität zu verbessern. 2012 wurde von der ESUR (European Society of Urogenital Radiology) die erste Version veröffentlicht (51), nach Überarbeitung zusammen mit dem American College of Radiology (ACR) 2015 die zweite Version, PI-RADS 2.0 (52). Seit 2019 gibt es nun die erneut aktualisierte Klassifikation, PI-RADS 2.1.

Bei der Befundung nach PI-RADS 2.0 (s. Tabelle 3 ) wird jede Läsion zonenabhängig anhand strukturierter Kriterien einer PI-RADS-Kategorie zugeordnet, die mit einem Punktwert von 1 bis 5 die Wahrscheinlichkeit für das Vorliegen eines klinisch signifikanten Prostatakarzinoms angibt. Dieses wird hier wie folgt definiert: Gleason-Score  $\geq 7$  (inklusive 3+4, mit prominenter, aber nicht vorherrschender Gleason 4-Komponente) und/oder Volumen  $\geq 0,5 \text{ cm}^3$  und/oder extraprostatische Ausbreitung. Eine Biopsie zur pathologischen Abklärung wird im Allgemeinen ab einem PI-RADS-Score von 4 empfohlen. Wichtige Erneuerungen in PI-RADS 2.0 zur vorherigen Version sind, dass es nun einen PI-RADS-Gesamtscore gibt, der dem Score der Läsion mit der höchsten PI-RADS-Kategorie entspricht. Weiters gibt es für die Prostatazonen jeweils eine „dominante Sequenz“, die hauptsächlich für die Beurteilung herangezogen wird. Für die periphere Zone entspricht diese der DWI beziehungsweise ADC und für die Transitionalzone der T2w. Auf die Änderungen in der 2019 geupdateten PI-RADS 2.1 Version wird an dieser Stelle nicht näher eingegangen, da alle Läsionen in der vorliegenden Arbeit gemäß der PI-RADS 2.0 Version erhoben wurden.

| PIRADS (T2w) | PZ   | TZ   | PIRADS (DWI) | PZ oder TZ   |
|--------------|--|--|--------------|--|
| 1            | Uniforme Hyperintensität (normal)  | Homogene mittlere Signalintensität (normal)  | 1            | Keine Auffälligkeiten in ADC und DWI mit hohen b-Werten (normal)   |
| 2            | Lineare, keilförmige oder diffuse, milde Hypointensität, meist unscharfer Rand                                 | Umschriebene hypointense oder heterogene(r), eingekapselte(r) Knoten (BPH)                                     | 2            | Unschärfe Hypointensität in ADC  |
| 3            | Heterogene Signalintensität oder nicht-umschriebene, rundliche, moderate Hypointensität                        | Heterogene Signalintensität mit dunklem Rand   | 3            | Fokale milde/moderate Hypointensität in ADC und Isointensität/milde Hypointensität in DWI mit hohen b-Werten                     |
| 4            | Umschriebene, homogene, moderate Hypointensität, begrenzt auf die Prostata und < 1,5 cm im größten Durchmesser | Linsenförmige oder nicht-umschriebene, homogene, moderate Hypointensität und < 1,5 cm im größten Durchmesser   | 4            | Fokale, deutliche Hypointensität in ADC und deutliche Hyperintensität in DWI mit hohen b-Werten; < 1,5 cm im größten Durchmesser |
| 5            | Wie 4, nur ≥ 1,5cm im größten Durchmesser oder eindeutiges extraprostatistisches Wachstum/ invasives Verhalten | Wie 4, nur ≥ 1,5cm im größten Durchmesser oder eindeutiges extraprostatistisches Wachstum/ invasives Verhalten | 5            | Wie 4, nur ≥ 1,5cm im größten Durchmesser oder eindeutiges extraprostatistisches Wachstum/ invasives Verhalten                   |

**a**

**b**

| PIRADS (DCE) | PZ oder TZ   |
|--------------|--|
| -            | Kein frühes Enhancement;<br>diffuses Enhancement, nicht mit einer fokalen Läsion in T2w und/oder DWI korrespondierend;<br>fokales Enhancement, mit einer BPH-typischen Läsion in T2w korrespondierend; |
| +            | Fokales Enhancement + früheres oder zeitgleiches Enhancement im Vergleich zu angrenzendem normalen Prostatagewebe + korrespondierend zu Läsionen in T2w und/oder DWI                                   |

**c**

| PI-RADS in DWI (dominante Sequenz für die PZ) | PI-RADS in T2w | PI-RADS in DCE | PI-RADS gesamt | PI-RADS in T2w (dominante Sequenz für die TZ) | PI-RADS in DWI | PI-RADS in DCE | PI-RADS gesamt |
|---|----------------|----------------|----------------|---|----------------|----------------|----------------|
| 1   | alle           | alle           | 1              | 1   | alle           | alle           | 1              |
| 2   | alle           | alle           | 2              | 2   | alle           | alle           | 2              |
| 3   | alle           | -              | 3              | 3   | ≤ 4            | alle           | 3              |
| 3   | alle           | +              | 4              | 3   | 5              | alle           | 4              |
| 4   | alle           | alle           | 4              | 4   | alle           | alle           | 4              |
| 5   | alle           | alle           | 5              | 5   | alle           | alle           | 5              |

**d**

**e**

Tabelle 3: Beurteilung nach PI-RADS 2.0 (52).

Beurteilung der einzelnen PI-RADS-Scores (2.0) für die periphere Zone (PZ) und die Transitionalzone (TZ) in der (a) T2w-Sequenz, (b) Diffusionsbildgebung (DWI) und (c) Perfusionsbildgebung (DCE). (d) Beurteilung des PI-RADS-Gesamtscores (2.0) in der peripheren Zone (PZ) mit der DWI als dominanter Sequenz. (e) Beurteilung des PI-RADS-Gesamtscores (2.0) in der Transitionalzone (TZ) mit der T2w als dominanter Sequenz.

## Beispielfälle

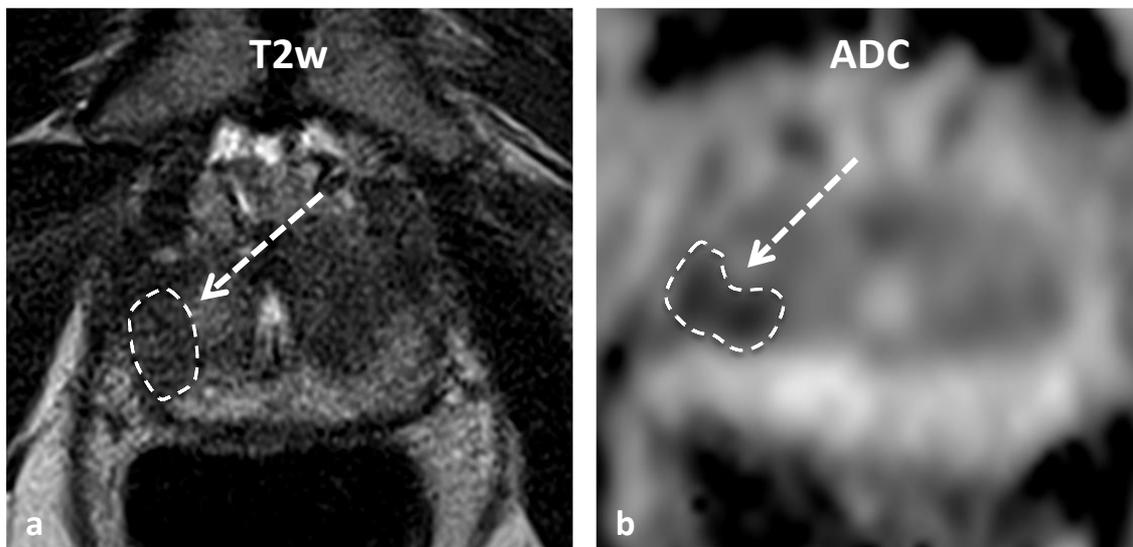


Abb. 3: 79-jähriger Patient mit einem PSA-Wert von 7,5 ng/ml.

Die axiale T2w (a) zeigt eine flau fokale T2w-Signalabsenkung in der rechtsseitigen posterolateralen peripheren Zone apexnah (Pfeil bzw. Umrandung; T2w Score: 3) mit deutlicher fokaler ADC-Absenkung (b), die korrespondierende DWI (b=1200; Bild nicht gezeigt) zeigte keine relevante Signalhyperintensität (DWI Score 3), somit PI-RADS Gesamtscore (2.0): 3. Das Prostataktomiepräparat zeigte ein klinisch signifikantes Prostatakarzinom mit einem Gleason-Score von 3+4.

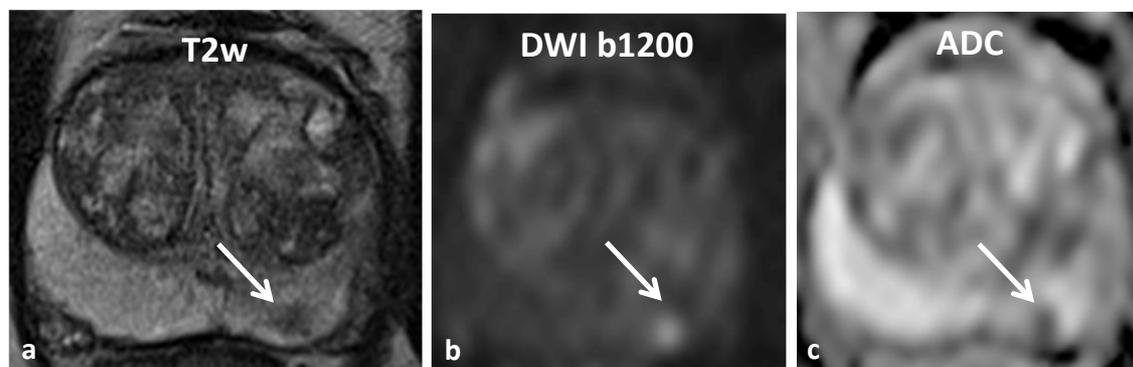


Abb. 4: 61-jähriger Patient mit einem PSA-Wert von 25 ng/ml.

(a) T2w-hypointense fokale Läsion in der posteromedialen peripheren Zone links in Drüsenmitte (Pfeil), T2w-Score 4. (b, c) Die Läsion zeigt eine deutliche fokale Diffusionsstörung mit hyperintensem Signal in den hohen b1200er Werten und entsprechender Signalabsenkung in der ADC-Map (Pfeile in b, c), DWI-Score 4. Es ergibt sich somit ein PI-RADS Gesamtscore (2.0) von 4. Die TRUS-Biopsie zeigte keinen Nachweis eines Prostatakarzinoms.

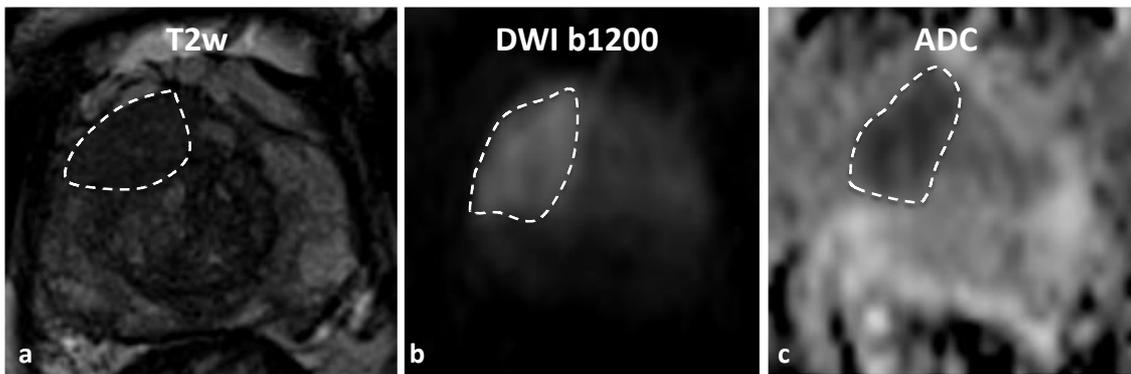


Abb. 5: 75-jähriger Patient mit einem PSA-Wert von 20,7 ng/ml.

Die axiale T2w (a) zeigt eine große hochgradig karzinomsuspekte T2w-hypointense Läsion in der rechten anterioren peripheren Zone in der Drüsenmitte (gestrichelte Umrandung; T2w Score von 5). Es zeigt sich in der korrespondierenden DWI (b) eine starke Signalhyperintensität (somit DWI Score von 5) sowie eine deutliche fokale ADC-Absenkung (c). Es ergibt sich ein PI-RADS Gesamtscore (2.0) von 5. Das Prostataektomiepräparat zeigte ein klinisch signifikantes Prostatakarzinom mit einem Gleason-Score von 5+4.

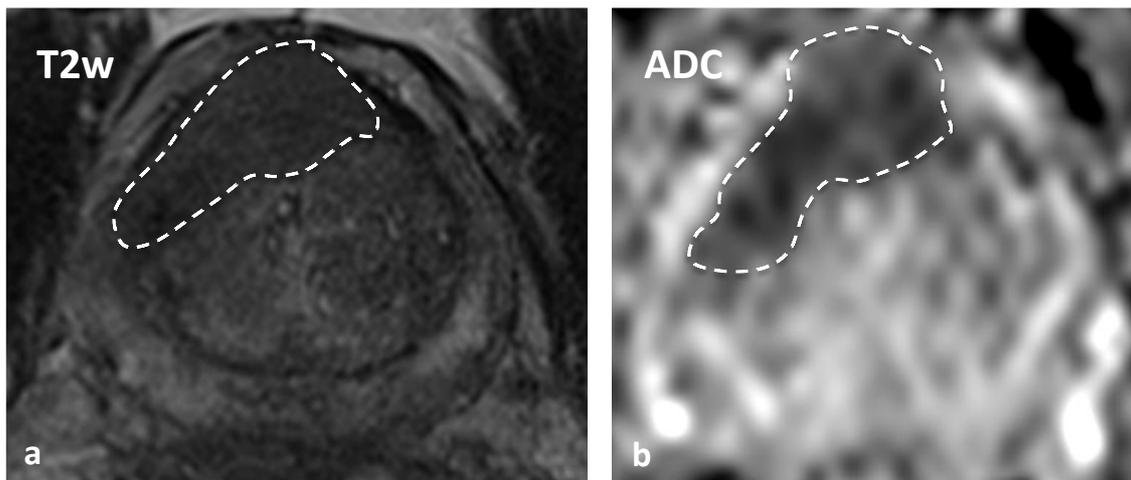


Abb. 6: 55-jähriger Patient mit einem PSA-Wert von 3,9 ng/ml.

Die axiale T2w (a) zeigt eine große hochgradig karzinomsuspekte T2w-hypointense Läsion mit lentikulärer Morphologie in der anterioren Transitionalzone in Drüsenmitte (rechts > links) (gestrichelte Umrandung; T2w Score von 5) mit deutlicher fokaler ADC-Absenkung (b), die korrespondierende DWI (b=800; Bild nicht gezeigt) zeigte eine starke Signalhyperintensität (somit DWI Score von 5). Es ergibt sich ein PI-RADS Gesamtscore (2.0) von 5. Die TRUS-Biopsie der Läsion ergab ein klinisch nicht signifikantes Prostatakarzinom mit einem Gleason Score von 3+3.

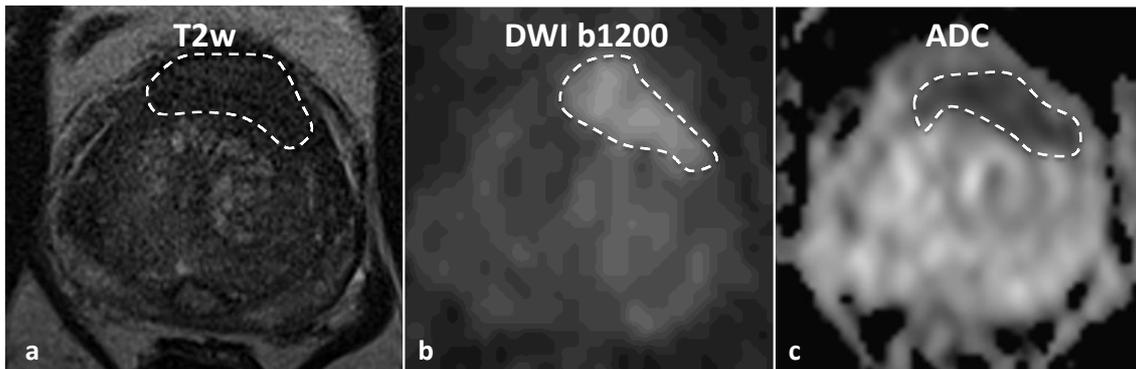


Abb. 7: 63-jähriger Patient mit einem PSA-Wert von 9,9 ng/ml.

Die axiale T2w (a) zeigt eine große hochgradig karzinomsuspekte T2w-hypointense Läsion mit lentikulärer Morphologie (bis 3,2 cm) in der Transitionalzone ventral in Drüsenmitte mit unscharfer Umrandung und breitbasigem Kapselkontakt > 1cm (gestrichelte Umrandung; T2w Score von 5), die von den multiplen T2w-hypointensen, teils heterogen kontrastierten Knoten in der restlichen Transitionalzone im Sinne einer BPH abweicht. Es zeigt sich eine Signalhyperintensität in der korrespondierenden DWI (b, somit DWI Score von 5) mit deutlicher fokaler ADC-Absenkung (c). Es ergibt sich ein PI-RADS Gesamtscore (2.0) von 5. Das Prostatektomiepräparat zeigte ein klinisch signifikantes Prostatakarzinom mit einem Gleason-Score von 3+4.

### **Verbesserungen in der Detektion klinisch signifikanter Prostatakarzinome durch die mpMRT sowie PI-RADS 2.0 und Limitationen**

In einer Metaanalyse von Woo et al. 2017 (53) wurden 21 Studien verglichen, die die diagnostische Genauigkeit des PI-RADS-Scores für die Detektion von Prostatakarzinomen untersuchten. Die gepoolte Sensitivität betrug hier 89% und die gepoolte Spezifität 73%. Sechs Studien verglichen PI-RADS 1.0 und 2.0 direkt, wobei die zweite Version zwar eine höhere Sensitivität (95% statt 88%), aber keine signifikant höhere Spezifität (73% statt 75%) in der Detektion klinisch signifikanter Karzinome zeigte. Futterer et al. untersuchten 12 PI-RADS 1.0-Studien zwischen 2011 und 2014, wobei sie nur auf eine sehr ungenaue Detektionsrate von klinisch signifikanten Prostatakarzinomen von 44 – 87% kamen (54). Außerdem zeigten zwei bereits zuvor genannte, große multizentrische Studien, die PROMIS- (45) und PRECISION-Studie (46), dass die mpMRT die Überdiagnose von klinisch insignifikanten Karzinomen verringern und gleichzeitig die Detektion von klinisch signifikanten Karzinomen erhöhen konnte. Dass hierbei eine strukturierte Vorlage mit Dropdown-Listen für die Befundung nach PI-RADS 2.0 die Compliance erhöht, den PI-RADS-Score auch wirklich zu nutzen, legten Shaish et al. dar, was demnach in einem Anstieg der

Detektionsrate von klinisch signifikanten Prostatakarzinomen resultierte. Neben einer höheren Sensitivität für klinisch signifikante Prostatakarzinome konnte durch die strukturierte Befundung mit PI-RADS 2.0 auch die Kommunikation zwischen Radiologen und Urologen verbessert werden, was bei Urologen in einem geringeren Bedürfnis resultierte, Radiologen aufgrund von Unklarheiten zu kontaktieren (55).

Hingegen unterliegt die mpMRT in ihrer aktuellen Form weiterhin relevanten Limitationen, die eine adäquate Risikostratifizierung signifikant beeinträchtigen. Erstens fällt die Detektion von Tumoren mit einem Volumen  $< 0,5$  ml aber einem Gleason-Score  $\geq 7$  relativ gering aus (20-26%, (56)). Generell wird das Tumolvolumen teilweise unterschätzt sowie ein relevanter Anteil der klinisch signifikanten Prostatakarzinome nicht erkannt, beispielsweise 16% in der Studie von Borofsky et al. (57). Zweitens verfügt die mpMRT nur über eine geringe Sensitivität in der Erkennung der Kapselinfiltration ((58), wodurch das Risiko eines positiven Schnittrandes insbesondere im Rahmen eines nervenerhaltenden operativen Vorgehens steigt. Des Weiteren sind unklare, beziehungsweise PI-RADS 3-Läsionen, eine recht heterogene Gruppe, die doch in 9 – 21% der Fälle signifikante Prostatakarzinome verbirgt ((59), (60)). Im Gegensatz dazu haben PI-RADS 4-Läsionen mit nur etwa 22% eine sehr geringe Detektionsrate für klinisch signifikante Prostatakarzinome, zumal ab PI-RADS-Kategorie 4 auf jeden Fall eine Biopsie durchgeführt werden sollte (61). Außerdem bleibt trotz Einführung der PI-RADS-Klassifikation eine gewisse Interobservervariabilität (62) und somit der Wunsch nach einer weiterführenden Befundobjektivierung bestehen, welche für eine exakte Risikostratifizierung unerlässlich ist.

### **1.3.4 Quantitative Beurteilung mit dem ADC-Wert**

#### **Definition**

Der ADC-Wert wurde im Rahmen der Diffusionsbildgebung bereits in Abschnitt 1.3.1 erläutert. Er wird basierend auf den DWI-Daten mit mindestens zwei b-Werten pro Pixel errechnet.

## **Verbesserungen in der Detektion klinisch signifikanter Prostatakarzinome durch den ADC-Wert und Limitationen**

Durch die Beurteilung mit dem ADC-Wert wurde vor allem die Möglichkeit einer einerseits objektiven, quantitativen Bildanalyse und andererseits bildbasierten Risikoabschätzung eingeführt: Es zeigte sich, dass der ADC-Wert einer Läsion mit deren Gleason-Score korreliert (63) und so zwischen klinisch signifikanten und klinisch insignifikanten Karzinomen unterschieden werden kann (64), und das mit einer Sensitivität von 91% und einer Spezifität von 63% (65).

Dennoch erfahren 20 – 40% aller Patienten nach radikaler Prostatektomie ((66), (67)) oder 30 – 50% nach Radiotherapie (68) innerhalb von 10 Jahren ein biochemisches Rezidiv und bis zu 50% der Patienten nach fokaler Therapie haben einen persistierenden Tumornachweis in der Prostata (69). Eine Optimierung, Fortentwicklung und klinische Evaluation der bisher eingesetzten Bildgebungsverfahren für eine bessere Risikostratifizierung, die letztlich eine Grundvoraussetzung für neue maßgeschneiderte Therapiekonzepte ist, scheint daher erstrebenswert. Aufgrund der Multifaktorialität der Parameter, die der Bestimmung des Tumorstadiums zu Grunde liegen, sowie einer hohen subjektiven Beeinflussung in der Bewertung der einzelnen Faktoren ist für eine verbesserte Risikostratifizierung eine Objektivierung dieser Parameter durch integrative Analysealgorithmen mit maschinellem Lernen unabdingbar.

### **1.3.5 Radiomics: Automatisierte, quantitative Bildanalyse und Machine-learning Verfahren**

#### **Definition**

Mit der Überschrift ihres Papers „Images are more than pictures, they are data“ akzentuierten Gillies et al. 2015 die Grundidee von Radiomics, einem Begriff, der sich im letzten Jahrzehnt etabliert hat: Durch die Umwandlung von medizinischen Bildern in quantitative Datensätze, die Informationen über die zugrundeliegende Pathophysiologie beinhalten, können diese analysiert werden und als Entscheidungshilfe dienen ((70), (71)). Genauer gesagt lautet die zentrale Hypothese, dass sogenannte quantitative „Imaging Biomarker“ die verschiedenen phänotypischen Unterschiede von Tumoren mitsamt der inter- als auch die intraläsionalen Tumorerheterogenität erfassen (72). Zusammen mit

klinischen Informationen können dann prädiktive Modelle entwickelt werden, welche die diagnostische, prognostische und prädiktive Genauigkeit klinischer Entscheidungen verbessern können (73). Radiomics kann auch als Weiterentwicklung der CAD (computer-aided diagnosis)-Systeme verstanden werden, welche Radiologen bei der Detektion und Diagnose des Prostatakarzinoms unterstützen sollen ((74), (75)).

Das Potential von Radiomics zeigte sich neben der Charakterisierung von Prostatakarzinomen bereits bei zahlreichen anderen Studien zur Karzinomdetektion, -charakterisierung und zur Entwicklung von Prädiktionsmodellen (unter anderem bei Glioblastomen, Kopf-Hals-, Lungen-, Mamma- oder Lebermalignomen) bei verschiedenen Bildgebungen (MRT, PET, CT). So war es mit Hilfe der Extraktion quantitativer prognostischer Biomarker aus der Bildgebung möglich, Vorhersagen zu verschiedensten klinischen Endpunkten wie Prognose, Therapieergebnis, Ursprung von CUPs (cancers of unknown primaries) oder Subtypen von Karzinomen zu treffen ((76), (77), (78), (79), (80), (81)).

### **Radiomics-Workflow**

Der Prozess, quantitative Tumoreigenschaften, sogenannte Features, aus den radiologischen Bildern zu gewinnen und diese mit spezifischen Ergebnissen zu verknüpfen, besteht grundlegend aus folgenden vier Arbeitsschritten ((82), s. Abb. 8):

Nach der Bildakquisition erfolgt zunächst die Segmentierung, bei der manuell oder automatisch bestimmte ROIs (regions of interest) umrandet werden, zum Beispiel die Tumorkläsion, um die Grenzen für das zu analysierende Volumen festzulegen.

Daraus werden dann mit einer Software im „feature extraction“-Prozess die quantitativen Tumoreigenschaften bestimmt. Diese lassen sich im Allgemeinen in drei Gruppen von Parametern einteilen: Form- (Volumen, Oberfläche, etc.), Intensitäts- und Texturparameter (GLCM, GLRLM, GLSZM, etc.) (83). Letztere Matrizen versuchen die Texturen von der genauen räumlichen Pixelverteilung zu lösen und beschreiben die Graustufenmuster benachbarter Voxel (s. 3.5.3).

Im nächsten Schritt erfolgt mittels maschinellen Lernens die Entwicklung eines prädiktiven Modells beziehungsweise eines Klassifikators. Dazu wird typischerweise zuvor der Datensatz in einen Trainings- und einen Testdatensatz unterteilt, wobei letzterer nur für die abschließende Evaluation verwendet wird. Maschinelles Lernen meint, dass Rechenmodelle aus bekannten Daten „lernen“ und auf neue, unbekannte Daten akkurate Rückschlüsse ziehen können (84). Für radiomische Modelle werden häufig SVMs (support vector machines), logistische Regressionen, oder Random Forests als Klassifikatoren verwendet. Die letzten beiden sind für diese Arbeit relevant: Die (binäre) logistische Regression untersucht, ob ein Zusammenhang zwischen einer abhängigen (binären) Variable und einer oder mehreren unabhängigen Variablen besteht. Random Forests sind ein Klassifikationsalgorithmus, der die Vorhersagen vieler einzelner Entscheidungsbäume kombiniert und dadurch eine Reihe von günstigen Eigenschaften aufweist (85).

Vor dem Trainieren eines Klassifikators kann die Vielzahl der zuvor extrahierten Features auf eine sinnvolle Anzahl reduziert werden. Eine Methode ist das „Minimum Redundancy Maximum Relevance (mRMR)“-Schema, das die relevantesten, aber nicht redundanten Features herausfiltert. Für das Training wird dann häufig die Kreuzvalidierung (crossvalidation, CV) verwendet, wobei beispielsweise bei der k-fachen Kreuzvalidierung (k-fold CV) der Datensatz in k kleinere Datensätze geteilt wird und jeweils 1 kleinerer gegen die restlichen  $1 - k$  Datensätze getestet wird. Die Performance des Modells wird schließlich anhand k Testdatensätzen evaluiert, gängige Maße sind dazu Sensitivität und Spezifität oder weiterführend die ROC (receiver operating characteristic)-Kurve mit der AUC (area under curve).

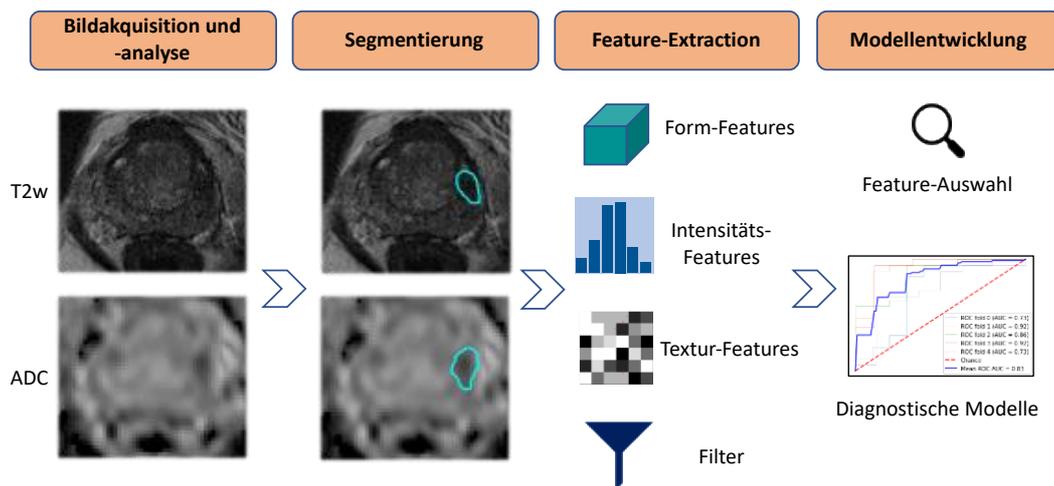


Abb. 8: Radiomischer Workflow.

Zuerst erfolgt die Bildakquisition und -analyse, danach die Segmentierung, Feature-Extraction sowie Modellentwicklung und -validierung.

### Radiomics-Ansätze zur Charakterisierung des Prostatakarzinoms aus der aktuellen Literatur

In den letzten Jahren wurden zahlreiche Studien veröffentlicht, die sich mit der Detektion von Prostatakarzinomen durch Radiomics-Modelle auseinandersetzen. Sukzessiv rückte die genauere Charakterisierung beziehungsweise Risikostratifizierung von Prostatakarzinomen immer mehr in den Fokus des Interesses, indem versucht wurde, mittels Radiomics klinisch signifikante von klinisch nicht signifikanten Prostatakarzinomen zu unterscheiden. Folgende Studien sind hierfür von besonderer Bedeutung:

Bonekamp et al. verglichen 2018 in einer retrospektiven, monozentrischen Studie die radiologische Beurteilung anhand des PI-RADS-Scores, des mittleren ADC-Wertes und eines Radiomics-Ansatzes zur Charakterisierung von Prostataläsionen (86). Die quantitative Klassifizierung der Läsionen in klinisch signifikant (Gleason Grade Group  $\geq 2$ ) und klinisch nicht signifikant anhand der ADC-Werte war hierbei genauer als die qualitative Klassifizierung durch die Radiologen (PI-RADS 2: Sensitivität 89%, Spezifität 43%, mittlerer ADC-Wert: Sensitivität 93%, Spezifität 51%). Das Radiomics-Modell erbrachte allerdings keinen signifikanten Benefit in der Detektion klinisch signifikanter Prostatakarzinome (Sensitivität 96%, Spezifität 51%).

Fehr et al. publizierten 2015 eine ebenfalls retrospektive, monozentrische Studie, in der sie mehrere Radiomics-Modelle zur Klassifikation der

Prostatakrebsaggressivität präsentierten (87). Sie hatten zum Ziel, einerseits Prostatakarzinome mit einem Gleason-Score von 6 (3+3) von denjenigen mit einem Gleason-Score  $\geq 7$  zu unterscheiden, sowie andererseits Prostatakarzinome mit einem Gleason-Score von 3+4 von denjenigen mit einem Gleason-Score von 4+3 zu differenzieren. Ihr Radiomics-Ansatz basierend auf T2w-Sequenzen und mADC-Werten erzielte in beiden Fällen eine Genauigkeit von über 90%, wohingegen ein von ihnen ebenso untersuchter, nur auf den mADC-Werten basierender Klassifikator lediglich eine Genauigkeit von circa 60% erreichte. Ihre Studie zeigte somit, dass Radiomics-Modelle mit einer hohen Sensitivität und Spezifität klinisch signifikante Prostatakarzinome erkennen können, jedoch die mADC-Werte alleine für einen entsprechend genauen Klassifikator eine unzureichende Basis darstellen und eine Kombination aus T2w-Sequenzen und mADC-Werten für das Training der Radiomics-Modelle die akkuratesten Ergebnisse liefert. Darüber hinaus wurden in dieser Studie die unbalancierten Klassen beachtet, die für viele machine-learning-Modelle ein Problem darstellen. Mit entsprechenden statistischen Methoden (s. 5.2) konnte so die Performance eines sowohl die T2w-Sequenzen als auch die mADC-Werte einbeziehenden Radiomics-Klassifikators deutlich angehoben werden, wodurch dieser erst dann genauer war als der rein auf den mADC-Werten trainierte Radiomics-Klassifikator.

Schelb et al. führten 2019 eine erneut retrospektive, monozentrische Studie durch (88): Die Aufgabe, klinisch signifikante Prostatakarzinome (Gleason Score  $\geq 3+4$ ) zu detektieren und segmentieren, wurde einerseits von Radiologen mittels PI-RADS-Score und manueller Segmentierung durchgeführt, andererseits automatisch von einem deep-learning-Radiomics-Modell, was einer weiterentwickelten Unterform des maschinellen Lernens entspricht. Anschließend wurde deren Performance verglichen: Das deep-learning-Modell erzielte ähnlich gute Ergebnisse (Sensitivität 92%, Spezifität 47%) wie die Radiologen in der Detektion klinisch signifikanter Prostatakarzinome (Sensitivität 88%, Spezifität 50%), und das, obwohl das deep-learning-Modell nur anhand von bpMRT-Daten (nur T2w- und DWI-Sequenz) trainiert wurde. Im Rahmen der PI-RADS-Beurteilung waren hingegen mpMRT-Daten mit der zusätzlichen DCE-Sequenz verfügbar.

Radiomics verspricht also eine objektive, bildbasierte, nicht-invasive Diagnose und Risikostratifizierung von Prostatakarzinomen. Bisherige Studien sind jedoch im Wesentlichen auf retrospektiven Daten basierende Proof-of-Concept-Studien ohne entsprechende Validierung oder klaren Weg in die Praxis. Den Bogen von experimenteller Forschung zu klinischer Anwendung zu spannen ist hier eine Herausforderung und bedarf noch tiefgründiger, weiterer Forschung, wozu diese Arbeit beitragen soll.

## 2 Zielsetzung

Ziel der vorliegenden Arbeit ist es, anhand eines retrospektiven, monozentrischen Datensatzes unter Verwendung von Radiomics-Verfahren die prätherapeutische Tumoraggressivität und klinische Signifikanz eines Prostatakarzinoms vorherzusagen, um zukünftig die Prognoseabschätzung und Therapieplanung bei Prostatakarzinompatienten verbessern und individualisieren zu können.

### 2.1 Arbeitshypothese

Diese Arbeit untersucht somit folgende Arbeitshypothese:

**Radiomics ermöglicht die bildbasierte Klassifizierung von klinisch signifikanten versus klinisch nicht signifikanten Prostatakarzinomen und ist herkömmlichen bildgebenden und klinischen Verfahren, wie beispielsweise dem mADC-Wert, PI-RADS-Score, PSA-Wert oder der PSA-Dichte, überlegen.**

Demnach soll mittels der Extraktion quantitativer „Imaging Biomarker“ aus MRT-Daten unter Berücksichtigung klinischer Parameter ein objektiver Analysealgorithmus entwickelt werden, der es erlaubt, schon vor einer Biopsie beziehungsweise präoperativ Patienten mit einem signifikanten Prostatakarzinom zu identifizieren. Um zu belegen, dass hier der Radiomics-Ansatz gegenüber den bisher etablierten bildbasierten Methoden einen Mehrwert bietet, soll dieser zum einen mit der konventionellen Klassifizierungsmethode anhand des PI-RADS-Scores, zum anderen mit der etablierten quantitativen Klassifizierungsmethode anhand des mittleren ADC-Wertes verglichen werden.

Durch die anvisierte prätherapeutische, bildbasierte Beurteilbarkeit der Tumoraggressivität mittels Verfahren des maschinellen Lernens entsteht in der Zukunft ein direkter Mehrwert für Patienten mit Prostatakarzinomen. Einerseits können so unnötige Biopsien sowie Übertherapien reduziert werden, andererseits leitet sich daraus prospektiv die Weiterentwicklung etablierter, individualisierter Therapieverfahren ab. Die vorliegende Arbeit setzt somit unmittelbar an einem weitverbreiteten klinischen Problem, der bisher unzureichenden Möglichkeit einer bildbasierten, prätherapeutischen Beurteilung

der Tumoraggressivität, an und versucht dieses auf innovative Art und Weise zu lösen.

## **3 Material und Methoden**

### **3.1 Studiendesign**

#### **3.1.1 Studientyp**

Im Rahmen dieser Dissertation wurde eine retrospektive, monozentrische Studie am Klinikum der Universität München (Campus Großhadern) durchgeführt. Ein positives Ethikvotum (Zeichen 613-15) liegt vor.

#### **3.1.2 Studienpopulation**

Retrospektiv wurden konsekutiv 477 Patienten erfasst, die im Zeitraum von Januar 2012 bis November 2018 regulär im Klinikalltag eine mpMRT-Bildgebung erhielten. Alle Patientendaten stammen aus den Patientenakten der Klinik und Poliklinik für Radiologie der Ludwig-Maximilians-Universität München in Kooperation mit der Urologischen Klinik und Poliklinik der Ludwig-Maximilians-Universität München (Direktor: Prof. Dr. med. Christian Stief). Von diesen 477 Patienten wurden final 140 Patienten in die Studie eingeschlossen (s. Abb. 9), von welchen wiederum folgende Daten erhoben wurden: Prostatakarzinom ja/nein, Gleason-Score der Biopsie, Methode der Gewebebegewinnung (Stanzbiopsie, Prostatektomie oder Prostatamorcellat), Prostatektomie erfolgt ja/nein (falls ja, zusätzlich Gleason-Score des Prostatektomiepräparates, TNM-Stadium, UICC-Stadium, Grading, R-Status), PI-RADS-Score (2.0), PSA-Wert, PSA-Dichte. Für die Analysen wurde, sofern vorhanden, der aus der Prostatektomie bestimmte Gleason-Score dem aus der Biopsie bestimmten Gleason-Score vorgezogen. Dies war bei 83 der 140 Patienten des finalen Kollektivs der Fall, da hier zusätzlich zur initialen Prostatastanzbiopsie eine Prostatektomie erfolgte.

#### **Einschlusskriterien**

- Patienten mit pathologisch gesicherter Diagnose eines Prostatakarzinoms
- Patienten mit Verdacht auf ein Prostatakarzinom, erhobener MRT-Bildgebung und negativem Histologiebefund < 6 Monate nach Erhalt der Bildgebung (beispielsweise Patienten mit Karzinomverdacht und erhöhtem PSA-Wert bei benigner Prostatahyperplasie oder chronischer Prostatitis)

- Ausreichende Datenqualität mit lückenloser Abdeckung des gesamten Tumors sowie der gesamten Prostata
- Vorliegen von mindestens T1w-, T2w- und DWI-Aufnahmen (Untersuchung ohne Verwendung einer endorektalen Spule)
- Vorliegen eines histologischen Gradings sowie Gleason-Scores bei gesichertem Prostatakarzinom beziehungsweise einer negativen Histologie

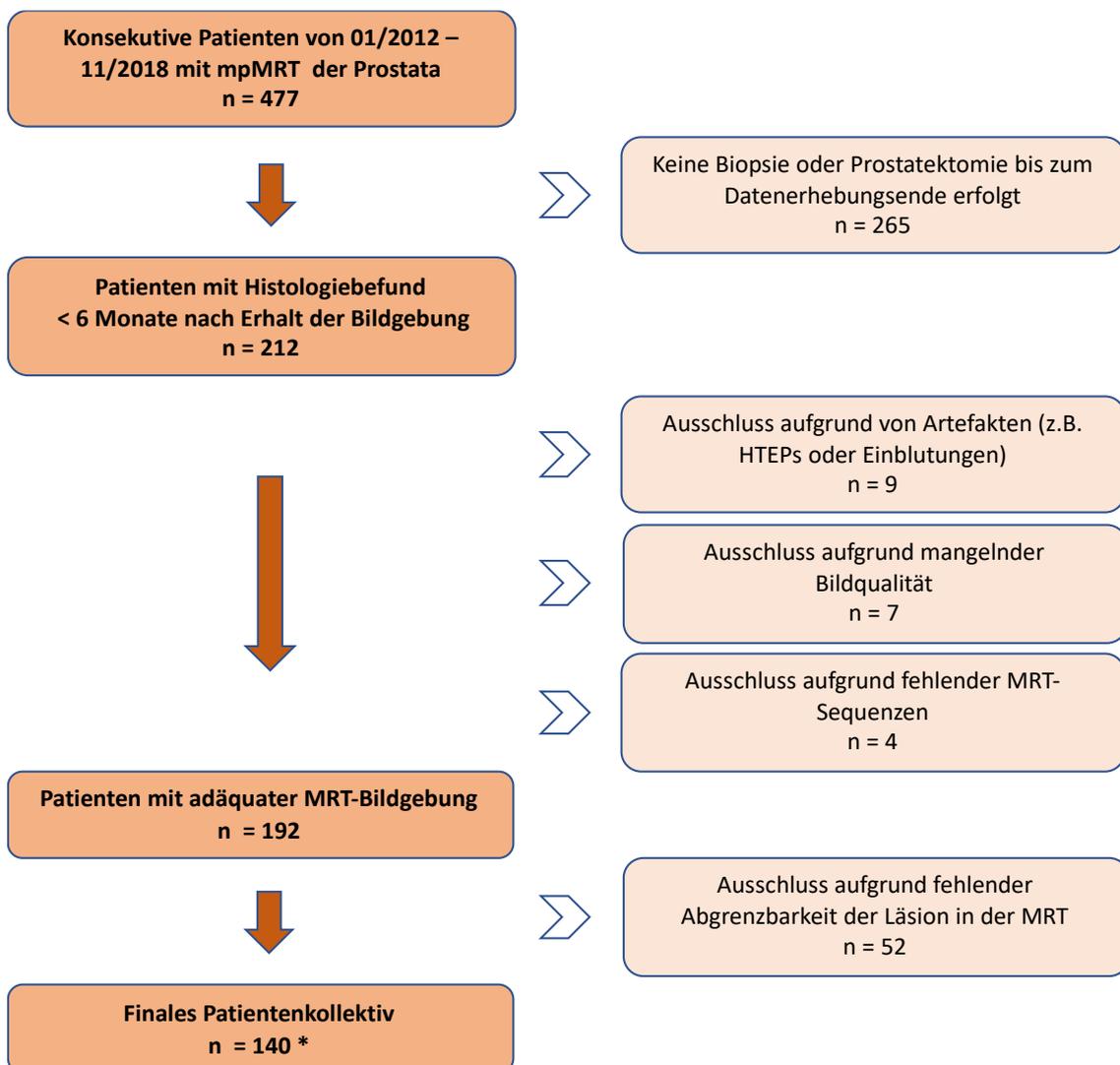


Abb. 9: Patienteneinschluss.

mpMRT = Multiparametrische MRT, HTEP = Hüft-Totalendoprothese.  
 \* Bei 83 der 140 Patienten erfolgte zusätzlich zur initialen Prostatastanzbiopsie eine Prostatektomie.

## 3.2 Bildgebung: MRT-Untersuchung

Für die MRT-Untersuchungen wurde hauptsächlich (bei 126 von 140 Patienten) der klinikinterne 3-Tesla-Scanner (Magnetom Skyra, Siemens) benutzt. Bei 14 Patienten lag eine Bildgebung von extern vor, wobei davon 3 Untersuchungen an einem 1,5-Tesla-Scanner durchgeführt wurden. Entsprechend des heterogenen Kollektivs mit unter anderem wenigen externen Datensätzen wurden die MRT-Bilddaten der eingeschlossenen Patienten nicht alle einheitlich akquiriert, dies spiegelt jedoch die klinische Realität wider und muss bei der Auswahl von Patienten für quantitative Analysen berücksichtigt werden. In Tabelle 4 ist der Übersicht halber das am häufigsten verwendete MRT-Protokoll am Standort Großhadern aufgeführt, welches als Screeninguntersuchung vorrangig ohne eine Kontrastmittelgabe (biparametrisch) erfolgte:

| MRT-Sequenzen   | Ausrichtung | TE (ms) | TR (ms) | Schichtdicke (mm) | Bildfeld (mm) | Matrix    |
|-----------------|-------------|---------|---------|-------------------|---------------|-----------|
| t2_tse_cor      | koronar     | 114     | 4000    | 3                 | 200 x 200     | 320 x 320 |
| t2_tse_tra      | transversal | 119     | 4000    | 3                 | 202 x 239     | 448 x 340 |
| t1_vibe_dixon   | transversal | 1,260   | 4,020   | 2                 | 325 x 400     | 320 x 221 |
| ep2d_diff_tra   | transversal | 61      | 4600    | 3                 | 449 x 449     | 96 x 57   |
| resolve_DWI_tra | transversal | 55      | 5010    | 3                 | 449 x 449     | 112 x 112 |

Tabelle 4: MRT-Protokoll des klinikinternen 3-Tesla-Scanners (Magnetom Skyra, Siemens).

## 3.3 Strukturierte qualitative Befundung (PI-RADS)

Die qualitative Befundung der MRT-Untersuchungen erfolgte im Rahmen der klinischen Routine durch radiologische Fachärzte. Dabei wurde die Prostata-MRT sämtlicher Befunde im klinischen Setting strukturiert nach PI-RADS 2.0 Kriterien beurteilt. Somit wurde für jede Läsion ein definitiver PI-RADS-Score bestimmt (s. 1.3.3).

## 3.4 Quantitative Analyse der DW-MRT

Die quantitative Analyse der MRT-Untersuchungen erfolgte über die Ermittlung der mittleren ADC-Werte, die aus DWI-Akquisitionen mit verschiedenen b-

Werten extrahiert wurden. Dabei wurde der mittlere ADC-Wert (mADC) aus den segmentierten Läsionen und in der gesamten Prostata bestimmt.

## **3.5 Radiomics-Analyse**

### **3.5.1 Segmentierung**

Die Segmentierung fokaler, tumorsuspekter Läsionen (PI-RADS  $\geq 3$ ) wurde manuell teils von zwei Fachärzten für Radiologie, teils von einer Assistenzärztin für das Fach Radiologie und der Verfasserin dieser Arbeit mit anschließender fachärztlicher Revision durchgeführt. Dabei wurde die Software MITK (The Medical Imaging Interaction Toolkit) (89)) verwendet. Die Segmentierungen wurden auf den T2w- und DWI-Sequenzen sowie ADC-Karten durchgeführt. In der T2-gewichteten Sequenz wurden sowohl die gesamte Prostata, die periphere Zone (PZ) und Transitionalzone (TZ) als auch die tumorsuspekten Läsionen (mit einem PI-RADS  $\geq 3$ ) in allen aufeinanderfolgenden axialen Schichten segmentiert. Lagen pro Patient mehrere tumorsuspekte Läsionen vor, wurde gemäß der Indexläsionstheorie (15) die Indexläsion als die Läsion mit dem höchsten PI-RADS-Score definiert und segmentiert. Auf Basis der Segmentierungen der einzelnen axialen Schichten wurde mit Hilfe der MITK-Software ein 3D-Modell des segmentierten Volumens errechnet. In den ADC-Karten wurden jeweils die gesamte Prostata sowie die korrespondierenden tumorsuspekten Läsionen segmentiert (s. Abb. 10). Wenn die Läsion in der MRT nicht sicher abgrenzbar war oder keine Läsion sichtbar war (häufig bei Prostatakarzinomen mit niedrigem Risiko (Gleason-Score  $\leq 6$ )), konnte die Läsion nicht segmentiert werden. Final lagen bei den eingeschlossenen 140 Prostatakarzinompatienten somit 140 segmentierte Indexläsionen vor.

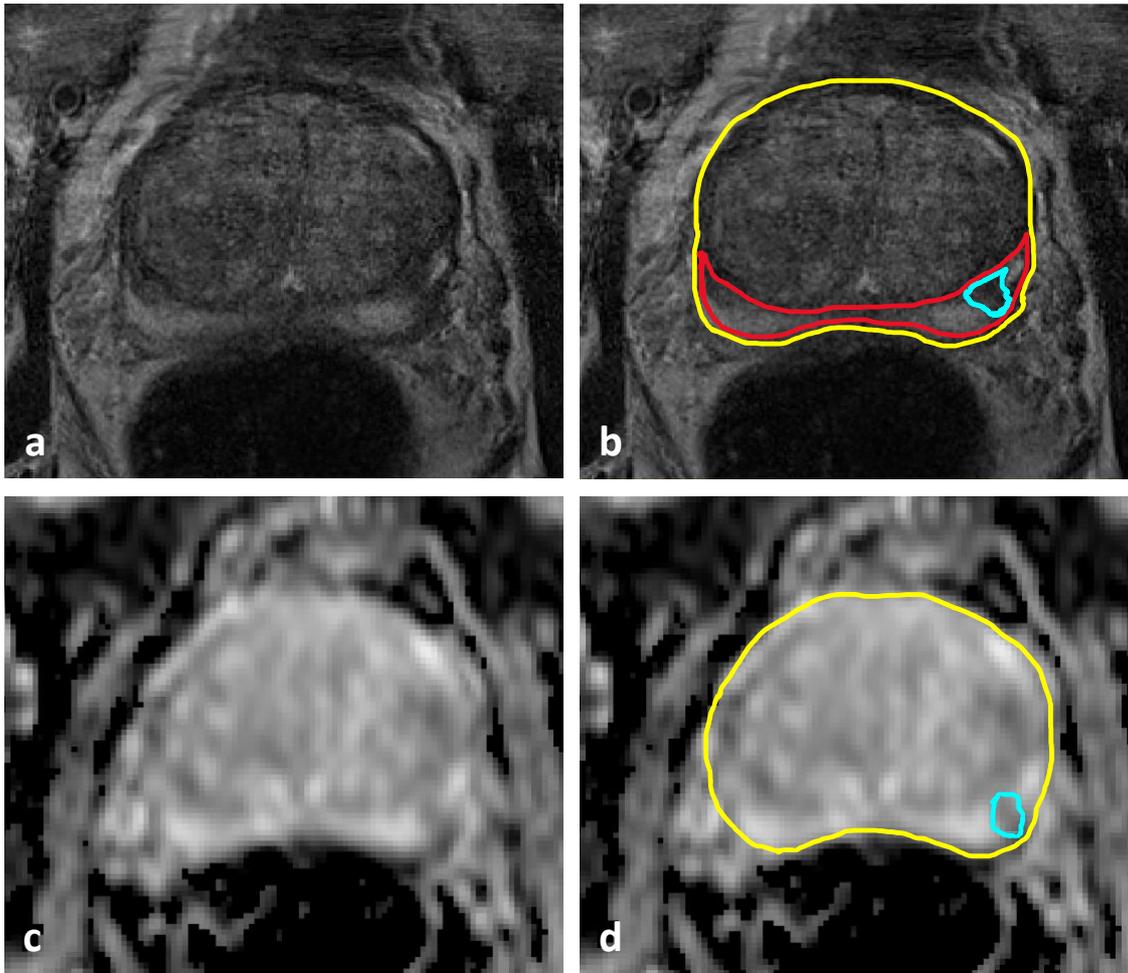


Abb. 10: Beispiel der Segmentierung eines Prostatakarzinoms bei einem 73-jährigen Patienten. PSA-Wert 12,0 ng/ml. Verdächtiger Herdbefund in der posteromedialen peripheren Zone links in der Drüsenmitte (PI-RADS 4, T3a).

(a) T2-Gewichtung: Signalthypointensität des tumorsuspekten Areals mit entsprechender Segmentierung in (b) (zyan) sowie Segmentierung der gesamten Prostata (gelb) und der peripheren Zone (rot). (c) ADC-Karte: Korrespondierende fokale Signalthypointensität des tumorsuspekten Herdbefundes. (d) Segmentierungen der gesamten Prostata (gelb) und der Läsion (zyan) in der ADC-Karte. Das postoperative Prostatektomiepräparat zeigte ein Prostatakarzinom mit einem Gleason-Score von 3+4.

### 3.5.2 Bildnachbearbeitung

Die T2-gewichteten Bilder wurden normiert, da die Intensitätswerte in dieser Sequenz nicht geräteunabhängig sind. Bei den ADC-Bildern musste keine Normierung durchgeführt werden, weil der ADC-Wert an sich ein unabhängiges, quantitatives Maß ist.

### 3.5.3 Feature-Berechnung

Aus den multiparametrischen Bilddaten zusammen mit den manuellen Segmentierungen wurden dann eine Reihe von quantitativen Tumoreigenschaften bestimmt. Dieser Prozess der Feature-Bestimmung wurde mit Hilfe der Open-Source-Software „PyRadiomics“ durchgeführt, die in der für wissenschaftliches Rechnen oft verwendeten Programmiersprache „Python“ implementiert ist. PyRadiomics wurde mit dem Ziel entwickelt, als Referenzstandard für radiomische Analysen zu dienen, um reproduzierbare Ergebnisse zu erhalten (90). Die einzelnen Schritte der Feature-Bestimmung waren wie folgt: Zuerst wurden die T2-Bilder, ADC-Bilder und Segmentierungen in die PyRadiomics-Plattform geladen. Dann wurden insgesamt acht in die Software bereits integrierte Filter auf die MRT-Bilder angewendet, darunter zum Beispiel der „Laplacian of Gaussian“- oder „Wavelet transform“-Filter, um einen breiteren Basisdatensatz für die Feature-Bestimmung zu haben. Die Features wurden jeweils aus den ADC- und T2-Bildern sowie aus den mit Filtern bearbeiteten MRT-Bildern in PyRadiomics berechnet. Dies erfolgte nach den Empfehlungen der Image Biomarker Standardisation Initiative (IBSI) (83). Die berechneten Features ließen sich in drei Parameterkategorien einteilen:

- Formparameter (n = 13 pro Bildkontrast)  
z.B. Volumen, Oberfläche, „Kugelförmigkeit“, Kompaktheit  
Weil sie unabhängig von den Intensitätswerten sind, werden diese als einzige Kategorie nur aus den ungefilterten Bildern berechnet.
- Intensitätsparameter (n = 18 x 8 Filter = 144 pro Bildkontrast)  
Sogenannte „first order features“, z.B. Mittelwert, Verteilungsbreite in der T2- und ADC-Gewichtung
- Texturparameter (n = 73 x 8 Filter = 584 pro Bildkontrast)  
Diese Matrizen beschreiben die Graustufenmuster benachbarter Voxel (s. Abb. 11)  
z.B.
  - „gray level cooccurrence matrix (GLCM)“ (91)
  - „gray level run length matrix (GLRLM)“ (92)
  - „gray level size zone matrix (GLSZM)“ (93)

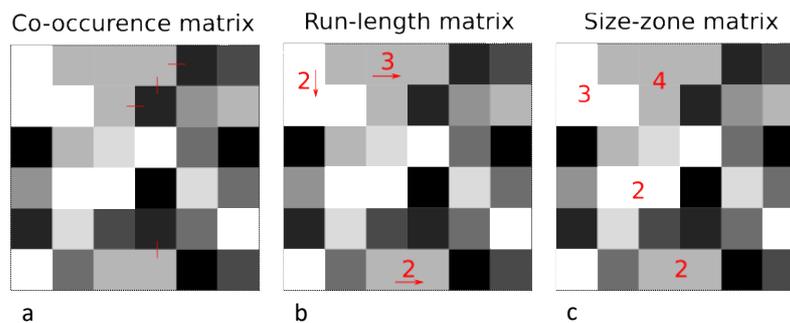


Abb. 11: Prinzip der Texturparameter.

(a) GLCM: Es werden Paare von nebeneinanderliegenden Voxel zweier Graustufen erkannt. (b) GLRLM: Es wird die Anzahl und Richtung in einer Reihe nebeneinanderliegender Voxel derselben Graustufe erkannt. (c) GLSZM: Es wird die Anzahl nebeneinanderliegender Voxel derselben Graustufe erkannt.

Somit wurden insgesamt 741 Features pro Patient pro Bildkontrast (T2-gewichtete und ADC-Sequenz) aus den MRT-Bilddatensätzen extrahiert, das entspricht 1482 Features pro Patient in toto.

### 3.5.4 Training der Klassifikatoren und Feature-Auswahl

Insgesamt wurden sieben Klassifikatoren trainiert: Ein Algorithmus basierend auf der logistischen Regression und sechs Random-Forest-Algorithmen, wobei davon zuerst ein „Basis“-Random-Forest-Algorithmus erstellt wurde und dieser danach variiert wurde, um zu sehen, welche Stellschraube einen relevanten Einfluss hat. Diese Stellschrauben werden auch Hyperparameter bezeichnet, deren Optimierung ein elementarer Teil einer möglichst präzisen Modellentwicklung ist (94).

Der logistische-Regressions-Algorithmus wurde reguliert, um die Interpretierbarkeit des Modells zu erhöhen und ein sogenanntes „Overfitting“ zu verhindern. Das meint, dass ein Modell zu genau Eigenschaften der Trainingsdaten lernt, die jedoch nicht Eigenschaften des gesamten Datensatzes beziehungsweise der gesamten Population sind. Dazu wurde die „elastic-net-Regularisierung“ angewandt, wobei die Features um 0 zentriert und auf eine Standardabweichung von 1 skaliert wurden.

Für das Training des „Basis“-Random-Forest-Algorithmus wurde zuvor eine Technik für unbalancierte Klassen angewandt („class weights“), wie es in dieser Studie der Fall war. Es erfolgte eine Gewichtung, die umgekehrt proportional zur Klassengröße war, also ausglich, dass Patienten mit klinisch signifikanten (93

Fälle) und Patienten mit klinisch nicht signifikanten oder keinen Prostatakarzinomen (47 Fälle) unterschiedlich häufig in dem Datensatz auftraten. Ebenfalls wurde das „Minimum Redundancy Maximum Relevance (mRMR)“-Schema (s. 1.3.5) angewandt, um aus der Vielzahl der zuvor beschriebenen extrahierten Features, die auch hochkorrelierte oder „noisy“ Features umfasste, die relevanten auszuwählen. Das Schema erstellte dann eine Rangliste dieser, aus welcher wiederum für diesen Algorithmus die besten 25 Features ausgewählt wurden. Auf den reduzierten Features wurde dann der Klassifikator trainiert, um die Dignität der Läsionen vorherzusagen. Dabei wurden die reduzierten Features der Patienten abhängig vom jeweiligen histologischen Ergebnis mit dem Label „klinisch signifikantes Prostatakarzinom positiv/ Gleason-Score  $\geq 7$ “ oder „klinisch signifikantes Prostatakarzinom negativ/ Gleason-Score  $\leq 6$ “ versehen.

Die restlichen fünf abgewandelten Random-Forest-Algorithmen (RF-Algorithmen) waren wie folgt:

- Ein RF-Algorithmus wie oben beschrieben mit „class weights“ aber auf den besten 50 Features trainiert
- Ein RF-Algorithmus wie oben beschrieben mit „class weights“ aber auf den besten 100 Features trainiert
- Ein RF-Algorithmus wie oben beschrieben nur ohne „class weights“
- Ein RF-Algorithmus wie oben beschrieben nur ohne Feature-Auswahl
- Ein RF-Algorithmus wie oben beschrieben nur statt „class weights“ mit einer anderen Technik für unbalancierte Datensätze, der überrepräsentierten Stichprobe („oversampling“). Dadurch wird die Anzahl der unterrepräsentierten Fälle angehoben und ein balancierter Datensatz geschaffen. SMOTE (synthetic minority sampling technique) ist eine Methode dafür, dessen genaue Funktionsweise aber über den Rahmen dieser Arbeit hinaus geht.

Für alle sieben Klassifikatoren folgte eine fünffache Kreuzvalidierung, wobei der Datensatz zufällig in fünf kleinere Datensätze geteilt wurde, und jeweils vier kleinere Datensätze für das Training sowie ein kleinerer Datensatz für das Testen verwendet wurden. Die Datensätze waren dabei für alle Klassifikatoren die gleichen. Die Genauigkeit der Klassifikatoren wurde anhand jeder Testrunde

evaluiert (fünf Ergebnisse pro Klassifikator) und quantitativ mit einer ROC(receiver operating characteristics)-Kurve beziehungsweise der AUC (area under curve) ausgewertet.

### **3.6 Statistische Analysen**

Alle statistischen Analysen wurden in R beziehungsweise in Python 3.7 durchgeführt. Das beinhaltete die Erhebung statistischer sowie klinischer Merkmale des Patientenkollektivs und die Berechnung der Genauigkeit der Histologieergebnisse der Biopsie versus der Prostatektomie.

Um die drei Methoden (PI-RADS-Score, mADC, Radiomics) hinsichtlich der Genauigkeit ihrer Klassifizierung von klinisch signifikanten versus klinisch nicht signifikanten Prostatakarzinomen zu evaluieren, wurde mit Hilfe des pROC-Pakets in R (95) für alle Methoden eine Receiver Operating Characteristic-Kurve (ROC-Kurve) erstellt und die Fläche unter der ROC-Kurve (AUC) berechnet. Ein klinisch signifikantes Prostatakarzinom wurde dabei entsprechend der Leitlinie und vorherigen Studien als ein Prostatakarzinom mit einem Gleason-Score von  $\geq 7$  (3+4) in mindestens einem Gewebezylinder definiert. Die Vergleichsgruppe der klinisch nicht signifikanten Prostatakarzinome bestand demnach aus Patienten mit einem risikoarmen Prostatakarzinom (Gleason-Score  $\leq 6$ ), sowie Patienten ohne Karzinomnachweis in der Histologie. Des Weiteren wurde bei der PI-RADS- und ADC-Analyse der optimale Cut-Off-Wert nach Youden berechnet, welcher als der Grenzwert definiert ist, der am weitesten von der Identitätslinie (Diagonale) entfernt ist (Optimalitätskriterium =  $\max(\text{Sensitivität} + \text{Spezifität})$ ). Bei der Radiomics-Analyse wurde die mittlere Fläche unter der ROC-Kurve der fünf Evaluierungsrunden im Rahmen der fünffachen Kreuzvalidierung berechnet.

## 4 Ergebnisse

### 4.1 Demographische und klinische Merkmale der Patientenstichprobe

In der Patientenstichprobe dieser Studie, die 140 Männer mit Prostataläsionen in der MRT umfasste, lag ein Medianalter von 68 (IQR 61–75) Jahren vor (s. Tabelle 5). Der mediane PSA-Wert betrug 7,5 (IQR 5,6–14,1) ng/ml. Das mediane Prostatavolumen betrug 41,3 (IQR 29,6–59,7). Die mediane PSA-Dichte betrug 0,21 (IQR 0,12–0,36) ng/ml/cm<sup>3</sup>.

Bei 19 der 140 Patienten (14%) lag nach histologischer Beurteilung kein Prostatakarzinom vor, bei 121 (86%) wurde das Vorliegen eines Prostatakarzinoms bestätigt. Von den 140 analysierten Läsionen waren 24 (17%) in der Transitionalzone lokalisiert, 116 (83%) hingegen in der peripheren Zone.

Bei 93 von 140 Patienten (66%) war ein klinisch signifikantes Prostatakarzinom vorhanden, bei 47 Patienten (34%) war ein klinisch nicht signifikantes Prostatakarzinom oder kein Prostatakarzinom vorhanden. In diese beiden Gruppen sollten die Patienten durch die drei Methoden PI-RADS-Score, mittlerer ADC-Wert und Radiomics klassifiziert werden.

Unter den 140 Patienten betrug der mediane mADC-Wert der gesamten Prostata 1197 (IQR 1129–1273) mm/sec<sup>2</sup>. Dabei betrug der mediane mADC-Wert aller eingeschlossenen Läsionen 851 (IQR 763–982) mm/sec<sup>2</sup>, der mediane mADC-Wert der Läsionen nur in der peripheren Zone 864 (IQR 769–1002) mm/sec<sup>2</sup> beziehungsweise 825 (IQR 715–905) mm/sec<sup>2</sup> nur in der Transitionalzone.

| Variable   | Alle Patienten <sup>1,2</sup> | Patienten mit GS ≤ 6 / kein PCA <sup>1,2</sup> | Patienten mit GS ≥ 7 <sup>1,2</sup> |
|--|-------------------------------|--|-------------------------------------|
| Medianalter  | 68 [61 – 75]                  | 68 [61 – 75]                                   | 68 [62 – 75]                        |
| Median-PSA-Wert <sup>3</sup> (ng/ml)                                       | 7,5 [5,6 – 14,1]              | 7,5 [5,6 – 13,9]                               | 8,0 [5,9 – 14,6]                    |
| Patienten mit PSA-Wert <sup>3</sup> > 10 (ng/ml)                           | 51/136 (36) <sup>5</sup>      | 10/43 (23)                                     | 41/93 (44)                          |
| Patienten mit PSA-Wert <sup>3</sup> ≤ 10 (ng/ml)                           | 85/136 (63) <sup>5</sup>      | 33/43 (77)                                     | 52/93 (56)                          |
| Medianprostatavolumen <sup>3,4</sup> (ml)                                  | 41,3 [29,6 – 59,7]            | 43,1 [33,6 – 70,3]                             | 39,0 [27,4 – 54,1]                  |
| Median-PSA-Dichte <sup>3,4</sup> (ng/ml/cm <sup>3</sup> )                  | 0,21 [0,12 – 0,36]            | 0,14 [0,08 – 0,22]                             | 0,25 [0,14 – 0,43]                  |
| Patienten mit PSA-Dichte <sup>3,4</sup> > 0,15 (ng/ml/cm <sup>3</sup> )    | 88/134 (66)                   | 21/43 (49)                                     | 67/91 (74)                          |
| Patienten mit PSA-Dichte <sup>3,4</sup> ≤ 0,15 (ng/ml/cm <sup>3</sup> )    | 46/134 (34)                   | 22/43 (51)                                     | 24/91 (26)                          |
| Eingeschlossene Patienten  | 140                           | 47   | 93                                  |
| - davon PI-RADS 2-Läsion   | 4/140 (3)                     | 3/47 (6) <sup>5</sup>                          | 1/93 (1) <sup>5</sup>               |
| - davon PI-RADS 3-Läsion   | 37/140 (26)                   | 26/47 (55) <sup>5</sup>                        | 12/93 (13) <sup>5</sup>             |
| - davon PI-RADS 4-Läsion   | 28/140 (20)                   | 8/47 (17) <sup>5</sup>                         | 20/93 (22) <sup>5</sup>             |
| - davon PI-RADS 5-Läsion   | 71/140 (51)                   | 10/47 (21) <sup>5</sup>                        | 60/93 (65) <sup>5</sup>             |
| - davon Läsion in der TZ   | 24/140 (17)                   | 10/47 (21)                                     | 14/93 (15)                          |
| - davon Läsion in der PZ   | 116/140 (83)                  | 37/47 (79)                                     | 79/93 (85)                          |
| Patienten ohne PCA   | 19/140 (14)                   | 19/47 (40)                                     | 0/93 (0)                            |
| Patienten mit PCA  | 121/140 (86)                  | 28/47 (60)                                     | 93/93 (100)                         |
| - davon Gleason-Score 6 (3+3)  | 28/121 (23)                   | 28/47 (60)                                     | 0/93 (0)                            |
| - davon Gleason-Score 7a (3+4)   | 37/121 (31)                   | 0/47 (0)                                       | 37/93 (40)                          |
| - davon Gleason-Score 7b (4+3)   | 24/121 (20)                   | 0/47 (0)                                       | 25/93 (27)                          |
| - davon Gleason-Score 8 (4+4)  | 12/121 (10)                   | 0/47 (0)                                       | 11/93 (12)                          |
| - davon Gleason-Score 9a (4+5)   | 17/121 (14)                   | 0/47 (0)                                       | 17/93 (18)                          |
| - davon Gleason-Score 9b (5+4)   | 3/121 (2)                     | 0/47 (0)                                       | 3/93 (3)                            |
| Patienten mit ksPCA (Gleason-Score ≥ 7)                                    | 93/140 (66)                   | 0/47 (0)                                       | 93/93 (100)                         |
| Patienten mit knsPCA (Gleason-Score ≤ 6) oder keinem PCA                   | 47/140 (34)                   | 47/47 (100)                                    | 0/93 (0)                            |
| Patienten mit knsPCA (Gleason-Score ≤ 6)                                   | 28/140 (20)                   | 28/47 (60)                                     | 0/93 (0)                            |
| Median-mADC-Wert der gesamten Prostata <sup>4</sup> (mm/sec <sup>2</sup> ) | 1197 [1129 – 1273]            | 1231 [1151 – 1303]                             | 1185 [1108 – 1259]                  |
| Median-mADC-Wert aller Läsionen (mm/sec <sup>2</sup> )                     | 851 [763 – 982]               | 1021 [799 – 1133]                              | 831 [747 – 900]                     |
| - der Läsionen in der PZ (mm/sec <sup>2</sup> )                            | 864 [769 – 1002]              | 1039 [828 – 1159]                              | 832 [756 – 904]                     |
| - der Läsionen in der TZ (mm/sec <sup>2</sup> )                            | 825 [715 – 905]               | 841 [776 – 1035]                               | 810 [615 – 873]                     |
| Patienten mit Biopsie  | 134/140 (96)                  | 35/47 (74)                                     | 84/93 (90)                          |
| Patienten mit Prostatektomie   | 87/140 (62)                   | 14/47 (30)                                     | 77/93 (83)                          |

Tabelle 5: Merkmale der eingeschlossenen Patienten (n = 140).

PSA = Prostata-spezifisches Antigen, PI-RADS = Prostate Imaging Reporting and Data System, TZ = Transitionalzone, PZ = periphere Zone, PCA = Prostatakarzinom, ksPCA = klinisch signifikantes Prostatakarzinom, knsPCA = klinisch nicht signifikantes Prostatakarzinom, mADC = mean apparent diffusion coefficient. <sup>1</sup> Werte in eckigen Klammern sind Interquartilsabstände. <sup>2</sup> Werte in runden Klammern sind Prozente. <sup>3</sup> Bei 4 Patienten fehlte der PSA-Wert. <sup>4</sup> Bei 2 Patienten fehlte die Segmentierung der gesamten Prostata. <sup>5</sup> Durch Rundungen summieren sich diese Prozente nicht auf 100.

## 4.2 Beispielfälle

### 4.2.1 Bioptisch gesichertes klinisch signifikantes Prostatakarzinom

#### Fall 1

Die mittlere-ADC-Wert- und Radiomics-Methode erkannten ein von der PI-RADS-Methode nicht detektiertes, klinisch signifikantes Prostatakarzinom (Gleason-Score 3+4).

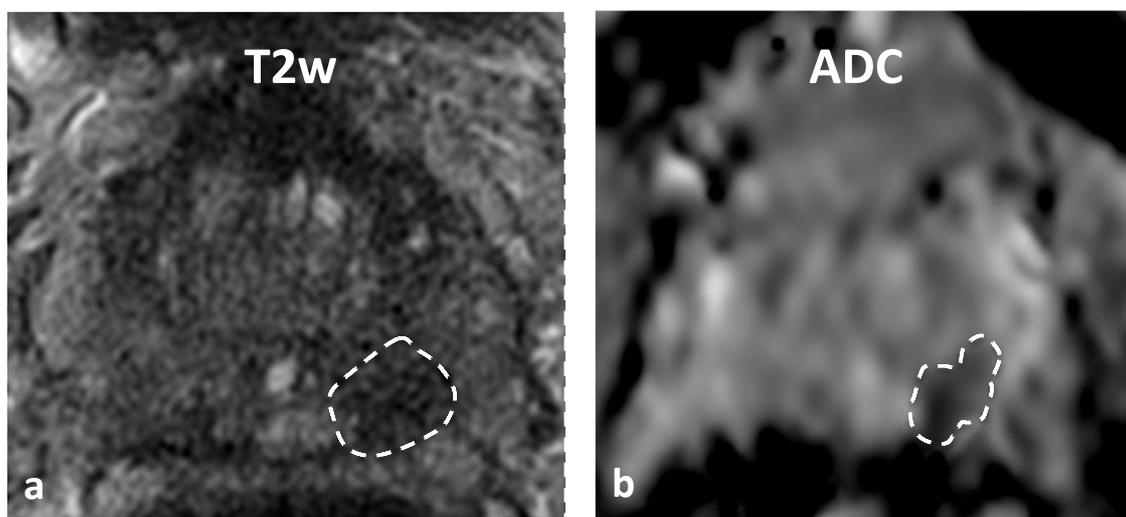


Abb. 12: 60-jähriger Patient mit einem PSA-Wert von 5,9 ng/ml.

Die axiale T2w (a) zeigt eine flau fokale T2w-Signalabsenkung in der posteromedialen peripheren Zone links in Drüsenmitte (Umrandung: T2w Score von 3) mit deutlicher fokaler ADC-Absenkung (b), die korrespondierende DWI (b = 1200; Bild nicht gezeigt) zeigte keine relevante fokale Signalhyperintensität (gemäß DWI Score von 3), somit PI-RADS (2.0) Gesamtscore von 3. Entgegen der Bewertung nach PI-RADS wurde die Läsion mit einem mADC-Wert von 882,9 mm<sup>2</sup>/sec als klinisch signifikant bewertet (Cut-off: 1016,5 mm<sup>2</sup>/sec). Auch der radiomische „Basis“-Random-Forest-Algorithmus schätzte die Läsion mit einem Score für das Vorhandensein eines klinisch signifikanten Prostatakarzinoms von 0,90 als klinisch signifikant ein. Das Prostatektomiepräparat zeigte ein klinisch signifikantes Prostatakarzinom mit einem Gleason-Score von 3+4.

## Fall 2

Nur die PI-RADS-Methode erkannte ein von der mittleren-ADC-Wert- und Radiomics-Methode nicht detektiertes, klinisch signifikantes Prostatakarzinom (Gleason-Score 3+4).

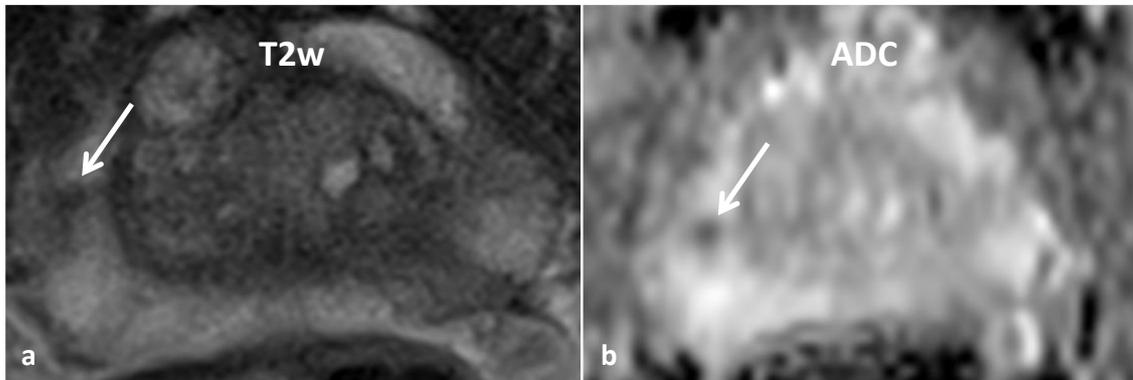


Abb. 13: 73-jähriger Patient mit einem PSA-Wert von 4,6 ng/ml.

Die axiale T2w (a) zeigt eine ca. 0,7 cm messende, T2w-hypointense fokale Läsion (T2w Score von 4) mit deutlicher fokaler ADC-Absenkung (b), einem DWI Score von 4 entsprechend. Es ergibt sich somit ein PI-RADS (2.0) Gesamtscore von 4. Entgegen der Bewertung nach PI-RADS wurde die Läsion mit einem mADC-Wert von 1305,5 mm<sup>2</sup>/sec als klinisch nicht signifikant bewertet (Cut-off: 1016,5 mm<sup>2</sup>/sec). Der radiomische „Basis“-Random-Forest-Algorithmus schätzte die Läsion ebenfalls mit einem Score für das Vorhandensein eines klinisch signifikanten Prostatakarzinoms von 0,29 als klinisch nicht signifikant ein. Das Prostatektomiepräparat zeigte ein klinisch signifikantes Prostatakarzinom mit einem Gleason-Score von 3+4.

## 4.2.2 Bioptisch gesichertes klinisch nicht signifikantes Prostatakarzinom beziehungsweise kein Karzinomnachweis in der Histologie

### Fall 3

Die mittlere-ADC-Wert- und Radiomics-Methode erkannten im Gegenteil zur PI-RADS-Methode, dass kein klinisch signifikantes Prostatakarzinom vorlag (Gleason-Score 3+3).

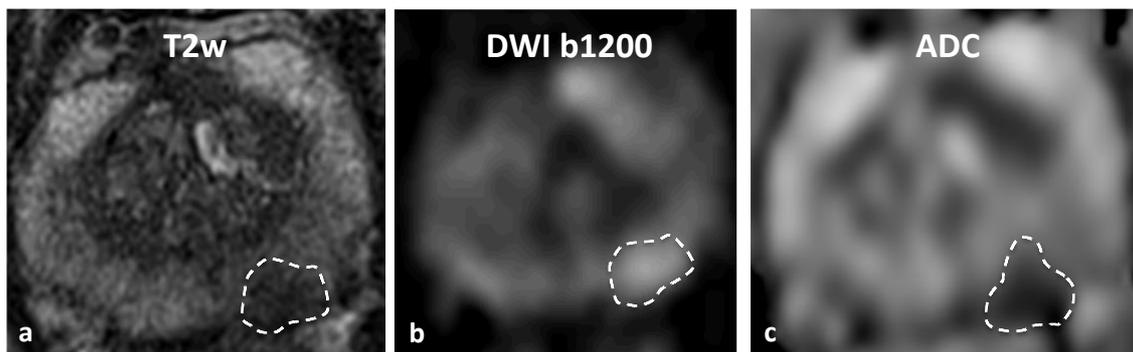


Abb. 14: 60-jähriger Patient mit einem PSA-Wert von 3,6 ng/ml.

(a) T2w-hypointense fokale Läsion in der posteromedialen peripheren Zone links in Drüsenmitte (Pfeil), T2w-Score von 4. (b, c) Die Läsion zeigt eine deutliche fokale Diffusionsstörung mit hyperintensem Signal in den hohen b1200er Werten und entsprechender Signalabsenkung in der ADC-Map (Pfeile in b, c), die Läsion misst ca. 1,2 cm, einem DWI-Score von 4 entsprechend. Es ergibt sich somit ein PI-RADS Gesamtscore (2.0) von 4. Entgegen der Bewertung nach PI-RADS wurde die Läsion mit einem mADC-Wert von 1039,1 mm<sup>2</sup>/sec als klinisch nicht signifikant bewertet (Cut-off: 1016,5 mm<sup>2</sup>/sec). Auch der radiomische „Basis“-Random-Forest-Algorithmus schätzte die Läsion mit einem Score für das Vorhandensein eines klinisch signifikanten Prostatakarzinoms von 0,35 als klinisch nicht signifikant ein. Das Prostatektomiepräparat zeigte ein klinisch nicht signifikantes Prostatakarzinom mit einem Gleason-Score von 3+3.

#### Fall 4

Keine der Methoden (Radiomics, PI-RADS oder mittlerer ADC-Wert) erkannte, dass kein Prostatakarzinom vorlag.

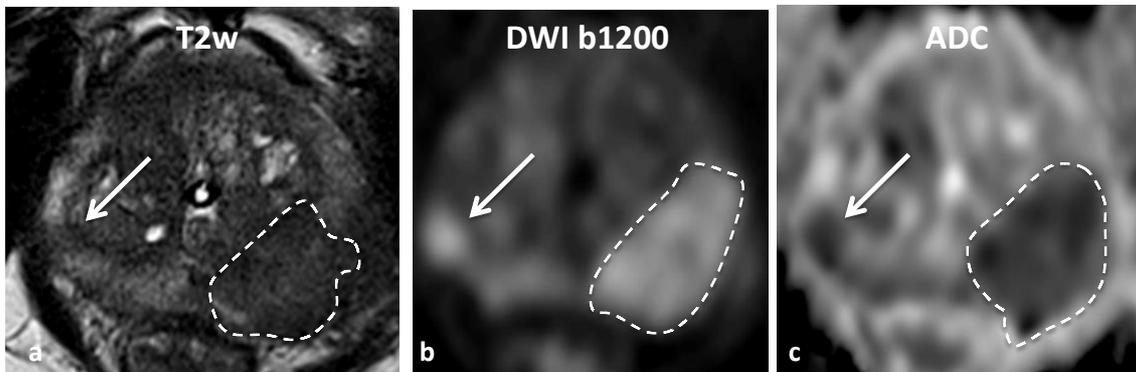


Abb. 15: 66-jähriger Patient mit einem PSA-Wert von 17,1 ng/ml und klinischem Verdacht auf ein Prostatakarzinom.

In der axialen T2w zeigt sich links paramedian in der posteromedialen und posterolateralen peripheren Zone eine moderat T2w-hypointense Läsion > 2,5 cm mit Kapselbulging (gestrichelte Umrandung) mit ausgeprägter Diffusionsstörung (b, c), PI-RADS 5 entsprechend. Nebenbefundlich zeigt sich eine kleine, 1,0 cm messende moderat T2w-hypointense und diffusionsrestringierte Läsion in der posterolateralen peripheren Zone rechts in Drüsenmitte (Pfeile in a – c), PI-RADS 4. Hier ergibt sich bildmorphologisch der Verdacht auf ein multifokales Prostatakarzinom. Ebenso wurde die Läsion mit einem mADC-Wert von 783,6 mm<sup>2</sup>/sec als klinisch signifikant bewertet (Cut-off: 1016,5 mm<sup>2</sup>/sec). Auch der radiomische „Basis“-Random-Forest-Algorithmus schätzte die Läsion mit einem Score für das Vorhandensein eines klinisch signifikanten Prostatakarzinoms von 0,98 als klinisch signifikant ein. Die MRT-Ultraschall-Fusionsbiopsie dieser karzinomverdächtigen Läsionen ergab keinen Anhalt für ein Prostatakarzinom sondern Zeichen einer granulomatösen Prostatitis.

## 4.3 ROC-Analysen

### 4.3.1 Strukturierte qualitative Befundung (PI-RADS)

Die ROC(receiver operating characteristics)-Kurve der strukturierten qualitativen Befundung nach PI-RADS ist in Abb. 16 dargestellt. Die AUC (area under curve) betrug 0,78 (95%-Konfidenzintervall = [0,70 – 0,86]). Das bedeutet, dass mit einer Wahrscheinlichkeit von 78% ein zufälliges Paar eines auffälligen und eines nicht auffälligen PI-RADS-Scores korrekt als klinisch signifikantes (Gleason-Score  $\geq 7$ ) oder nicht signifikantes Prostatakarzinom (Gleason-Score  $< 6$ ) klassifiziert werden konnte (96). Das 95%-Konfidenzintervall gibt dabei jenes Intervall an, welches 95 Mal den tatsächlichen Wert der AUC beinhalten würde, würde diese Prozedur 100 Mal wiederholt werden (97). Der optimale Cut-off-Wert nach Youden wurde bei einem PI-RADS-Score von 4 (Sensitivität 0,87, Spezifität 0,62) erreicht und entspricht dem Punkt, der am weitesten von der Identitätslinie (Diagonale) entfernt ist. Das bedeutet, dass nach Youden ein PI-RADS-Score von 4 den besten Schwellenwert darstellt, um klinisch signifikante von klinisch nicht signifikanten Prostatakarzinomen zu unterscheiden. Hierbei liegt bei einem PI-RADS-Score  $\geq 4$  mit einer Sensitivität von 87% sowie einer Spezifität von 62% ein klinisch signifikantes Prostatakarzinom vor.

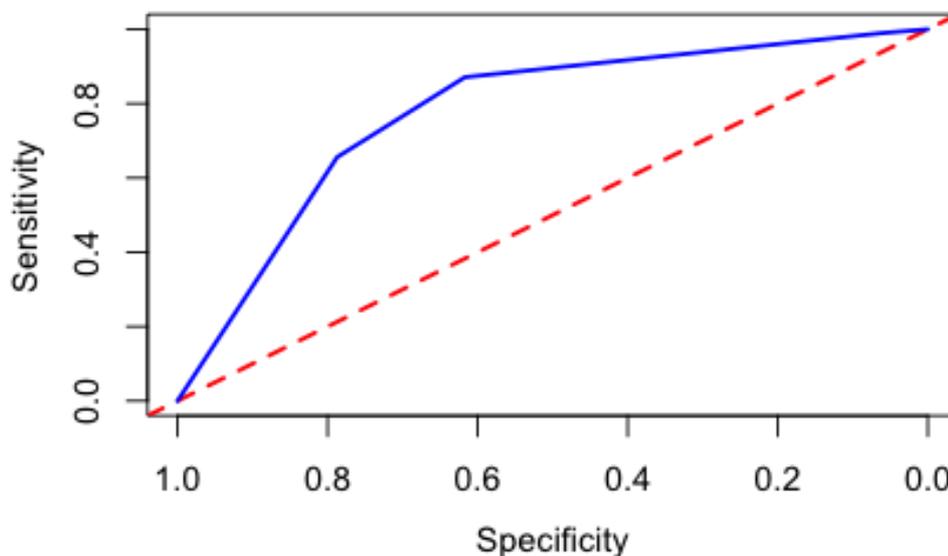


Abb. 16: ROC-Kurve der Befundung nach PI-RADS.

AUC = 0,78 (95%KI = [0,70 – 0,86]). Bester Cut-off-Wert bei einem PI-RADS-Score von 4. (Darstellung mit dem pROC-Paket in R)

### 4.3.2 Quantitative Analyse der DW-MRT (mADC)

Die ROC(receiver operating characteristics)-Kurve der quantitativen Analyse anhand des mittleren ADC-Wertes ist in Abb. 17 dargestellt. Die AUC (area under curve) betrug 0,72 (95%-Konfidenzintervall = [0,61 – 0,82]). Das bedeutet, dass mit einer Wahrscheinlichkeit von 72% ein zufälliges Paar eines auffälligen und eines nicht auffälligen mADC-Wertes korrekt als klinisch signifikantes (Gleason-Score  $\geq 7$ ) oder nicht signifikantes Prostatakarzinom (Gleason-Score  $< 6$ ) klassifiziert werden konnte (96). Das 95%-Konfidenzintervall gibt dabei jenes Intervall an, welches 95 Mal den tatsächlichen Wert des mADC beinhalten würde, würde diese Prozedur 100 Mal wiederholt werden (97). Der optimale Cut-off-Wert nach Youden wurde bei einem mADC-Wert von 1016,5 (Sensitivität 0,95, Spezifität 0,51) erreicht und entspricht dem Punkt, der am weitesten von der Identitätslinie (Diagonale) entfernt ist. Das bedeutet, dass nach Youden ein mADC von 1016,5 den besten Schwellenwert darstellt, um klinisch signifikante von klinisch nicht signifikanten Prostatakarzinomen zu unterscheiden. Hierbei liegt bei einem mADC  $\geq 1016,5$  mit einer Sensitivität von 95% sowie einer Spezifität von 51% ein klinisch signifikantes Prostatakarzinom vor.

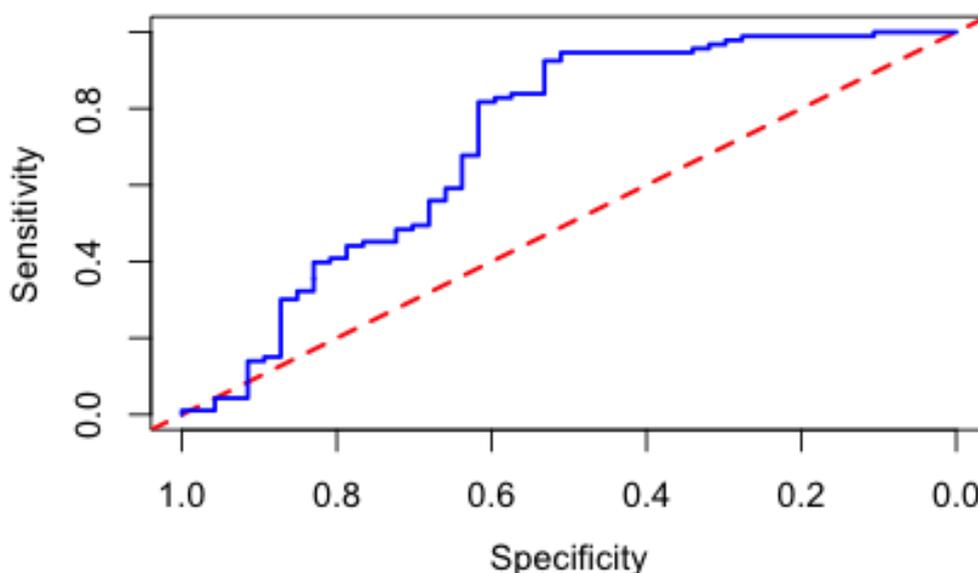


Abb. 17: ROC-Kurve der quantitativen Analyse anhand des mittleren ADC-Wertes.

AUC = 0,72 (95%KI = [0,61– 0,82]). Bester Cut-off-Wert bei einem mittleren ADC-Wert von 1016,5. (Darstellung mit dem pROC-Paket in R)

### 4.3.3 Radiomics-Analyse

Die ROC(receiver operating characteristics)-Kurven der einzelnen Algorithmen sind in den folgenden Abbildungen dargestellt. Dabei ergab sich jede der fünf Kurven durch die Evaluation auf dem entsprechenden „testing“-Fold der Kreuzvalidierung. Die AUC-Werte entsprechen jeweils der Wahrscheinlichkeit, dass ein zufälliges Paar eines auffälligen und eines nicht auffälligen Radiomics-Scores korrekt als klinisch signifikantes (Gleason-Score  $\geq 7$ ) oder nicht signifikantes Prostatakarzinom (Gleason-Score  $< 6$ ) klassifiziert werden konnte (96).

Abb. 18 zeigt die fünf ROC-Kurven des auf der logistischen Regression basierenden Algorithmus. Die mittlere AUC (mean area under curve) betrug 0,75.

Abb. 19 zeigt die fünf ROC-Kurven des „Basis“-Random-Forest-Algorithmus („class weights“, mRMR-Feature-Auswahl mit 25 Features). Die mittlere AUC (mean area under curve) betrug 0,82.

Abb. 20 zeigt die fünf ROC-Kurven des Random-Forest-Algorithmus mit „class weights“ und mRMR-Feature-Auswahl mit 50 Features. Die mittlere AUC (mean area under curve) betrug 0,84.

Abb. 21 zeigt die fünf ROC-Kurven des Random-Forest-Algorithmus mit „class weights“ und mRMR-Feature-Auswahl mit 100 Features. Die mittlere AUC (mean area under curve) betrug 0,84.

Abb. 22 zeigt die fünf ROC-Kurven des Random-Forest-Algorithmus ohne „class weights“ und mRMR-Feature-Auswahl mit 25 Features. Die mittlere AUC (mean area under curve) betrug 0,83.

Abb. 23 zeigt die fünf ROC-Kurven des Random-Forest-Algorithmus mit „class weights“ ohne mRMR-Feature-Auswahl. Die mittlere AUC (mean area under curve) betrug 0,82.

Abb. 24 zeigt die fünf ROC-Kurven des Random-Forest-Algorithmus mit SMOTE (Technik der überrepräsentierten Stichprobe für unbalancierte Datensätze) statt „class weights“ und mRMR-Feature-Auswahl mit 25 Features. Die mittlere AUC (mean area under curve) betrug 0,80.

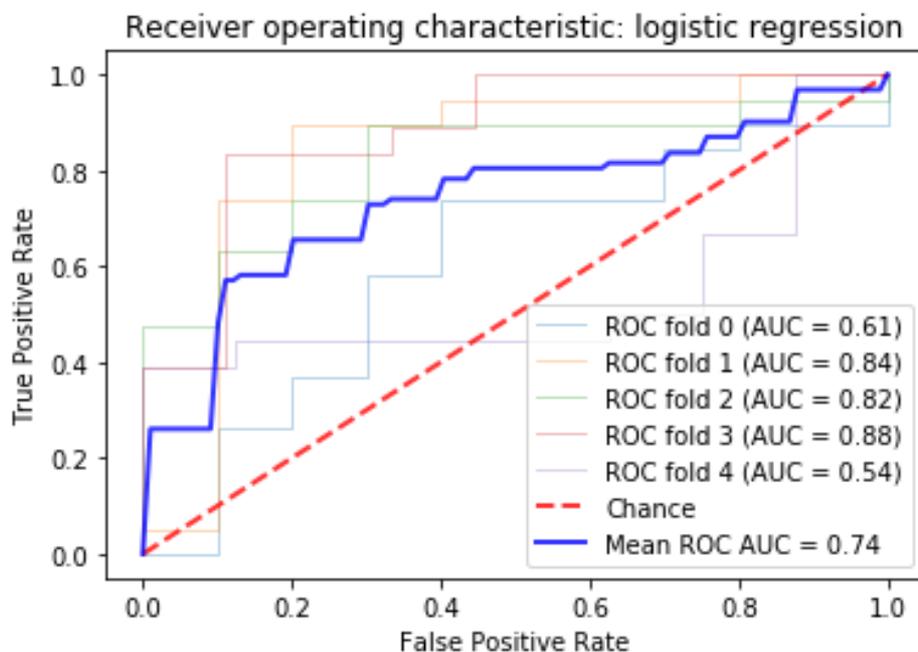


Abb. 18: ROC-Kurven der Radiomics-Analyse: logistische Regression.

Jede der fünf ROC-Kurven entsprach einer Evaluierungsrunde (fold 0 – 4) im Rahmen der fünffachen Kreuzvalidierung. Der Mittelwert aller fünf Flächen unter der Kurve (mean ROC AUC) betrug 0,75.

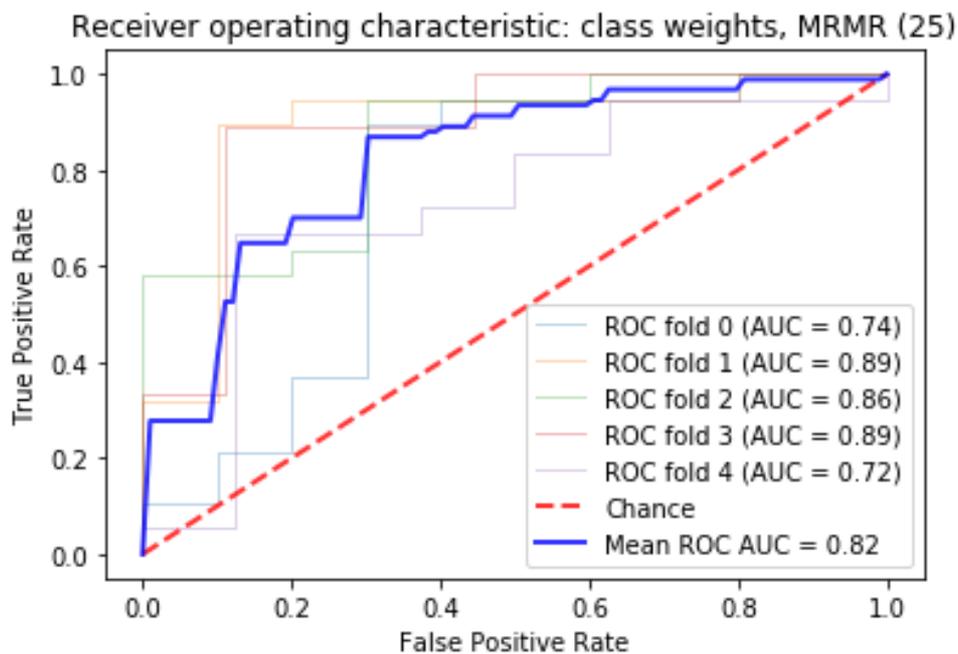


Abb. 19: ROC-Kurven der Radiomics-Analyse: „Basis“-Random-Forest-Algorithmus („class weights“, mRMR-Feature-Auswahl mit 25 Features).

Jede der fünf ROC-Kurven entsprach einer Evaluierungsrunde (fold 0 – 4) im Rahmen der fünffachen Kreuzvalidierung. Der Mittelwert aller fünf Flächen unter der Kurve (mean ROC AUC) betrug 0,82.

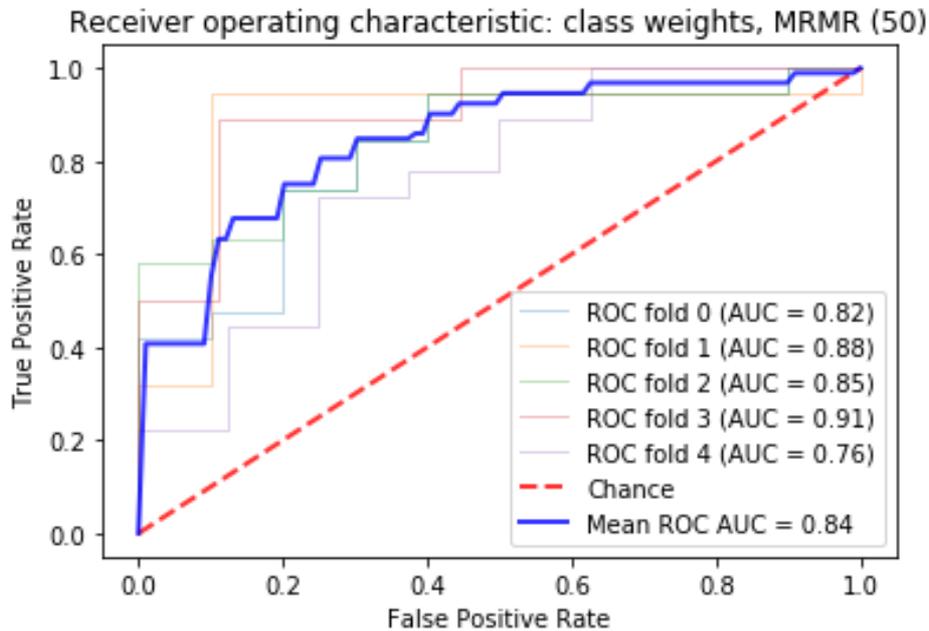


Abb. 20: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus mit „class weights“ und mRMR-Feature-Auswahl mit 50 Features.

Jede der fünf ROC-Kurven entsprach einer Evaluierungsrunde (fold 0 – 4) im Rahmen der fünffachen Kreuzvalidierung. Der Mittelwert aller fünf Flächen unter der Kurve (mean ROC AUC) betrug 0,84.

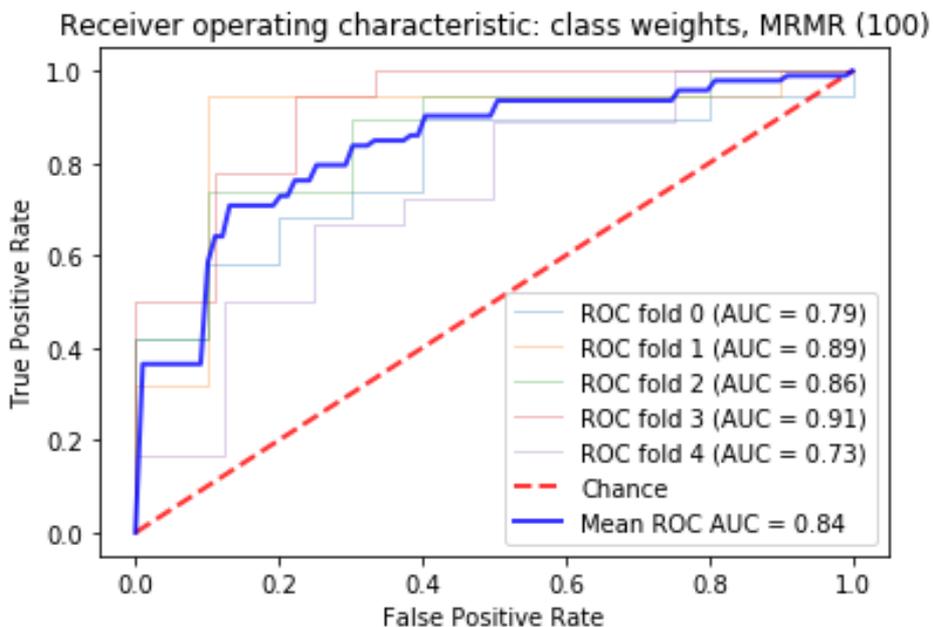


Abb. 21: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus mit „class weights“ und mRMR-Feature-Auswahl mit 100 Features.

Jede der fünf ROC-Kurven entsprach einer Evaluierungsrunde (fold 0 – 4) im Rahmen der fünffachen Kreuzvalidierung. Der Mittelwert aller fünf Flächen unter der Kurve (mean ROC AUC) betrug 0,84.

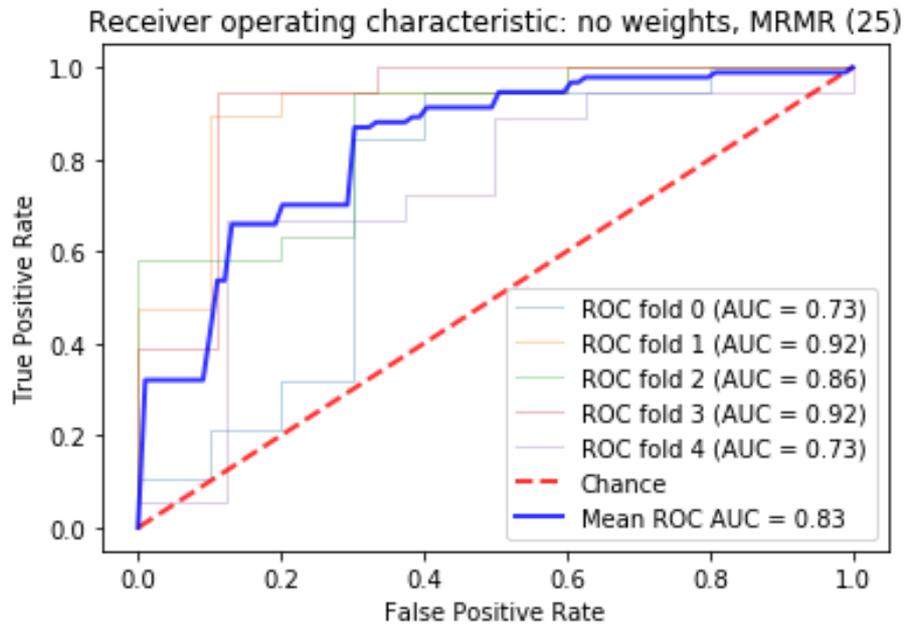


Abb. 22: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus ohne „class weights“ und mRMR-Feature-Auswahl mit 25 Features.

Jede der fünf ROC-Kurven entsprach einer Evaluierungsrunde (fold 0 – 4) im Rahmen der fünffachen Kreuzvalidierung. Der Mittelwert aller fünf Flächen unter der Kurve (mean ROC AUC) betrug 0,83.

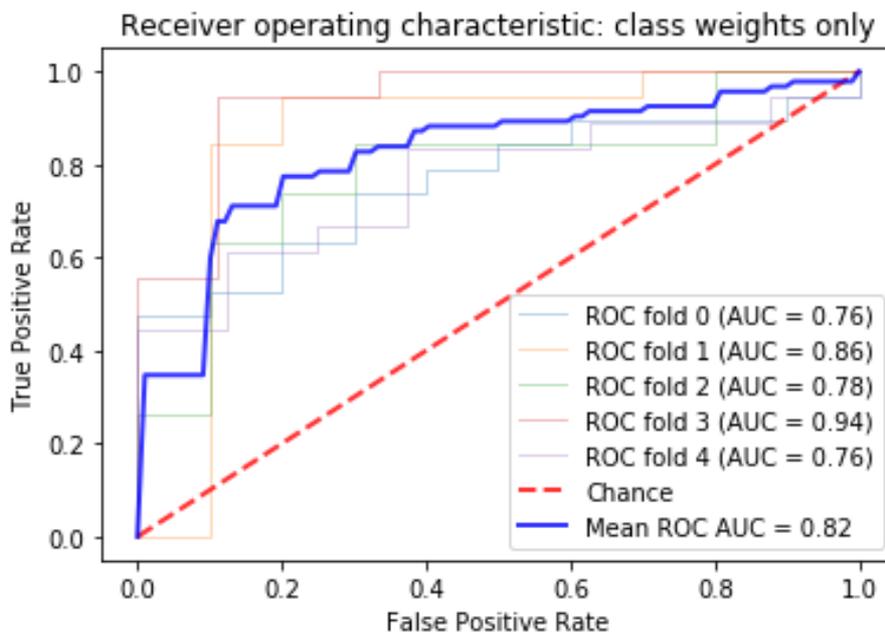


Abb. 23: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus mit „class weights“ ohne mRMR-Feature-Auswahl.

Jede der fünf ROC-Kurven entsprach einer Evaluierungsrunde (fold 0 – 4) im Rahmen der fünffachen Kreuzvalidierung. Der Mittelwert aller fünf Flächen unter der Kurve (mean ROC AUC) betrug 0,82.

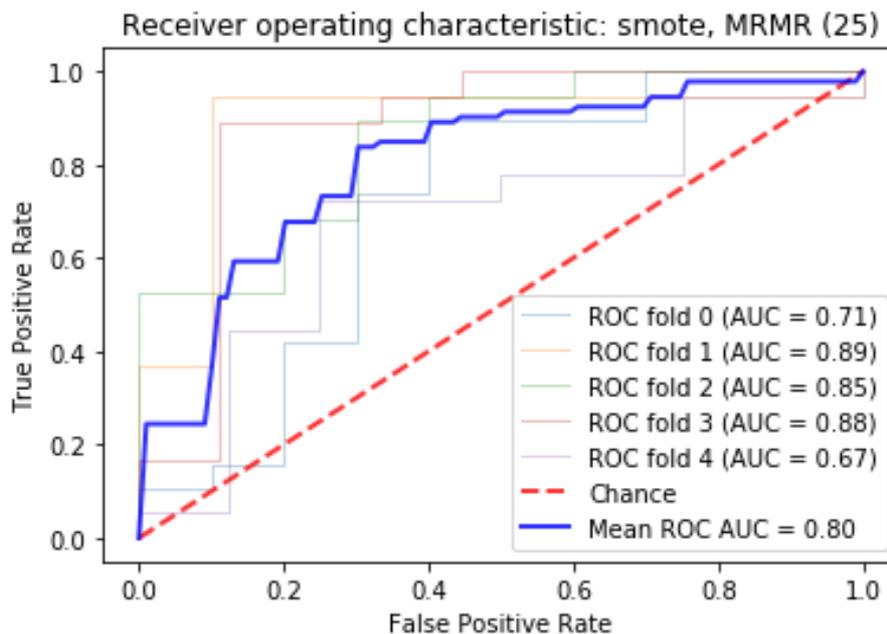


Abb. 24: ROC-Kurven der Radiomics-Analyse: Random-Forest-Algorithmus mit SMOTE statt „class weights“ und mRMR-Feature-Auswahl mit 25 Features.

Jede der fünf ROC-Kurven entsprach einer Evaluierungsrunde (fold 0 – 4) im Rahmen der fünffachen Kreuzvalidierung. Der Mittelwert aller fünf Flächen unter der Kurve (mean ROC AUC) betrug 0,80.

#### 4.3.4 Vergleich PI-RADS, mADC und Radiomics

Um die Ergebnisse der PI-RADS-, mADC- und Radiomics-Methode besser vergleichen zu können wurden ein Scatterplot erstellt, der die verschiedenen AUC-Werte der Methoden hinsichtlich ihrer Differenzierungsfähigkeit von klinisch signifikanten versus klinisch nicht signifikanten Prostatakarzinomen visuell in Bezug setzt (s. Abb. 25).

Eine Tabelle bietet weiters noch einmal eine gesammelte Übersicht über die AUC-Werte der einzelnen Methoden (s. Tabelle 6).

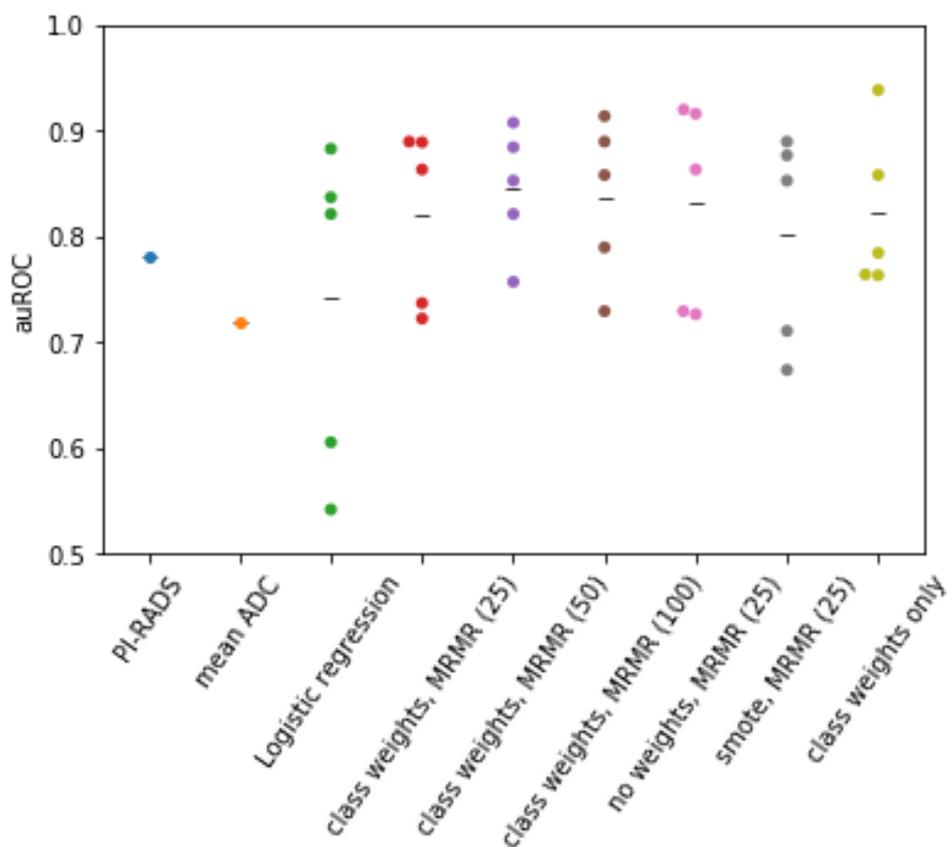


Abb. 25: Scatterplot zur Übersicht über die AUC-Werte aller Methoden hinsichtlich ihrer Differenzierungsfähigkeit von klinisch signifikanten versus klinisch nicht signifikanten Prostatakarzinomen.

In Farbe sind die Werte der Fläche unter der Kurve (AUC) der verschiedenen Methoden dargestellt. Bei den radiomischen Algorithmen sind es entsprechend den fünf Validierungsteilmengen jeweils fünf, der schwarze Strich zeigt jeweils den Mittelwert der fünf AUC-Werte an. auROC = area under receiver operating characteristics, PI-RADS = prostate imaging reporting and data system, ADC = apparent diffusion coefficient, mRMR = minimum redundancy maximum relevance, () = Anzahl der verwendeten Features, smote = synthetic minority sampling technique.

| Method   | AUC   |
|--|-------|
| PI-RADS  | 0,78  |
| Mean ADC   | 0,72  |
| Logistische-Regression-Algorithmus   | 0,75* |
| „Basis“-Random-Forest-Algorithmus<br>(class weights, mRMR (25))              | 0,82* |
| Random-Forest-Algorithmus<br>(class weights, mRMR (50))                      | 0,84* |
| Random-Forest-Algorithmus<br>(class weights, mRMR (100))                     | 0,84* |
| Random-Forest-Algorithmus<br>(no class weights, mRMR (25))                   | 0,83* |
| Random-Forest-Algorithmus<br>(smote statt class weights, mRMR (25))          | 0,80* |
| Random-Forest-Algorithmus<br>(nur class weights, keine mRMR-Feature-Auswahl) | 0,82* |

Tabelle 6: Übersicht über die AUC-Werte aller Methoden hinsichtlich ihrer Differenzierungsfähigkeit von klinisch signifikanten versus klinisch nicht signifikanten Prostatakarzinomen.

\* Mittelwerte der Flächen unter der Kurve der fünf Evaluierungsrunden im Rahmen der fünffachen Kreuzvalidierung (mean AUC)

## 4.4 Zusammenhänge zwischen verschiedenen Variablen und klinisch signifikanten Prostatakarzinomen

### 4.4.1 PSA-Wert und klinisch signifikante Prostatakarzinome

Der Zusammenhang zwischen PSA-Wert und klinisch signifikanten Prostatakarzinomen ist in Abb. 26 veranschaulicht. Weil es sich um eine kontinuierliche und eine diskrete Variable handelt, wurde ein Boxplot verwendet. Dieser Plot stellt dar, dass der mediane PSA-Wert der klinisch signifikanten Prostatakarzinome (Gleason-Score  $\geq 7$ ) höher ist als der der klinisch nicht signifikanten Prostatakarzinome (Gleason-Score  $\leq 6$ ). Ebenso lässt sich erkennen, dass die PSA-Werte der klinisch signifikanten Prostatakarzinome einer sichtlich breiteren Streuung unterliegen und mehr Ausreißer ausweisen als die Werte der klinisch nicht signifikanten.

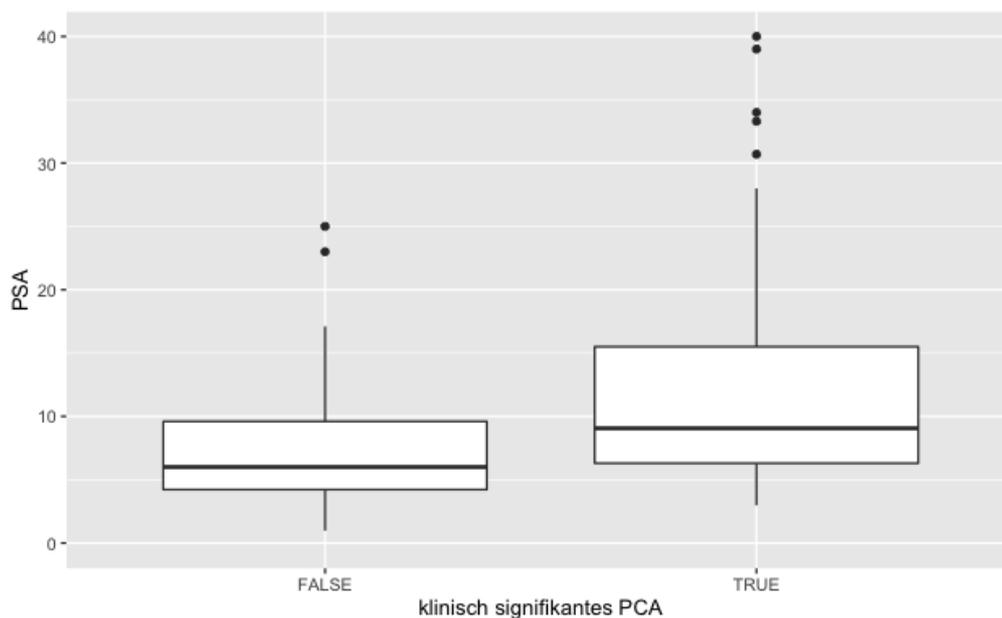


Abb. 26: Boxplot der PSA-Werte bei klinisch signifikanten und nicht signifikanten Prostatakarzinomen.

Drei PSA-Werte sind  $> 40$  ng/ml (332 ng/ml, 180 ng/ml, 74 ng/ml) und nicht abgebildet.

#### 4.4.2 PSA-Dichte und klinisch signifikante Prostatakarzinome

Der Zusammenhang zwischen PSA-Dichte und klinisch signifikanten Prostatakarzinomen ist in Abb. 27 veranschaulicht. Weil es sich um eine kontinuierliche und eine diskrete Variable handelt, wurde ein Boxplot verwendet. Dieser Plot stellt dar, dass die mediane PSA-Dichte der klinisch signifikanten Prostatakarzinome (Gleason-Score  $\geq 7$ ) höher ist als die der klinisch nicht signifikanten Prostatakarzinome (Gleason-Score  $\leq 6$ ). Ebenso lässt sich erkennen, dass die PSA-Dichten der klinisch signifikanten Prostatakarzinome einer sichtlich breiteren Streuung unterliegen und mehr Ausreißer ausweisen als die der klinisch nicht signifikanten.

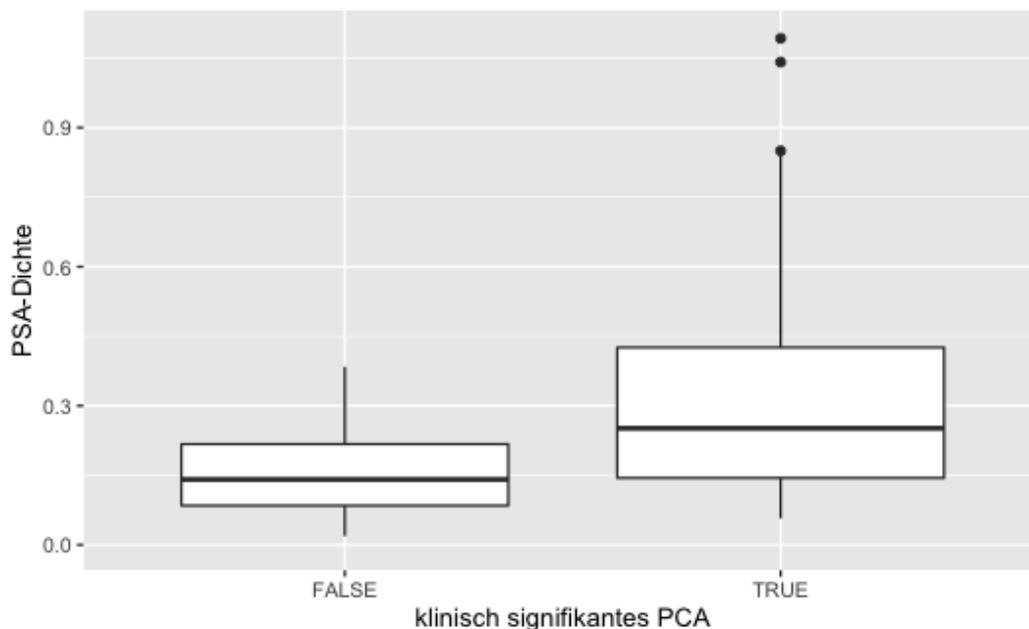


Abb. 27: Boxplot der PSA-Dichte bei klinisch signifikanten und klinisch nicht signifikanten Prostatakarzinomen.

Zwei PSA-Dichten sind  $> 1,5 \text{ ng/ml/cm}^3$  ( $1,76 \text{ ng/ml/cm}^3$ ,  $2,94 \text{ ng/ml/cm}^3$ ) und nicht abgebildet.

### 4.4.3 Prostatavolumen und klinisch signifikante Prostatakarzinome

Der Zusammenhang zwischen Prostatavolumen und klinisch signifikanten Prostatakarzinomen ist in Abb. 28 veranschaulicht. Weil es sich um eine kontinuierliche und eine diskrete Variable handelt, wurde ein Boxplot verwendet. Dieser Plot stellt dar, dass das mediane Prostatavolumen der klinisch signifikanten Prostatakarzinome (Gleason-Score  $\geq 7$ ) und der klinisch nicht signifikanten Prostatakarzinome (Gleason-Score  $\leq 6$ ) relativ ähnlich ist. Ebenso lässt sich erkennen, dass die Prostatavolumina der klinisch nicht signifikanten Prostatakarzinome zwar einer breiteren Streuung unterliegen, allerdings weniger Ausreißer ausweisen als die der klinisch nicht signifikanten.

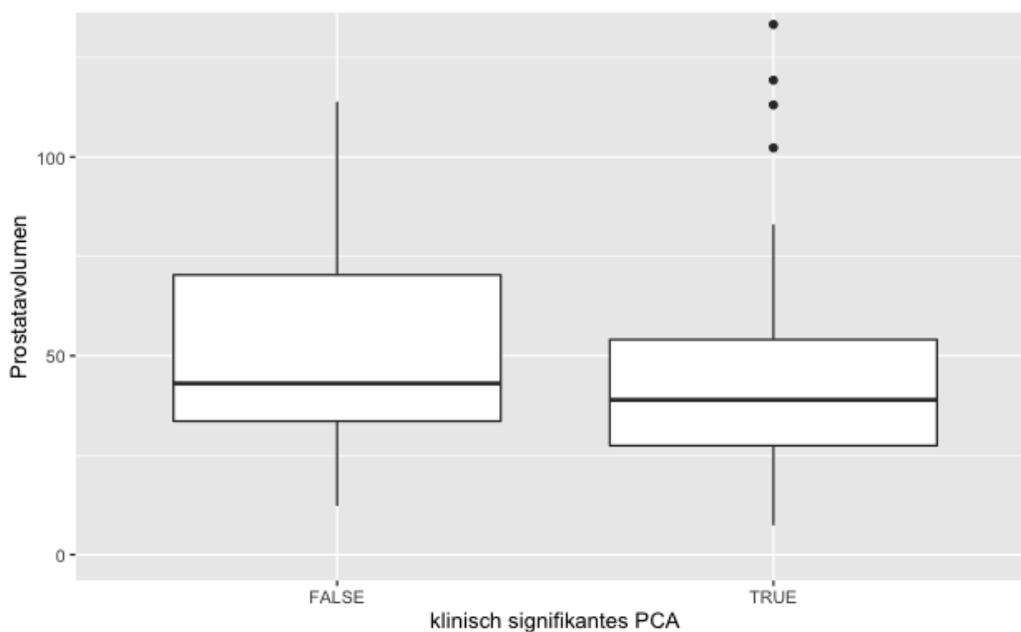


Abb. 28: Boxplot der Prostatavolumina bei klinisch signifikanten und nicht signifikanten Prostatakarzinomen.

Ein Prostatavolumen ist  $> 135$  ml (169,5 ml) und nicht abgebildet.

#### 4.4.4 PI-RADS-Score und klinisch signifikante Prostatakarzinome

Der Zusammenhang zwischen PI-RADS-Score und klinisch signifikanten Prostatakarzinomen ist in Abb. 29 veranschaulicht. Weil es sich um zwei diskrete Variablen handelt, wurde ein Punktplot verwendet. Die Größe der Punkte ist hier durch die jeweilige Fallzahl bestimmt. Dieser Plot stellt dar, dass die Mehrzahl der klinisch signifikanten Prostatakarzinome (Gleason-Score  $\geq 7$ ) mit einem PI-RADS-Score  $\geq 4$  bewertet wurde (die meisten davon mit einem PI-RADS-Score von 5). Ebenso lässt sich erkennen, dass die Mehrheit der klinisch nicht signifikanten Prostatakarzinome (Gleason-Score  $\leq 6$ ) mit einem PI-RADS-Score  $< 4$  bewertet wurde (die meisten davon mit einem PI-RADS-Score von 3).

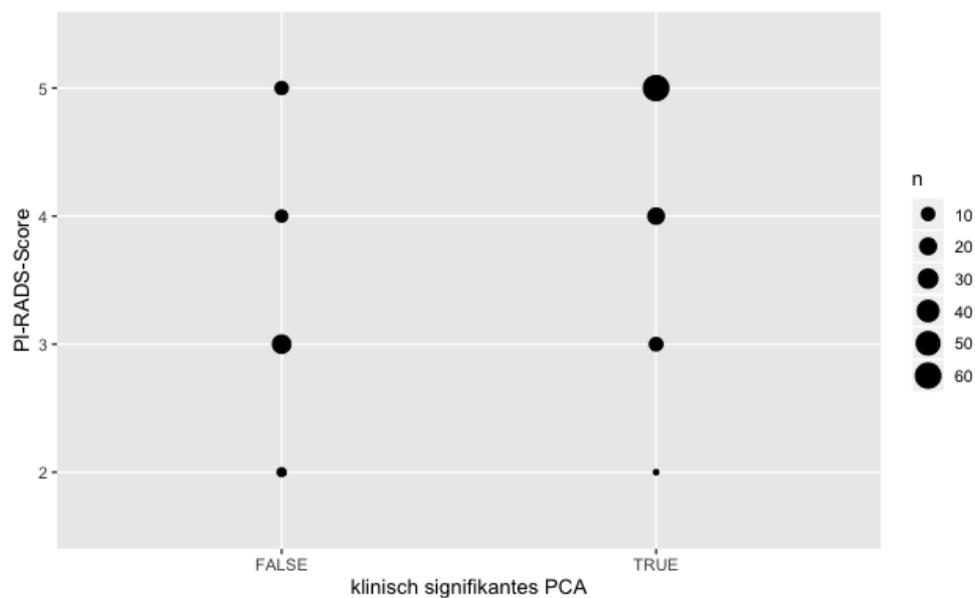


Abb. 29: Punktplot der PI-RADS-Scores bei klinisch signifikanten und klinisch nicht signifikanten Prostatakarzinomen. n = Anzahl der Fälle.

#### 4.4.5 Mittlerer ADC-Wert und klinisch signifikante Prostatakarzinome

Der Zusammenhang zwischen mittlerem ADC-Wert und klinisch signifikanten Prostatakarzinomen ist in Abb. 30 veranschaulicht. Weil es sich um eine kontinuierliche und eine diskrete Variable handelt, wurde ein Boxplot verwendet. Dieser Plot stellt dar, dass der mediane mADC-Wert der klinisch signifikanten Prostatakarzinome (Gleason-Score  $\geq 7$ ) deutlich niedriger ist als der mediane mADC-Wert der klinisch nicht signifikanten Prostatakarzinome (Gleason-Score  $\leq 6$ ). Ebenso lässt sich erkennen, dass die mADC-Werte der klinisch signifikanten Prostatakarzinome einer sichtlich kleineren Streuung unterliegen als die der klinisch nicht signifikanten.

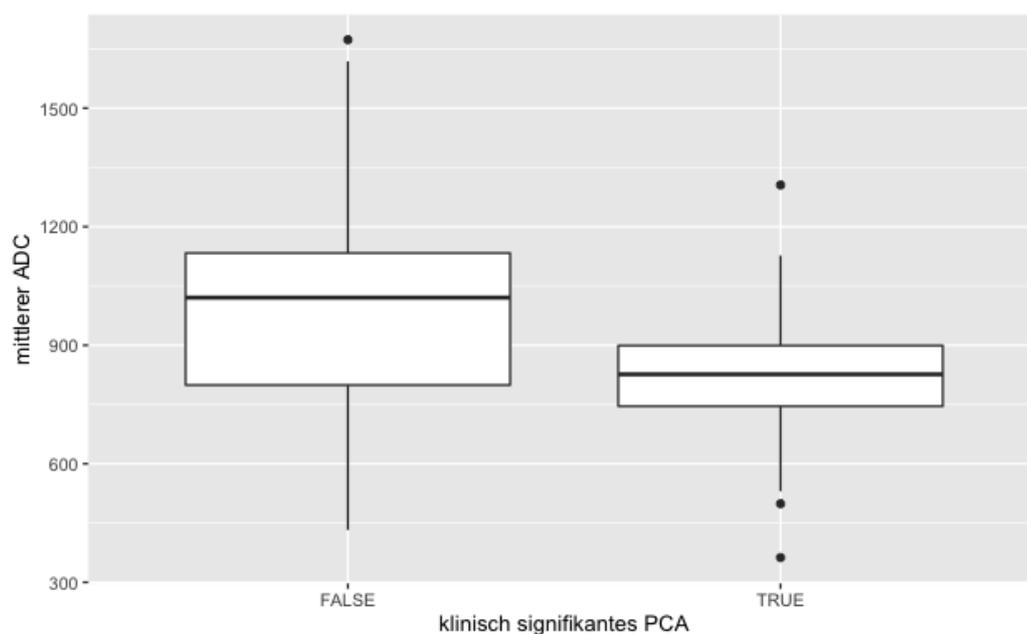


Abb. 30: Boxplot der mittleren ADC-Werte bei klinisch signifikanten und klinisch nicht signifikanten Prostatakarzinomen.

## 4.5 Genauigkeit des Histologie-Ergebnisses der Biopsie im Vergleich zum Prostatektomiepräparat

Bei 83 der 140 Patienten erfolgte zusätzlich zur initialen Stanzbiopsie eine Prostatektomie, sodass hier ein postoperatives Histologieergebnis vorlag. In dieser Studie konnten durch das Histologieergebnis der Biopsie 65 von 73 Patienten mit einem klinisch signifikantem Prostatakarzinom (Gleason-Score  $\geq 7$ ) richtig erkannt werden. 9 von 10 Patienten mit einem klinisch nicht signifikanten Prostatakarzinom (Gleason-Score  $\leq 6$ ) oder keinem Prostatakarzinom wurden ebenfalls richtig erkannt. Bei 8 Patienten kam es nach erfolgter Prostatektomie zu einem Upgrading auf ein klinisch signifikantes Prostatakarzinom.

Die Sensitivität, die Patienten anhand des Biopsie-Ergebnisses in die Kategorien 1) klinisch signifikantes Prostatakarzinom und 2) klinisch nicht signifikantes Prostatakarzinom oder kein Prostatakarzinom einzuteilen, betrug somit 0,89, die Spezifität 0,90 (s. Tabelle 7).

|  | <b>Prostatektomie:<br/>GS &gt; 6</b> | <b>Prostatektomie:<br/>GS <math>\leq</math> 6 oder kein PCA</b> |
|--|--------------------------------------|---|
| <b>Biopsie: GS &gt; 6</b>                            | 65 RP                                | 1 FP  |
| <b>Biopsie: GS <math>\leq</math> 6 oder kein PCA</b> | 8 FN                                 | 9 RN  |
|  | <b>Sensitivität<br/>= 0,89</b>       | <b>Spezifität<br/>= 0,90</b>                                    |

Tabelle 7: Kontingenztafel zur Genauigkeit des Biopsieergebnisses (mit der Histologie aus dem Prostatektomiepräparat als Goldstandard).

GS = Gleason-Score, PCA = Prostatakarzinom, RP = richtig Positive, FP = falsch Positive, FN = falsch Negative, RN = richtig Negative.

## 5 Diskussion

Im Rahmen dieser Studie wurde untersucht, ob anhand von machine learning-Modellen basierend auf quantitativen Radiomics-Features aus Prostata-mpMRT-Aufnahmen besser als anhand des PI-RADS-Scores (2.0) oder des mittleren ADC-Wertes klinisch signifikante von klinisch nicht signifikanten Prostatakarzinomen differenziert werden können. Solche Modelle könnten die prätherapeutischen Tumoraggressivität vorhersagen und so die Prognoseabschätzung, den Auswahlprozess für Patienten für eine Prostatabiopsie und die Therapieplanung bei Prostatakarzinompatienten verbessern und individualisieren.

### 5.1 Interpretation der Ergebnisse

#### 5.1.1 PI-RADS

Bei der Befundung nach PI-RADS (2.0) ergab die ROC-Analyse, dass mit einer Wahrscheinlichkeit von 78% ein zufälliges Paar eines auffälligen und eines nicht auffälligen PI-RADS-Scores korrekt als klinisch signifikantes (Gleason-Score  $\geq 7$ ) oder nicht signifikantes Prostatakarzinom (Gleason-Score  $< 6$ ) klassifiziert werden konnte. Das 95%-Konfidenzintervall war mit 0,70 – 0,86 relativ groß. Das Konfidenzintervall hätte sich mit einem größeren Patientenkollektiv verkleinern lassen, da man sich mit einer größeren Stichprobe immer mehr dem „tatsächlichen“ Wert annähern kann (solange man keine systematischen Fehler macht). Ebenso scheint die Streuung recht groß gewesen zu sein, was zu einem breiteren Konfidenzintervall führte. Erwartungsgemäß lag der optimale Cut-Off-Wert nach Youden bei einem PI-RADS-Score von 4. Auch PI-RADS 2.0 empfiehlt Prostatakarzinome in der MRT ab diesem Score als klinisch signifikant zu betrachten und eine Biopsie durchführen zu lassen. Bei diesem Cut-Off-Wert wurde in unserer Studie eine moderate Sensitivität von 87% und Spezifität von 62% für das Vorliegen eines klinisch signifikanten Prostatakarzinoms erreicht. Dieses Ergebnis entspricht der Literatur (s. 5.2) und ist natürlich abhängig von der Einschätzung und Erfahrung der Ärzte, die die Bilder der eingeschlossenen Patienten befundeten, da Prostatakarzinome oft nicht einfach zu differenzieren

sind und sich beispielsweise ähnlich einer benignen Prostatitis präsentieren können.

### **5.1.2 ADC**

Bei der Klassifizierung in klinisch signifikante und klinisch nicht signifikante Prostatakarzinome anhand des mittleren ADC-Wertes ergab die ROC-Analyse eine Fläche unter der Kurve (AUC) von 0,72. Das Konfidenzintervall war mit 0,61 – 0,82 recht groß, wahrscheinlich ebenfalls aufgrund der nicht allzu großen Fallzahl. Der optimale Cut-Off-Wert nach Youden lag bei einem mittleren ADC-Wert von 1016,5, mit dem eine Sensitivität von 95% und eine Spezifität von 51% für das Vorliegen eines klinisch signifikanten Prostatakarzinoms erreicht wurde. Unsere Ergebnisse suggerieren, dass der objektive mittlere ADC-Wert allein also nicht dem subjektiven PI-RADS-Score überlegen ist. Es erscheint sinnvoll, dass ein einziger Durchschnittswert der Diffusionsrestriktion nicht adäquat die Gewebestände widerspiegeln kann beziehungsweise dies nicht besser kann als ausgebildete Radiologen, deren Beurteilung anhand des von Experten entwickelten PI-RADS-Scores auf relevanten Eigenschaften des Gewebes fußt.

### **5.1.3 Radiomics**

Bei der Radiomics-Methode schnitt die logistische Regression in der Differenzierung klinisch signifikanter Prostatakarzinome mit einer mittleren AUC von 0,75 etwas schlechter ab als die Random-Forest-Algorithmen, deren mittlere AUC im Bereich von 0,80 – 0,84 lagen. Aufgrund der nicht allzu großen Fallzahl wurden fünf Runden bei der Kreuzvalidierung durchgeführt. Die recht breite Streuung hätte wahrscheinlich durch einen insgesamt homogeneren Datensatz mit höherer Fallzahl verkleinert werden können. Zwischen der Performance der Random-Forest-Algorithmen lässt sich kein großer Unterschied erkennen: Die Bereiche, in die die Ergebnisse der fünf Validierungsteilmengen eines jeden Algorithmus fallen, überschneiden sich größtenteils. Das bedeutet, dass der „Basis“-Random-Forest durch keine der jeweiligen Stellschrauben relevant verändert werden konnte, weder durch eine unterschiedliche Anzahl an Features, die miteinbezogen wurden, noch durch Methoden wie „oversampling“ oder „class weights“, die unserem unbalancierten Datensatz entgegenwirken sollten. Es scheint hier also noch unbekannte Faktoren zu geben, die Einfluss auf die Klassifizierung der Algorithmen haben und bei Veränderung diese

möglicherweise verbessern würden. Dennoch lässt sich das Potential von Radiomics erkennen, die Risikostratifizierung im Diagnoseprozess des Prostatakarzinoms zu unterstützen. Laut unseren Resultaten lassen die radiomischen Random-Forest-Modelle eine exaktere Klassifizierung in klinisch signifikante und klinisch nicht signifikante Prostatakarzinome zu als der PI-RADS-Score (2.0) oder der mittlere ADC-Wert (PI-RADS: AUC 0,78, mADC: AUC 0,72, bestes Radiomics-Modell: AUC 0,84). Das ist auch schlüssig, da Radiomics darauf abzielt, phänotypische Unterschiede mitsamt der intratumoralen Heterogenität quantitativ, bildbasiert zu erfassen und somit eine umfassende und genaue Charakterisierung der verdächtigen Läsionen durchführen kann. Erschwerend ist hierbei jedoch, dass sich benigne Differentialdiagnosen deutlich unterscheiden können. Beispielsweise können eine Prostatitis oder unklar umkapselte Knoten im Rahmen einer benignen Prostatahyperplasie ähnlich einem Prostatakarzinom zu einer Signalabschwächung in der T2w-Sequenz führen. Um dem zu begegnen und Radiomics-Modelle präziser zu entwickeln, müssten solche Subkategorien extra berücksichtigt werden.

## **5.2 Einordnung der Ergebnisse in den aktuellen Stand der Forschung**

Wie gut PI-RADS klinisch signifikante Prostatakarzinome detektieren kann, untersuchten Woo et al. 2017 im Rahmen einer Metaanalyse (53). Dabei wurden aus 21 Studien eine gepoolte Sensitivität von 89% und eine gepoolte Spezifität von 73% berechnet. Die Sensitivität entspricht in etwa den Ergebnissen der vorliegenden Arbeit der PI-RADS (2.0)-Methode (87%), die Spezifität ist in der vorliegenden Arbeit jedoch deutlich niedriger (62%). Diese Werte würden sich wahrscheinlich mit einer höheren Fallzahl zunehmend angleichen. Außerdem nutzten die einbezogenen Studien verschiedene Definitionen für klinisch signifikante Prostatakarzinome, da es dafür nach wie vor keine konkreten Vorgaben gibt (20), was allgemein einen Vergleich erschwert. Die größte Herausforderung bei der Beurteilung nach PI-RADS bleibt die subjektive Einschätzung sowie die selbst unter erfahrenen Radiologen bestehende Variabilität der Befundergebnisse. So wurden beispielsweise in einer Studie von

240 klinisch signifikanten Prostatakarzinomen nur 121 von allen sechs Radiologen als solche erkannt (98).

Der mittlere ADC-Wert, als quantitativer Parameter, soll die Variabilität zwischen befundenden Radiologen verringern. Wie andere Autoren zeigten auch Nowak et al., dass zwischen dem mittleren ADC-Wert und klinisch signifikanten Prostatakarzinomen eine Korrelation besteht, sowohl für Karzinome der peripheren Zone als auch der Transitionalzone (65). Der mittlere ADC-Wert liefert somit eine wertvolle, objektive Einschätzung, erkennt aber wie auch die Ergebnisse der vorliegenden Arbeit nahelegen, mit AUC von 0,70 – 0,81 alleine nicht zwingend genauer klinisch signifikante Prostatakarzinome ((99), (65), (100)). Die divergierenden Ergebnisse in der Literatur werden einerseits auf die meist recht kleinen Fallzahlen zurückzuführen sein. Andererseits beeinflussen diverse Parameter die Messung des ADC-Wertes, sodass beispielsweise die Wahl der b-Werte zu variablen Ergebnissen führen könnte. Bajgiran et al., die mit 218 eingeschlossenen Patienten eine verhältnismäßig große Studienpopulation wählten, bestimmten statt einem absoluten mittleren ADC-Wert ein ADC-Verhältnis, indem sie den mittleren ADC-Wert der Läsion durch den mittleren ADC-Wert einem benignen Referenzareal gegenüberstellten (99). So konnten sie statt der absoluten die relative ADC-Wert-Veränderung messen. Dieser Ansatz verspricht, den Einfluss der unterschiedlichen Untersuchungsparametereinstellungen auszuschalten.

Die Ergebnisse zur Detektion von klinisch signifikanten Prostatakarzinomen anhand von Radiomics-Modellen sind in der Literatur sehr heterogen. Neben den vielseitigen Parametereinstellung der MRT-Untersuchung an sich, die die radiomischen Biomarker, wie den mittleren ADC-Wert, beeinflussen können, gibt es viele weitere Stellschrauben im Rahmen des radiomischen Workflows, wie beispielsweise verschiedene Bildbearbeitungsmethoden, Klassifikationsalgorithmen oder Feature-Auswahl-Verfahren. Chaddad et al. untersuchten in ihrer Studie ebenfalls einen Random-Forest-Algorithmus und lagen mit einer AUC von 0,78 innerhalb den Schwankungsbreiten der vorliegenden Arbeit (101). Kwon et al. hatten mit 344 eingeschlossenen Patienten eine ungewöhnlich hohe Fallzahl und erreichten mit ihrem Random-Forest-Algorithmus eine AUC von 0,82, der ähnlich der vorliegenden Studie im Vergleich zu der Entscheidungsbaumanalyse CART (classification and

regression trees) und der Regressionsanalyse LASSO (least absolute shrinkage and selection operator) als bester Algorithmus abschnitt (102). Obwohl sie zusätzlich zu T2w-Sequenzen und ADC-Karten auch DCE- und PD(proton density weighted)-Sequenzen verwendeten und somit vielseitigere Tumorcharakteristika miteinbezogen, erreichten sie in ihrer Studie damit keine genauere Differenzierung klinisch signifikanter Prostatakarzinome als die vorliegende Arbeit.

Bonekamp et al. verglichen 2018 ähnlich zur vorliegenden Studie die Performance des PI-RADS-Scores, des mittleren ADC-Wertes und Radiomics (86). Sie untersuchten ebenfalls einen Random-Forest-Algorithmus für ihre Radiomics-Analyse und erreichten mit 316 Patienten eine mehr als doppelt so hohe Fallzahl. Sie erzielten mit den von ihnen optimal gesetzten Cut-off-Werten vergleichbare Ergebnisse für die PI-RADS- und mittlere ADC-Wert-Analysen (PI-RADS: Sensitivität 89%, Spezifität 43%, mADC: Sensitivität 93%, Spezifität 51%), jedoch leicht bessere Ergebnisse bei ihrer Radiomics-Analyse. Das könnte damit zusammenhängen, dass für die Bildaufnahmen bei allen Patienten dasselbe MRT-Gerät derselben Institution verwendet wurde und somit viele „Störfaktoren“ ausgeschaltet werden konnten.

Einige andere Studien haben Radiomics immer entweder mit dem PI-RADS-Score oder dem mittleren ADC-Wert verglichen: Orczyk et al. erreichten mit ihrem radiomischen Modell basierend auf einer logistischen Regression eine AUC von 0,93 verglichen zu 0,91 ihrer PI-RADS-Analyse, sodass 53% der MRT-Läsionen die Biopsie erspart hätte bleiben können (103). Algohary et al. testeten drei Klassifikatoren (quadratic discriminant analysis, support vector maschine, random forest), wobei sie sich auf die in der MRT falsch positiven und falsch negativen Fälle konzentrierten. Der Random-Forest-Algorithmus schnitt hierbei zwar hinsichtlich der Detektion von klinisch signifikante Prostatakarzinomen am schlechtesten ab, konnte aber trotzdem die Anzahl der falsch Positiven und falsch Negativen um circa jeweils ein Drittel im Vergleich zum PI-RADS-Score reduzieren (104). Beide Studien erreichten somit scheinbar bessere Ergebnisse als die vorliegende Arbeit, wobei diese in Anbetracht der sehr geringen Fallzahl (Orczyk et al. n = 20, Algohrary et al. n = 56) gering aussagekräftig sind.

Eine größere Studienpopulation, die mit 147 eingeschlossenen Patienten sehr ähnlich der vorliegenden Studie war, untersuchten Fehr et al.. Sie verglichen drei Radiomics-Modelle (t-test support vector machine, adaptive boosting und recursive feature selection support vector machine) mit einem mADC-Modell, für das sie einen support-vector-machine-Algorithmus auf den mittleren ADC-Werten trainierten (87). In dieser Studie wurden nur Patienten mit erfolgter Prostatektomie und damit exakter histologischer Klassifizierung des Prostatakarzinoms eingeschlossen. Fehr et al. versuchten in dieser Studie, eine wichtige Limitation zu begrenzen: Unbalancierte Klassen stellen für viele machine-learning-Modelle ein Problem dar und können sich ungünstig auf die Performance der Klassifikatoren auswirken. Mithilfe statistischer Methoden unter dem Überbegriff der sogenannten überrepräsentierten Stichprobe („oversampling“) kann die Anzahl der unterrepräsentierten Fälle angehoben werden und so wieder ein balancierter Datensatz geschaffen werden. Dafür gibt es verschiedene Techniken, Fehr et al. wandten die SMOTE(synthetic minority sampling technique)- und Gibbs-Methode an. Interessant war, dass durch die SMOTE-Methode die Performance vor allem eines Klassifikators deutlich angehoben werden konnte und so erst besser war als die Performance der ADC-Analyse. Auch in der vorliegenden Arbeit wurde aufgrund der unbalancierten Stichprobe SMOTE angewandt, wodurch jedoch keine Verbesserung erzielt werden konnte. Chen et al., mit einer fast dreifach so großen Patientenzahl, konnten bei ihrem radiomischen Modell dadurch ebenfalls keine Performanceverbesserung feststellen (105).

### **5.3 Besondere Stärken und Limitationen**

Eine Stärke der vorliegenden Studie ist, dass unsere radiomischen Modelle durch fünffache Kreuzvalidierung in fünf verschiedenen Szenarien evaluiert wurden. Somit wurde nicht wie in den meisten anderen Studien der Datensatz rein in einen Test- und Trainingsdatensatz geteilt und die Evaluierung in nur einem festen Szenario durchgeführt. Umfassender Literaturrecherche zufolge gibt es bis dato keine ähnliche Studie, die nicht nur ihr bestes Ergebnis, sondern alle Ergebnisse der Kreuzvalidierung veröffentlichte. Dadurch lässt sich besser einschätzen, wie sich die „tatsächliche“ Performance der radiomischen Modelle

unterscheidet.

Des Weiteren wurde in der Literatur bis dato kaum das Problem von unbalancierten Daten beachtet. Die vorliegende Studie wandte als eine der Wenigen aufgrund ungleicher Patientenzahl in den zwei Klassen (Gleason-Score  $\geq 7$  beziehungsweise Gleason-Score  $\leq 6$  oder kein Prostatakarzinom) die Methoden „oversampling“ und „class weights“ an, um diese Imbalance auszugleichen und eine dadurch entstehende Ergebnisverzerrung zu minimieren. Auch wenn sich final herausstellte, dass diese Methoden kaum Einfluss hatten, ist es eine wichtige Feststellung, die einerseits anderen Studien widerspricht (87) und andererseits auffordert weitere Stellschrauben in der Klassifikationsmodellentwicklung zu suchen.

Außerdem ist es durch die vorliegende Arbeit zum ersten Mal gelungen, sieben verschiedene radiomische Modelle direkt mit der Performance des mittleren ADC-Wertes und des PI-RADS-Scores (2.0) zu vergleichen und zu zeigen, dass diese neuen Methoden basierend auf maschinellem Lernen ein großes Potential besitzen, die Risikostratifizierung von Prostatakarzinomen im aktuellen Diagnoseprozess als Ergänzung zu den beiden konventionelleren Methoden zu verbessern. Unsere Fragestellung zielt darauf ab, nicht die allgemeine Detektion von Prostatakarzinomen zu untersuchen, sondern den Forschungskurs auf die Detektion klinisch signifikanter Prostatakarzinome zu setzen, was von deutlich höherer klinischer Relevanz ist.

Unsere Studie unterliegt einigen Limitationen. Erstens wurde ein monozentrisches, retrospektives Studiendesign mit einer nicht allzu großen Fallzahl verwendet. Eine multizentrische, prospektive Vorgehensweise hätte das Design aufwerten können. Allerdings ist es für eine Hypothesengenerierung ein übliches Vorgehen, das zur Validierung der Hypothese im nächsten Schritt multizentrische, prospektive, randomisierte, kontrollierte Studien mit großen Patientenzahlen bedarf.

Zweitens gibt es beim Einschluss der Patienten in unsere Studie Einschränkungen: Einerseits wurde mit dem Kriterium, dass innerhalb von maximal 6 Monaten nach Bildgebung ein Histologiebefund einer Prostatabiopsie oder -ektomie vorliegen muss, ein relativ großer Zeitraum gewählt. Hier ist ein kürzerer Zeitraum besser, allerdings hätte das wiederum unsere Patientenzahl reduziert. Darüber hinaus wäre es für eine genauere Differenzierung von

Patienten mit einem klinisch signifikanten Prostatakarzinom besser gewesen, in die Vergleichsgruppe der Patienten mit einem klinisch nicht signifikanten Prostatakarzinom nicht auch Patienten mit Karzinomverdacht, aber negativer Histologie aufzunehmen. Dies nahmen wir jedoch hinsichtlich einer ansonsten zu kleinen Vergleichsgruppe in Kauf (28 Patienten mit klinisch nicht signifikantem Prostatakarzinom statt 47 Patienten mit klinisch nicht signifikantem oder keinem Prostatakarzinom als Vergleichsgruppe bei 93 Patienten mit klinisch signifikantem Prostatakarzinom) und setzten somit den Fokus auf die Detektion klinisch signifikanter und somit therapierelevanter Prostatakarzinome.

Drittens wurde in der vorliegenden Studie keine Validierung an einem separaten Testdatensatz durchgeführt. Hätte man jedoch von den eingeschlossenen Patienten einen Teil für ein Testen nach Entwicklung der Modelle aufbehalten, wäre die Anzahl der Patienten, die für das Training verwendet werden hätte können recht gering gewesen, was unsere Ergebnisgenauigkeit beeinflusst hätte. Darüber hinaus erlaubte die verwendete Kreuzvalidierungsstrategie eine realistische Einschätzung der Schwankungsbreite der einzelnen Algorithmen. Viertens wurde nicht bei allen eingeschlossenen Patienten eine Prostatektomie durchgeführt, sodass teils nur das histologische Ergebnis der Biopsie vorlag. Dies war bei etwa einem Drittel der Patienten der Fall (49/140). Es wurden jedoch zu einem großen Teil MRT-TRUS-Fusionsbiopsien durchgeführt, die vergleichsweise zuverlässig sind (106). Darüber wurde bei den Patienten, bei denen eine Biopsie und eine Prostatektomie durchgeführt wurde, ausgewertet, wie genau der histologische Befund von beiden übereinstimmte, und dies zusammen mit den Ergebnissen offengelegt. Eine weitere Limitation stellt hier überdies dar, dass bei Patienten mit multifokalem Prostatakarzinom im Rahmen der vorliegenden Studie nur die Läsion mit dem höchsten PI-RADS-Score ausgewählt wurde. Somit war nicht klar, aus welcher Läsion der Gleason-Score bestimmt wurde. Dies traf allerdings nur auf insgesamt 17/140 Patienten zu, von denen wiederum nur zwei keine Prostatektomie erhielten. Fünftens wurden die MRT-Bilder an unterschiedlichen Geräten und Institutionen aufgenommen und auch unterschiedliche Protokolle gefahren. Dies wurde jedoch bewusst im Hinblick auf unser Ziel akzeptiert, radiomische Features entdecken und Modelle entwickeln zu wollen, die robust gegen diese Variationen sind, um ein zukünftiges Umsetzen in die Praxis zu erleichtern. Sechstens

wurden in der vorliegenden Studie die Segmentierungen manuell durchgeführt, was sehr zeitintensiv und grundsätzlich subjektiv ist. Eine Verbesserung wären hier semi-automatische oder gänzlich automatische Segmentierungstechniken (107).

Zuletzt stellten noch die unbalancierten Klassen eine Limitation dar, welche aber mit dem Einsatz entsprechender Methoden beachtet und ausgeglichen wurden (s. oben).

## **5.4 Konsequenzen**

### **5.4.1 Zukünftige Rolle der mpMRT in der Prostatadiagnostik**

Eine frühe und exakte Detektion von klinisch signifikanten Prostatakarzinomen ist essentiell für einen schnellen Therapiebeginn und damit unmittelbar verantwortlich für den Behandlungserfolg. Die Heilung bei lokalisierter Erkrankung beträgt 80 – 99% (108), nach Metastasierung beträgt die 5-Jahres-Überlebensrate nur noch circa 30% (109). Die mpMRT und in weiterer Folge die MRT-TRUS-Fusionsbiopsie wurde in den letzten Jahren fest in die Prostatakarzinomdiagnostik integriert und wird angesichts ihres großen Potentials auch weiter an Bedeutung gewinnen (42).

#### **Biopsienaive Patienten**

Zunehmend befürworten Studien im Rahmen der Primärdiagnose des Prostatakarzinoms einen „MRI-only-workflow“ (110). Die mpMRT könnte als Triage eingesetzt werden, ob die Durchführung einer Biopsie überhaupt nötig ist, um unnötige invasive Eingriffe und die Diagnose klinisch insignifikanter Prostatakarzinome zu reduzieren (111). Es scheint evident, dafür den PI-RADS-Score anzuwenden, der die Wahrscheinlichkeit angibt, mit der eine Prostataläsion in der mpMRT ein klinisch signifikantes Prostatakarzinom darstellt (s. 1.3.3). Padhani et al. empfehlen ähnlich der S3-Leitlinie folgendes Vorgehen (111): Bei PI-RADS 1- sowie 2-Läsionen spricht die Studienlage für das Unterlassen einer Biopsie, zumal die mpMRT mit strukturierter PI-RADS-Befundung einen sehr hohen negativen prädiktiven Wert aufweist. Dieser kann durch Einbezug klinischer Parameter weiter erhöht werden, vor allem durch Betrachtung der PSA-Dichte (112). Bei PI-RADS 3-Läsionen muss individuell

anhand der klinischen Gesamtsituation entschieden werden, bei PI-RADS 4- und 5-Läsionen wird zur Biopsie geraten. Bei diesem Vorgehen sollte der PI-RADS-Score möglichst präzise der Wahrscheinlichkeit eines klinisch signifikanten Prostatakarzinoms entsprechen, sodass Biopsien bei hoher Wahrscheinlichkeit durchgeführt, aber unnötige Biopsien vermieden werden. Das scheint jedoch vor allem bei den als PI-RADS 4 klassifizierten Läsionen nicht zuzutreffen, hier lagen in einer Studie beispielsweise nur in etwa 22% tatsächlich klinisch signifikante Prostatakarzinome vor (61). Gupta et al. sahen einen potentiellen Grund im Hochstufen einer PI-RADS 3- zu einer PI-RADS 4-Läsion durch eine positive DCE-Sequenz. Speziell in dieser Subgruppe der PI-RADS 4-Läsionen war die Rate klinisch signifikanter Prostatakarzinome gering (112).

Es ergeben sich zwei grundlegende Fragen: Welche minimale (beziehungsweise maximale) Detektionsrate von klinisch signifikanten Prostatakarzinomen (CDR = cancer detection rate) sollte in jeder PI-RADS-Kategorie vorliegen? Und, wie kann diese verbessert werden?

Bezüglich der ersten Frage wäre beispielsweise bei PI-RADS 4-Läsionen eine  $CDR > 0,5$  sinnvoll, zumal das Vorliegen eines klinisch signifikanten Prostatakarzinoms „wahrscheinlich“ sein soll. Letztlich mündet es in ein Abwägen von Sensitivität und Spezifität. Liegt die Präferenz auf einer möglichst hohen Detektionsrate von klinisch signifikanten Prostatakarzinomen (und wird somit eine erhöhte Diagnose von klinisch insignifikanten Karzinomen in Kauf genommen) oder auf einer möglichst geringen Detektionsrate von klinisch insignifikanten Prostatakarzinomen (und wird somit ein eventuelles Verkennen eines klinisch signifikanten Karzinoms toleriert)? Das lässt sich nicht pauschal beantworten und ist abhängig von der klinischen Situation. Es bedarf hier weiterer Diskussion und Forschung, was über den Rahmen dieser Arbeit hinausgeht.

Bezüglich der zweiten Frage wäre einerseits eine Integration klinischer Parameter anzustreben. Andererseits ist auch die Reduktion der Interobservervariabilität bei der mpMRT-Befundung ein valider Ansatz, denn diese ist trotz Einführung des PI-RADS-Scores nach wie vor hoch (112). Die quantitative MRT-Bildgebung setzt hier an und möchte mittels ADC-Wertes oder des Einsatzes von machine-learning-Algorithmen mehr Objektivität schaffen und

so eine bessere Detektionsrate von klinisch signifikanten Prostatakarzinomen erzielen. Diese und andere Arbeiten (s. 5.2) zeigten, dass eine Risikostratifizierung von Prostatakarzinomen anhand des PI-RADS-Scores oder des mittleren ADC-Wertes durch radiomische Modelle übertroffen werden kann. Greer et al. haben insbesondere dargelegt, dass durch CAD (computer aided diagnosis)-Systeme (s. 1.3.5) die Übereinstimmung von Radiologen von 57% auf 72% angehoben werden konnte (113). Gaur et al. betonten in ihrer 2018 publizierten Studie, dass CAD-Systeme vor allem die Spezifität und Karzinomdetektionsrate von moderat erfahrenen Radiologen für klinisch signifikante Prostatakarzinome in der Transitionalzone verbesserten (114).

### **MRT-TRUS-Fusionsbiopsie**

Die Fusionsbiopsie bietet einige Vorteile im Vergleich zur systematischen TRUS-Biopsie. Sie reduziert die Anzahl der nötigen Biopsien, detektiert weniger klinisch insignifikante und mehr klinisch signifikante Prostatakarzinome (letzteres vor allem bei zuvor negativer TRUS-Biopsie) und schließt weniger Stenzen ein, was die Biopsie-assoziierte Morbidität senken könnte (111). Noch wird die Fusionsbiopsie nicht generell der systematischen Biopsie vorgezogen, vor allem bei biopsienaiven Patienten wird aktuell eine Kombination aus beiden empfohlen, da nach wie vor eine heterogene Gruppe klinisch signifikanter Prostatakarzinome nur von der TRUS-Biopsie detektiert wird (s. 1.3.2). Padhani et al. untersuchten dazu die genauen Zahlen 200 biopsienaiver Männer mit mpMRT-Bildgebung: Bei 100 Patienten mit PI-RADS 1- und 2-Läsionen wurde nur eine systematische Biopsie durchgeführt, bei 100 Patienten mit P-RADS 3-, 4- und 5-Läsionen zusätzlich eine Fusionsbiopsie (111). In der ersten Gruppe, in der trotz PI-RADS 1- oder 2-Läsion eine systematische Biopsie durchgeführt wurde, wurde dadurch in 8% ein klinisch signifikantes Prostatakarzinom detektiert (andererseits führte dieses Vorgehen zu einer Überdiagnose klinisch insignifikanter Karzinome von 18%). In der zweiten Gruppe wurden in 5% klinisch signifikante Prostatakarzinome nur durch die systematische TRUS-Biopsie erkannt (die Rate für klinisch insignifikante Karzinome lag hier für die TRUS-Biopsie nur bei 4%). Um die Rolle der Fusionsbiopsie alleine zu stärken und ihre CDR zu erhöhen, bedarf es weiterer Forschung. Ein Ansatz könnte beispielsweise sein, mehrere Stenzen zu entnehmen. So stellten Zhang et al. fest, dass mit steigender Anzahl gewonnener Stenzen mehr klinisch signifikante Karzinome detektiert wurden

(115). Auch Radiomics-Modelle könnten hier einen Mehrwert liefern, da sie, wie auch diese Arbeit zeigt, das Potential besitzen, die Detektionsrate von klinisch signifikanten Prostatakarzinomen zu optimieren.

### **Aktive Überwachung**

Die aktive Überwachung ist eine Therapieoption für klinisch nicht signifikante Prostatakarzinome mit geringem Risiko, bei der aktuell regelmäßig MRT-Untersuchungen und Biopsien zur Überwachung der Tumorkläsionen durchgeführt werden. Ziel ist es, so eine Übertherapie von indolenten Prostatakarzinomen zu verhindern. Sie, und damit auch die mpMRT, wird jedoch mit dem allgemeinen Ziel, eine höhere Detektionsrate von klinisch signifikanten Prostatakarzinomen zu erreichen, eine immer bedeutendere Rolle spielen. So werden wahrscheinlich zukünftig mehr Patienten, sofern ein detektiertes Prostatakarzinom nicht klinisch signifikant ist, zur Verlaufsbeobachtung der aktiven Überwachung zugeteilt werden. Die Entscheidung, ob ein Patient in die aktive Überwachung aufgenommen wird, basiert hauptsächlich auf dem Biopsieergebnis (meist Gleason-Score  $\leq 6$ ), dessen Aussagekraft jedoch begrenzt ist (s. 1.2.3). Yim et al. stellten fest, dass der ADC-Wert und der PIRADS-Score unabhängige Prädiktoren für klinisch insignifikante Prostatakarzinome vor der aktiven Überwachung sind (116). Auch die MRT-Fusionsbiopsie wird an Bedeutung gewinnen, denn mehrere Studien zeigten, dass durch eine mpMRT und gezielte Fusionsbiopsie im Vergleich zur systematischen TRUS-Biopsie über ein Drittel der Prostatakarzinome reklassifiziert und der aktiven Überwachung zugeführt werden konnten (112). Das deutet auf eine wachsende Rolle und Relevanz der aktiven Überwachung hin. Dennoch bleibt eine gewisse Unsicherheit, dass trotz mpMRT und Biopsie ein aggressiver Tumor übersehen wurde. Diese könnte durch Implementierung von Radiomics-Tools in die MRT-Aufnahme im Vergleich zu konventionellen Methoden reduziert werden und als Triage zur Auswahl von Patienten für die aktive Überwachung verwendet werden. Die Lebensqualität der Patienten könnte somit ohne unnötige invasive diagnostische oder therapeutische Eingriffe entsprechend länger hochgehalten werden. Allerdings würden Radiomics-Modelle nicht nur als Triage vor der aktiven Überwachung eine Verbesserung erzielen, sondern könnten genauso ein nicht-invasives Monitoring

währenddessen ermöglichen, was einem absoluten Fortschritt im Sinne des Patientenwohles und der Präzisionsmedizin entsprechen würde.

### **Biparametrische MRT (bpMRT)**

Weiters sollte vermehrt die biparametrische MRT (bpMRT) beleuchtet werden. Sie besteht nur aus T2w- und DWI-Sequenzen, kommt demnach ohne Gadolinium-Kontrastmittel aus, hat eine kürzere Aufnahmezeit und ist zudem kostengünstiger als die mpMRT (117). In bestimmten klinischen Situationen könnte sie also eine valide Alternative darstellen (40). Auch in der aktualisierten PI-RADS Version 2.1, die im März 2019 erschien, wird die bpMRT miteinbezogen (118): Einerseits kann die bpMRT für die Detektion von klinisch signifikanten Prostatakarzinomen wertvoll sein und durch oben genannte Vorteile eventuell breiter eingesetzt werden. Andererseits ist die bei der bpMRT fehlende DCE-Sequenz sowohl für die Beurteilung der Transitional- als auch der peripheren Zone oftmals eine bedeutende Ergänzung und kann die Sensitivität der MRT erhöhen, vor allem, wenn die DWI-Aufnahme Artefakte aufweist. Demnach wird die Durchführung einer mpMRT speziell empfohlen, wenn klinische Risikofaktoren (beispielsweise eine familiäre Belastung oder eine zuvor negative Biopsie) das Vorliegen eines klinisch signifikanten Prostatakarzinoms wahrscheinlich machen, eine zuvor durchgeführte bpMRT trotz anhaltendem Karzinomverdacht unauffällig ist oder durch technische Faktoren nur eine suboptimale Qualität der DWI-Aufnahmen vorliegt. Für die bpMRT spricht das PIRADS Steering Committee noch keine spezifischen Empfehlungen aus, dazu sind weitere, multizentrischen Studien erforderlich. Schelb et al. nutzten für ihre Studie bpMRT-Daten und zeigten, dass durch automatische Detektion und Segmentierung klinisch signifikanter Prostatakarzinome anhand eines deep-learning-Modells (entwickelt aus bpMRT-Daten) ähnliche Ergebnisse erzielt werden konnten wie anhand des PI-RADS-Scores, der anhand von mpMRT-Daten bestimmt wurde (88). Auch in der vorliegenden Arbeit wurden die im Vergleich zur PI-RADS- und mADC-Methode besser abschneidenden Radiomics-Modelle nur auf bpMRT-Daten trainiert (T2w- und DWI-Sequenz). Allerdings muss bei dieser Studie die beträchtlich hohe Rate an falsch positiven Ergebnissen beachtet werden (fast 50%): Setzt man jedoch solche Modelle im Screening-Setting im Sinne eines Entscheidungsunterstützungssystems (DSS = decision support system) für Radiologen ein, kann der Vorteil einer hohen,

Modell-basierten CDR genutzt, jedoch trotzdem entsprechend der klinischen Gesamtsituation die Läsion final bewertet werden (119).

#### **5.4.2 Zukünftige Rolle von Radiomics-Modellen**

Auch wenn sich die Radiomics-Forschung vorerst nur mit Proof-of-Concept-Studien beschäftigt, zeichnet sich ab, dass Radiomics das Potential hat, einen signifikanten Mehrwert zur Risikostratifizierung in der Prostatakarzinomdiagnostik beizutragen. Immer mehr Studien, inklusive der vorliegenden Arbeit, zeigen, dass Radiomics-Modelle bildbasiert klinisch signifikante und klinisch nicht signifikante Prostatakarzinome besser als bisher etablierte Methoden (PI-RADS-Score, mADC-Wert) differenzieren können (s. 5.2). Die Unterschiede sind jedoch oft nicht sehr groß und die Studien retrospektiv an nur einer Institution durchgeführt. Andererseits wurde teilweise auch nur ein vergleichbares Abschneiden, aber keine Verbesserung durch Radiomics-Modelle beschrieben, wie beispielsweise in einer Studie von Bonekamp et al., in der anhand Radiomics-Modellen nicht vorteilhafter zwischen klinisch signifikanten und klinisch nicht signifikanten Prostatakarzinomen unterschieden werden konnte als anhand des ADC-Wertes (86). Ein solches Ergebnis kann allerdings von vielen Faktoren abhängen, weswegen in Zukunft weitere, multizentrische, prospektive Studien mit größeren Studienpopulationen und verschiedenen Radiomics-Methoden durchgeführt werden müssen. Im Allgemeinen lässt sich sagen, dass sich Radiomics durch seine Objektivität und umfangreiche Datenintegration vielversprechend von anderen Risikostratifizierungsansätzen abgrenzt und somit einen additiven Benefit im Vergleich zur Beurteilung nur einzelner Parameter (PI-RADS-Score, mADC, PSA-Wert, PSA-Dichte, Prostatavolumen, s. 4.4) bietet. Naheliegend ist dann, mehrere Faktoren für eine noch umfassendere Risikostratifizierung zu verknüpfen. Beispielsweise konnte gezeigt werden, dass ein auf PI-RADS-Scores und Radiomics-Features trainiertes machine learning-Modell besser klinisch signifikante Prostatakarzinome detektieren konnte als das Modell, das jeweils nur einzeln auf PI-RADS-Scores oder Radiomics-Features trainiert wurde (AUC von 0,98 versus AUC von 0,94 – 0,97) (120). Ebenso präsentierten Zhang et al. ein Radiomics-Nomogramm aus einem Radiomics-Modell und dem mittleren ADC-Wert, welches in der Differenzierung von klinisch signifikanten Prostatakarzinomen eine AUC von 0,84 erreichte und damit besser als das

Radiomics-Modell einzeln mit einer AUC von 0,81 abschnitt (121). Auch Woznicki et al. stellten durch eine Kombination mehrerer Parameter eine Verbesserung in der Detektion klinisch signifikanter Prostatakarzinome fest: In der Testkohorte erzielte ihr machine-learning-Modell, das sowohl auf Radiomics-Features als auch auf klinischen Parametern (PI-RADS-Score, PSA-Dichte und digital-rektale Untersuchung) trainiert wurde, mit einer AUC von 0,84 ein besseres Ergebnis als der PI-RADS-Score mit einer AUC von 0,69 oder der mittlere ADC-Wert mit einer AUC von 0,57 (122).

Wie könnten Radiomics-Modelle im klinischen Alltag eingesetzt werden? Im ersten Schritt wäre eine Implementierung im Sinne eines Entscheidungsunterstützungssystems (DSS = decision support system) für Radiologen als Orientierungshilfe in mehreren Szenarien sinnvoll. Einerseits als Support bei der Detektion und Lokalisation, andererseits auch bei der Risikostratifizierung von Prostatakarzinomen sowohl in klinisch signifikant und nicht signifikant als auch in weitere Subgruppen. Das könnte vor allem weniger erfahrenere Radiologen in der Durchführung einer optimaleren und individuelleren Therapieevaluierung, -planung und Prognoseeinschätzung, sowie in der Beurteilung von Follow-up-MRTs und Kontrolluntersuchungen im Rahmen der aktiven Überwachung unterstützen. Auch für eine optimale Patientenselektion für MRT-gesteuerte Biopsien könnten solche Entscheidungsunterstützungssysteme einen großen Benefit bringen. Die Folge wäre eine effizientere Prostatakarzinomdiagnostik, die genauer klinisch signifikante Prostatakarzinome erkennt und somit Überdiagnose und Übertherapie von indolenten Prostatakarzinomen reduziert. Im zweiten Schritt könnten dann umfassendere machine-learning-Modelle eingesetzt werden, die mehrere, auch klinische Prädiktoren berücksichtigen (s. oben) und eine nicht-invasive, vom Befunder unabhängige Risikostratifizierung im Sinne eines Entscheidungsunterstützungssystems darstellen.

## 5.5 Herausforderungen und nächste Schritte in Radiomics

Um Radiomics-Modelle weiterzuentwickeln und eine Implementierung in die Praxis zu ermöglichen, bedarf es prospektiver, multi-zentrischer, randomisierter und kontrollierter Studien mit größeren Fallzahlen, die extern validiert werden.

Ebenso sollten weitere Studien durchgeführt werden, die die Sinnhaftigkeit von einzelnen Radiomics-Modellen für die verschiedenen Prostatazonen untersuchen, da hierzu aktuell noch keine eindeutigen Ergebnisse vorliegen. Beispielsweise unterscheiden sich die mittleren ADC-Werte von Läsionen in der peripheren Zone und der Transitionalzone (s. **Error! Reference source not found.**, (123)). Weiters scheinen auch peritumorale Radiomics-Features einen Mehrwert zu bieten. So zeigten Algohary et al. in einer multizentrischen Studie, dass Radiomics-Modelle, die zusätzlich zu den aus bpMRT-Bilddaten gewonnenen intraläsionalen auch auf peritumorale Radiomics-Features basierten, eine 3 – 6-prozentige Verbesserung der Risikostratifizierung von Prostatakarzinomen herbeiführen konnten (AUC von 0,87 versus AUC von 0,84) (124).

Auch die verschiedenen Methoden im Rahmen des radiomischen Workflows sollten weiter geprüft werden. Viswanath et al. führten aktuell eine der ersten multi-zentrischen Studien durch, um zwölf verschiedene Arten von Klassifikatoren zu vergleichen (125). Auch deep-learning-Methoden sind aktuell in den Fokus gerückt (88), die besonders attraktiv sind, da sie ohne manuelle Läsionssegmentierung arbeiten.

Eine große Herausforderung wird es sein, Standards bei allen Faktoren der Bildakquisition festzulegen, um die Daten vergleichbar und für alle nutzbar zu machen. Der Schlüssel zu Radiomics sind nicht perfekte Features, sondern stabile und reproduzierbare (126). Derzeit wird zwar viel Engagement für die Entwicklung von quantitativen Imaging Biomarkern (QIB) gezeigt, aber keine hinreichende Evaluierung der technischen und klinischen Performance der Marker durchgeführt, da dafür oft inkonsistente und inkorrekte Terminologien und Methoden verwendet werden (127). Dazu sollte auch an einem „open-data“-

Netzwerk gearbeitet werden, das Daten öffentlich zugänglich macht und den Wissenstransfer und die Reproduzierbarkeit von Studien erhöht.

Eine weitere Hürde wird die gewaltige Menge an Daten sein, die im Rahmen der auf künstlicher Intelligenz (KI) basierenden Modellen anfällt (128). Einerseits verlangt diese entsprechenden Speicherplatz, andererseits auch professionelles Management durch Fachkundige. Cloud-Lösungen werden hier ein attraktives Angebot darstellen, welche jedoch mit entsprechenden Kosten und Datenschutzdiskussionen verbunden sind. Weiters muss diskutiert werden, wie mit der fraglichen Transparenz und den Ergebnissen von solchen KI-Modellen umgegangen wird.

Abschließend muss hier noch erwähnt werden, dass auch an einer einheitlichen Definition eines klinisch signifikanten Prostatakarzinoms gearbeitet werden muss, um Klarheit zu schaffen und Forschungsergebnisse vergleichbar zu machen.

## **5.6 Ausblick**

„Das ideale Ziel von Computer-unterstützter Diagnostik ist es, nicht generell die Anzahl an Prostatakarzinomen, sondern vielmehr die klinisch signifikanten Prostatakarzinome zu erkennen.“ Turkbey et al. formulierten diese Zielvorstellung (129), die der der vorliegenden Studie gleicht. Radiomics scheint dazu sehr gut geeignet zu sein, zumal es eine objektive, grundsätzlich reproduzierbare, quantitative Messmethode ist. Die MRT-Bildgebung ist weit verbreitet, ihr Potential aber noch keineswegs ausgeschöpft, auch wenn noch einige Herausforderungen vor einer Implementierung in den klinischen Alltag angegangen werden müssen. Was bringt die Zukunft?

McGarry et al. konnten erstmals mithilfe von radiomischen Features einen neuartigen Bildkontrast, sogenannte Gleason-Score-Wahrscheinlichkeitskarten, entwickeln, die die Wahrscheinlichkeiten für einen bestimmten Gleason-Score anzeigen (130). Dies könnte es möglich machen, gezielte Biopsien noch effizienter einzusetzen. Nagpal et al. setzten hier an (131): Sie wollten das Korrelat von Radiomics in der Pathologie anstoßen, ein „Pathomics“ entwickeln. Die Pathologie zeigt viele Parallelen zur Radiologie. Auch der Gleason-Score ist

ähnlich dem PI-RADS-Score eine subjektive, visuelle Beurteilung des Präparates unter dem Mikroskop, wobei ebenso Bilderkennungsmechanismen mittels maschinellen Lernens eingesetzt werden könnten, um eine objektivere Beurteilung zu ermöglichen. Weiter vorangeschritten ist das verwandte Feld der Radiogenomics, das zum Ziel hat, schon aus Imaging Features (epi)genetische Informationen zu gewinnen. Der Hype vom „next generation sequencing“ wird vielleicht bald zum „next generation imaging“. Dieses Feld ermöglicht einen über Radiomics alleine hinausgehenden Einblick, um letztlich maßgeblich klinische Entscheidungen beeinflussen zu können. Sowalsky et al. fanden zum Beispiel heraus, dass eine gemeinsame Vorstufe der Gleason-Muster 3 und 4 bei manchen Gleason-Muster-3-Prostatakarzinomen durch eine bestimmte Mutation erkennen lässt, ob dieses sich in ein Gleason-Muster-4-Prostatakarzinom entwickeln und somit klinisch signifikant werden wird oder nicht (132).

## 6 Zusammenfassung

Das Prostatakarzinom ist weltweit die zweithäufigste Krebserkrankung des Mannes (1). Die Heterogenität seiner Histologie, Tumoraggressivität und lokalen Ausbreitung stellt eine Herausforderung für das Krankheitsmanagement dar und erfordert eine exakte Risikostratifizierung für eine optimale und individuelle Therapieplanung.

Diese Studie hatte zum Ziel zu testen, ob Radiomics eine bessere, bildbasierte Detektion von klinisch signifikanten und damit sofort therapiebedürftigen Prostatakarzinomen erlauben könnte als die herkömmlichen bildgebenden Verfahren (PI-RADS-Score 2.0 und mittlerer ADC-Wert).

Dazu wurden basierend auf den T2w-Sequenzen sowie ADC-Karten der mpMRT-Aufnahmen der 140 eingeschlossenen Patienten sieben Radiomics-Algorithmen (einer als logistische Regression, die anderen als Random Forests) entwickelt, um Patienten mit klinisch signifikanten (Gleason-Score  $\geq 7$ ) von Patienten mit klinisch nicht signifikanten (Gleason-Score  $\leq 6$ ) oder keinen Prostatakarzinomen zu differenzieren und so die Risikostratifizierung des Prostatakarzinoms im prätherapeutischen Diagnoseprozess zu unterstützen. Ebenso wurden der PI-RADS-Score (2.0) und der mittlere ADC-Wert hinsichtlich ihrer Differenzierungsfähigkeit klinisch signifikanter Prostatakarzinome untersucht, sodass alle drei Methoden anhand der ROC-Analyse verglichen werden konnten.

Unsere Resultate zeigen, dass die Radiomics-Random-Forest-Modelle eine exaktere Klassifizierung zulassen als der PI-RADS-Score (2.0) oder der mittlere ADC-Wert. Erstere erzielten eine mittlere Fläche unter der Kurve im Bereich von 0,80 – 0,84, der PI-RADS-Score hingegen nur 0,78 und der mittlere ADC-Wert nur 0,72.

Diese Ergebnisse bestätigen, dass Radiomics durch eine MRT-basierte Charakterisierung des Prostatakarzinoms die prätherapeutische Prädiktion der Tumoraggressivität und somit Risikostratifizierung unterstützen kann. Ganz im Sinne der personalisierten Medizin kann Radiomics somit einen wertvollen Beitrag zur Optimierung des Patientenselektionsprozesses für Prostatabiopsien sowie

der individuellen Therapieplanung und folglich des Patientenergebnisses leisten, als auch einen effektiveren Einsatz des Gesundheitsbudgets ermöglichen.

# Literaturverzeichnis

1. World Cancer Research Fund. Prostate cancer statistics 2018 [Available from: <https://www.wcrf.org/dietandcancer/cancer-trends/prostate-cancer-statistics>].
2. Zentrum für Krebsregisterdaten, Robert Koch-Institut, Prostatalkrebs (Prostatakarzinom) 2017 [Available from: [https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Prostatalkrebs/prostatalkrebs\\_no\\_de.html](https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Prostatalkrebs/prostatalkrebs_no_de.html)].
3. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF), . Interdisziplinäre Leitlinie der Qualität S3 zur Früherkennung, Diagnose und Therapie der verschiedenen Stadien des Prostatakarzinoms 2019. Available from: [https://www.leitlinienprogramm-onkologie.de/fileadmin/user\\_upload/Downloads/Leitlinien/Prostata\\_5\\_0/LL\\_Prostatakarzinom\\_Kurzversion\\_5.1.pdf](https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Prostata_5_0/LL_Prostatakarzinom_Kurzversion_5.1.pdf).
4. Adams J. The case of scirrhus of the prostate gland with corresponding affliction of the lymphatic glands in the lumbar region and in the pelvis. *Lancet*. 1853;1:393-.
5. Lane JA, Donovan JL, Davis M, Walsh E, Dedman D, Down L, et al. Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. *Lancet Oncol*. 2014;15(10):1109-18.
6. Bott SR, Birtle AJ, Taylor CJ, Kirby RS. Prostate cancer management: (1) an update on localised disease. *Postgrad Med J*. 2003;79(936):575-80.
7. Center MM, Jemal A, Lortet-Tieulent J, Ward E, Ferlay J, Brawley O, et al. International variation in prostate cancer incidence and mortality rates. *Eur Urol*. 2012;61(6):1079-92.
8. Robert Koch-Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland (GEKID). Krebs in Deutschland für 2013/2014 2014.
9. Trogdon JG, Falchook AD, Basak R, Carpenter WR, Chen RC. Total Medicare Costs Associated With Diagnosis and Treatment of Prostate Cancer in Elderly Men Total Medicare Costs Associated With Diagnosis and Treatment of Prostate Cancer in Elderly Men Total Medicare Costs Associated With Diagnosis and Treatment of Prostate Cancer in Elderly Men. *JAMA Oncology*. 2019;5(1):60-6.
10. Mottet N, Cornford P, van den Bergh RC, Briers E, De Santis M, Fanti S, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Arnhem, The Netherlands.: EAU Guidelines Office; 2020 [Available from: <https://uroweb.org/guideline/prostate-cancer/#1>].
11. The American Cancer Society. Signs and Symptoms of Prostate Cancer. 2019 [Available from: <https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/signs-symptoms.html>].
12. Thüroff JW. Operative Anatomie der Prostata. In: Michel MS, Thüroff JW, Janetschek G, Wirth M, editors. *Die Urologie*. 2. Berlin, Heidelberg: Springer-Verlag; 2016. p. 1121.
13. McNeal JE. The zonal anatomy of the prostate. *Prostate*. 1981;2(1):35-49.
14. Andreoiu M, Cheng L. Multifocal prostate cancer: biologic, prognostic, and therapeutic implications. *Hum Pathol*. 2010;41(6):781-93.
15. Liu W, Laitinen S, Khan S, Vihinen M, Kowalski J, Yu G, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat Med*. 2009;15(5):559-65.
16. Stamey TA, McNeal JM, Wise AM, Clayton JL. Secondary cancers in the prostate do not determine PSA biochemical failure in untreated men undergoing radical retropubic prostatectomy. *Eur Urol*. 2001;39 Suppl 4:22-3.
17. Ahmed HU. The index lesion and the origin of prostate cancer. *N Engl J Med*. 2009;361(17):1704-6.
18. Humphrey PA. Gleason grading and prognostic factors in carcinoma of the prostate. *Mod Pathol*. 2004;17(3):292-306.
19. Thompson I, Thrasher JB, Aus G, Burnett AL, Canby-Hagino ED, Cookson MS, et al. Guideline for the management of clinically localized prostate cancer: 2007 update. *The Journal of urology*. 2007;177(6):2106-31.
20. Matoso A, Epstein JI. Defining clinically significant prostate cancer on the basis of pathological findings. *Histopathology*. 2019;74(1):135-45.
21. Downing A, Wright P, Hounsborne L, Selby P, Wilding S, Watson E, et al. Quality of life in men living with advanced and localised prostate cancer in the UK: a population-based study. *The Lancet Oncology*. 2019;20(3):436-47.

22. Grossman DC, Curry SJ, Owens DK, Bibbins-Domingo K, Caughey AB, Davidson KW, et al. Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement. *Jama*. 2018;319(18):1901-13.
23. Srivastava S, Koay EJ, Borowsky AD, De Marzo AM, Ghosh S, Wagner PD, et al. Cancer overdiagnosis: a biological challenge and clinical dilemma. *Nat Rev Cancer*. 2019;19(6):349-58.
24. Gandaglia G, Albers P, Abrahamsson PA, Briganti A, Catto JWF, Chapple CR, et al. Structured Population-based Prostate-specific Antigen Screening for Prostate Cancer: The European Association of Urology Position in 2019. *Eur Urol*. 2019.
25. Ilic D, Djulbegovic M, Jung JH, Hwang EC, Zhou Q, Cleves A, et al. Prostate cancer screening with prostate-specific antigen (PSA) test: a systematic review and meta-analysis. *Bmj*. 2018;362:k3519.
26. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet*. 2014;384(9959):2027-35.
27. Wirth M, Fröhner M. Prostatakarzinom: Epidemiologie, Ätiologie, Prävention und Früherkennung/Screening. In: Michel MS, Thüroff JW, Janetschek G, Wirth M, editors. *Die Urologie*. 2. Berlin Heidelberg: Springer-Verlag; 2016. p. 1301.
28. Djavan B, Margreiter M. Biopsy standards for detection of prostate cancer. *World J Urol*. 2007;25(1):11-7.
29. Berglund RK, Masterson TA, Vora KC, Eggener SE, Eastham JA, Guillonneau BD. Pathological upgrading and up staging with immediate repeat biopsy in patients eligible for active surveillance. *The Journal of urology*. 2008;180(5):1964-7; discussion 7-8.
30. Porten SP, Whitson JM, Cowan JE, Cooperberg MR, Shinohara K, Perez N, et al. Changes in prostate cancer grade on serial biopsy in men undergoing active surveillance. *J Clin Oncol*. 2011;29(20):2795-800.
31. Epstein JI, Feng Z, Trock BJ, Pierorazio PM. Upgrading and downgrading of prostate cancer from biopsy to radical prostatectomy: incidence and predictive factors using the modified Gleason grading system and factoring in tertiary grades. *Eur Urol*. 2012;61(5):1019-24.
32. Gershman B, Dahl DM, Olumi AF, Young RH, McDougal WS, Wu CL. Smaller prostate gland size and older age predict Gleason score upgrading. *Urol Oncol*. 2013;31(7):1033-7.
33. Koksall IT, Ozcan F, Kadioglu TC, Esen T, Kilicaslan I, Tunc M. Discrepancy between Gleason scores of biopsy and radical prostatectomy specimens. *Eur Urol*. 2000;37(6):670-4.
34. Baumgartner EM, Porter KK, Nix JW, Rais-Bahrami S, Gordetsky JB. Detection of extraprostatic disease and seminal vesicle invasion in patients undergoing magnetic resonance imaging-targeted prostate biopsies. *Translational andrology and urology*. 2018;7(Suppl 4):S392-S6.
35. Thompson J, Lawrentschuk N, Frydenberg M, Thompson L, Stricker P. The role of magnetic resonance imaging in the diagnosis and management of prostate cancer. *BJU international*. 2013;112 Suppl 2:6-20.
36. van der Heide UA, Frantzen-Steneker M, Astreinidou E, Nowee ME, van Houdt PJ. MRI basics for radiation oncologists. *Clin Transl Radiat Oncol*. 2019;18:74-9.
37. Malayeri AA, Khouli RHE, Zaheer A, Jacobs MA, Corona-Villalobos CP, Kamel IR, et al. Principles and Applications of Diffusion-weighted Imaging in Cancer Detection, Staging, and Treatment Follow-up. *RadioGraphics*. 2011;31(6):1773-91.
38. Zehof B, Pickles M, Liney G, Gibbs P, Rodrigues G, Kraus S, et al. Correlation of diffusion-weighted magnetic resonance data with cellularity in prostate cancer. *BJU international*. 2009;103(7):883-8.
39. Pedler K, Kitzing YX, Varol C, Arianayagam M. The current status of MRI in prostate cancer. *Aust Fam Physician*. 2015;44(4):225-30.
40. Girometti R, Cereser L, Bonato F, Zuiani C. Evolution of prostate MRI: from multiparametric standard to less-is-better and different-is better strategies. *European radiology experimental*. 2019;3(1):5.
41. Nörenberg D, Solyanik O, Schlenker B, Magistro G, Ertl-Wagner B, Clevert DA, et al. [MRI of the prostate]. *Urologe A*. 2017;56(5):665-77.
42. Rouviere O, Moldovan PC. The current role of prostate multiparametric magnetic resonance imaging. *Asian Journal of Urology*. 2019;6(2):137-45.
43. Murphy G, Haider M, Ghai S, Sreeharsha B. The Expanding Role of MRI in Prostate Cancer. *American Journal of Roentgenology*. 2013;201(6):1229-38.
44. Schoots IG, Petrides N, Giganti F, Bokhorst LP, Rannikko A, Klotz L, et al. Magnetic Resonance Imaging in Active Surveillance of Prostate Cancer: A Systematic Review. *European Urology*. 2015;67(4):627-36.

45. Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *The Lancet*. 2017;389(10071):815-22.
46. Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. *New England Journal of Medicine*. 2018;378(19):1767-77.
47. Rouviere O, Puech P, Renard-Penna R, Claudon M, Roy C, Mege-Lechevallier F, et al. Use of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naive patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. *Lancet Oncol*. 2019;20(1):100-9.
48. Bratan F, Niaf E, Melodelima C, Chesnais AL, Souchon R, Mège-Lechevallier F, et al. Influence of imaging and histological factors on prostate cancer detection and localisation on multiparametric MRI: a prospective study. *European Radiology*. 2013;23(7):2019-29.
49. Thompson JE, Moses D, Shnier R, Brenner P, Delprado W, Ponsky L, et al. Multiparametric magnetic resonance imaging guided diagnostic biopsy detects significant prostate cancer and could reduce unnecessary biopsies and over detection: a prospective study. *J Urol*. 2014;192(1):67-74.
50. McClure TD, Margolis DJ, Reiter RE, Sayre JW, Thomas MA, Nagarajan R, et al. Use of MR imaging to determine preservation of the neurovascular bundles at robotic-assisted laparoscopic prostatectomy. *Radiology*. 2012;262(3):874-83.
51. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate MR guidelines 2012. *European radiology*. 2012;22(4):746-57.
52. American College of Radiology (ACR). Prostate Imaging - Reporting and Data System. Version 2.015. 16.08.2019. Available from: <https://www.acr.org/-/media/ACR/Files/RADS/PI-RADS/PIRADS-V2.pdf>.
53. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic Performance of Prostate Imaging Reporting and Data System Version 2 for Detection of Prostate Cancer: A Systematic Review and Diagnostic Meta-analysis. *Eur Urol*. 2017;72(2):177-88.
54. Futterer JJ, Briganti A, De Visschere P, Emberton M, Giannarini G, Kirkham A, et al. Can Clinically Significant Prostate Cancer Be Detected with Multiparametric Magnetic Resonance Imaging? A Systematic Review of the Literature. *Eur Urol*. 2015;68(6):1045-53.
55. Magnetta MJ, Donovan AL, Jacobs BL, Davies BJ, Furlan A. Evidence-Based Reporting: A Method to Optimize Prostate MRI Communications With Referring Physicians. *AJR Am J Roentgenol*. 2018;210(1):108-12.
56. Vargas HA, Hotker AM, Goldman DA, Moskowitz CS, Gondo T, Matsumoto K, et al. Updated prostate imaging reporting and data system (PIRADS v2) recommendations for the detection of clinically significant prostate cancer using multiparametric MRI: critical evaluation using whole-mount pathology as standard of reference. *European radiology*. 2016;26(6):1606-12.
57. Borofsky S, George AK, Gaur S, Bernardo M, Greer MD, Mertan FV, et al. What Are We Missing? False-Negative Cancers at Multiparametric MR Imaging of the Prostate. *Radiology*. 2018;286(1):186-95.
58. Woo S, Cho JY, Kim SY, Kim SH. Extracapsular Extension in Prostate Cancer: Added Value of Diffusion-Weighted MRI in Patients With Equivocal Findings on T2-Weighted Imaging. *American Journal of Roentgenology*. 2015;204(2):W168-W75.
59. Hansen NL, Koo BC, Warren AY, Kastner C, Barrett T. Sub-differentiating equivocal PI-RADS-3 lesions in multiparametric magnetic resonance imaging of the prostate to improve cancer detection. *European journal of radiology*. 2017;95:307-13.
60. Rosenkrantz AB, Meng X, Ream JM, Babb JS, Deng F-M, Rusinek H, et al. Likert score 3 prostate lesions: Association between whole-lesion ADC metrics and pathologic findings at MRI/ultrasound fusion targeted biopsy. *Journal of Magnetic Resonance Imaging*. 2016;43(2):325-32.
61. Mehralivand S, Bednarova S, Shih JH, Mertan FV, Gaur S, Merino MJ, et al. Prospective Evaluation of PI-RADS™ Version 2 Using the International Society of Urological Pathology Prostate Cancer Grade Group System. *The Journal of urology*. 2017;198(3):583-90.
62. Girometti R, Giannarini G, Greco F, Isola M, Cereser L, Como G, et al. Interreader agreement of PI-RADS v. 2 in assessing prostate cancer with multiparametric MRI: A study using whole-mount histology as the standard of reference. *Journal of magnetic resonance imaging : JMRI*. 2019;49(2):546-55.
63. Vargas HA, Akin O, Franiel T, Mazaheri Y, Zheng J, Moskowitz C, et al. Diffusion-weighted endorectal MR imaging at 3 T for prostate cancer: tumor detection and assessment of aggressiveness. *Radiology*. 2011;259(3):775-84.

64. deSouza NM, Riches SF, VanAs NJ, Morgan VA, Ashley SA, Fisher C, et al. Diffusion-weighted magnetic resonance imaging: a potential non-invasive marker of tumour aggressiveness in localized prostate cancer. *Clinical Radiology*. 2008;63(7):774-82.
65. Nowak J, Malzahn U, Baur AD, Reichelt U, Franiel T, Hamm B, et al. The value of ADC, T2 signal intensity, and a combination of both parameters to assess Gleason score and primary Gleason grades in patients with known prostate cancer. *Acta Radiol*. 2016;57(1):107-14.
66. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Dorey FJ, Walsh PC, et al. Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. *Jama*. 2005;294(4):433-9.
67. Roehl KA, Han M, Ramos CG, Antenor JA, Catalona WJ. Cancer progression and survival rates following anatomical radical retropubic prostatectomy in 3,478 consecutive patients: long-term results. *J Urol*. 2004;172(3):910-4.
68. Kupelian PA, Mahadevan A, Reddy CA, Reuther AM, Klein EA. Use of different definitions of biochemical failure after external beam radiotherapy changes conclusions about relative treatment efficacy for localized prostate cancer. *Urology*. 2006;68(3):593-8.
69. Ahdoot M, Lebastchi AH, Turkbey B, Wood B, Pinto PA. Contemporary treatments in prostate cancer focal therapy. *Current opinion in oncology*. 2019;31(3):200-6.
70. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016;278(2):563-77.
71. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48(4):441-6.
72. Aerts HJ. The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review. *JAMA Oncol*. 2016;2(12):1636-42.
73. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749-62.
74. Wang S, Burt K, Turkbey B, Choyke P, Summers RM. Computer aided-diagnosis of prostate cancer on multiparametric MRI: a technical review of current research. *Biomed Res Int*. 2014;2014:789561-.
75. Lemaitre G, Marti R, Freixenet J, Vilanova JC, Walker PM, Meriaudeau F. Computer-Aided Detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: a review. *Comput Biol Med*. 2015;60:8-31.
76. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*. 2014;5:4006.
77. Agner SC, Rosen MA, Englander S, Tomaszewski JE, Feldman MD, Zhang P, et al. Computerized image analysis for identifying triple-negative breast cancers and differentiating them from other molecular subtypes of breast cancer on dynamic contrast-enhanced MR images: a feasibility study. *Radiology*. 2014;272(1):91-9.
78. Ingrisich M, Schneider MJ, Norenberg D, Negrao de Figueiredo G, Maier-Hein K, Suchorska B, et al. Radiomic Analysis Reveals Prognostic Information in T1-Weighted Baseline Magnetic Resonance Imaging in Patients With Glioblastoma. *Investigative radiology*. 2017;52(6):360-6.
79. Ingrisich M, Schöppe F, Paprottka K, Fabritius M, Strobl FF, De Toni EN, et al. Prediction of 90Y Radioembolization Outcome from Pretherapeutic Factors with Random Survival Forests. *Journal of Nuclear Medicine*. 2018;59(5):769-73.
80. Kniep HC, Madesta F, Schneider T, Hanning U, Schönfeld MH, Schön G, et al. Radiomics of Brain MRI: Utility in Prediction of Metastatic Tumor Type. *Radiology*. 2019;290(2):479-87.
81. Cook GJR, Azad G, Owczarczyk K, Siddique M, Goh V. Challenges and Promises of PET Radiomics. *International Journal of Radiation Oncology\*Biophysics\*Physics*. 2018;102(4):1083-9.
82. Sun Y, Reynolds HM, Parameswaran B, Wraith D, Finnegan ME, Williams S, et al. Multiparametric MRI and radiomics in prostate cancer: a review. *Australasian physical & engineering sciences in medicine*. 2019;42(1):3-25.
83. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. *arXiv e-prints [Internet]*. 2016 December 01, 2016. Available from: <https://ui.adsabs.harvard.edu/abs/2016arXiv161207003Z>.
84. Mohri M, Rostamizadeh A, Talwalkar A. *Foundations of Machine Learning*: The MIT Press; 2012. 480 p.
85. Breiman L. Random Forests. *Machine Learning*. 2001;45(1):5-32.
86. Bonekamp D, Kohl S, Wiesenfarth M, Schelb P, Radtke JP, Gotz M, et al. Radiomic Machine Learning for Characterization of Prostate Lesions with MRI: Comparison to ADC Values. *Radiology*. 2018;289(1):128-37.

87. Fehr D, Veeraraghavan H, Wibmer A, Gondo T, Matsumoto K, Vargas HA, et al. Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proceedings of the National Academy of Sciences*. 2015;112(46):E6265.
88. Schelb P, Kohl S, Radtke JP, Wiesenfarth M, Kickingereder P, Bickelhaupt S, et al. Classification of Cancer at Prostate MRI: Deep Learning versus Clinical PI-RADS Assessment. *Radiology*. 2019;190938.
89. Wolf I, Vetter M, Wegner I, Bottger T, Nolden M, Schobinger M, et al. The medical imaging interaction toolkit. *Med Image Anal*. 2005;9(6):594-604.
90. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res*. 2017;77(21):e104-e7.
91. Haralick RM. Statistical and structural approaches to texture. *Proceedings of the IEEE*. 1979;67(5):786-804.
92. Galloway MM. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*. 1975;4(2):172-9.
93. Thibault G, Fertil B, Navarro C, Pereira S, Lévy N, Sequeira J, et al. Texture Indexes and Gray Level Size Zone Matrix Application to Cell Nuclei Classification 2009.
94. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *J Mach Learn Res*. 2012;13:281-305.
95. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12(1):77.
96. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29-36.
97. Greenfield ML, Kuhn JE, Wojtys EM. A statistics primer. Confidence intervals. *Am J Sports Med*. 1998;26(1):145-9.
98. Rosenkrantz AB, Ginocchio LA, Cornfeld D, Froemming AT, Gupta RT, Turkbey B, et al. Interobserver Reproducibility of the PI-RADS Version 2 Lexicon: A Multicenter Study of Six Experienced Prostate Radiologists. *Radiology*. 2016;280(3):793-804.
99. Bajgiran AM, Mirak SA, Sung K, Sisk AE, Reiter RE, Raman SS. Apparent Diffusion Coefficient (ADC) Ratio Versus Conventional ADC for Detecting Clinically Significant Prostate Cancer With 3-T MRI. *AJR Am J Roentgenol*. 2019;213(3):W134-w42.
100. Moraes MO, Roman DHH, Copetti J, de SSF, Agra A, Noronha JAP, et al. Effects of the addition of quantitative apparent diffusion coefficient data on the diagnostic performance of the PI-RADS v2 scoring system to detect clinically significant prostate cancer. *World J Urol*. 2019.
101. Chaddad A, Kucharczyk MJ, Niazi T. Multimodal Radiomic Features for the Predicting Gleason Score of Prostate Cancer. *Cancers*. 2018;10(8).
102. Kwon D, Reis IM, Breto AL, Tschudi Y, Gautney N, Zavala-Romero O, et al. Classification of suspicious lesions on prostate multiparametric MRI using machine learning. *Journal of medical imaging (Bellingham, Wash)*. 2018;5(3):034502.
103. Orczyk C, Villers A, Rusinek H, Lepennec V, Bazille C, Giganti F, et al. Prostate cancer heterogeneity: texture analysis score based on multiple magnetic resonance imaging sequences for detection, stratification and selection of lesions at time of biopsy. *BJU international*. 2018.
104. Algohary A, Viswanath S, Shiradkar R, Ghose S, Pahwa S, Moses D, et al. Radiomic features on MRI enable risk categorization of prostate cancer patients on active surveillance: Preliminary findings. *Journal of magnetic resonance imaging : JMRI*. 2018.
105. Chen T, Li M, Gu Y, Zhang Y, Yang S, Wei C, et al. Prostate Cancer Differentiation and Aggressiveness: Assessment With a Radiomic-Based Model vs. PI-RADS v2. *Journal of magnetic resonance imaging : JMRI*. 2019;49(3):875-84.
106. Kamrava M, Kishan AU, Margolis DJ, Huang J, Dorey F, Lieu P, et al. Multiparametric magnetic resonance imaging for prostate cancer improves Gleason score assessment in favorable risk prostate cancer. *Pract Radiat Oncol*. 2015;5(6):411-6.
107. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust Radiomics Feature Quantification Using Semiautomatic Volumetric Segmentation. *PLOS ONE*. 2014;9(7):e102107.
108. Porter CR, Kodama K, Gibbons RP, Correa R, Jr., Chun FK, Perrotte P, et al. 25-year prostate cancer control and survival outcomes: a 40-year radical prostatectomy single institution series. *The Journal of urology*. 2006;176(2):569-74.
109. Noone AM, N. H, M. K, D. M, A. B, M. Y, et al. SEER Cancer Statistics Review, 1975-2015, National Cancer Institute Bethesda, MD 2019 [Available from: <https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/survival-rates.html#references>].
110. Padhani AR, Weinreb J, Rosenkrantz AB, Villeirs G, Turkbey B, Barentsz J. Prostate Imaging-Reporting and Data System Steering Committee: PI-RADS v2 Status Update and Future Directions. *Eur Urol*. 2019;75(3):385-96.

111. Padhani AR, Barentsz J, Villeirs G, Rosenkrantz AB, Margolis DJ, Turkbey B, et al. PI-RADS Steering Committee: The PI-RADS Multiparametric MRI and MRI-directed Biopsy Pathway. *Radiology*. 2019;0(0):182946.
112. Gupta RT, Mehta KA, Turkbey B, Verma S. PI-RADS: Past, present, and future. *Journal of magnetic resonance imaging : JMRI*. 2019.
113. Greer MD, Lay N, Shih JH, Barrett T, Bittencourt LK, Borofsky S, et al. Computer-aided diagnosis prior to conventional interpretation of prostate mpMRI: an international multi-reader study. *European radiology*. 2018;28(10):4407-17.
114. Gaur S, Lay N, Harmon SA, Doddakashi S, Mehralivand S, Argun B, et al. Can computer-aided diagnosis assist in the identification of prostate cancer on prostate MRI? a multi-center, multi-reader investigation. *Oncotarget*. 2018;9(73):33804-17.
115. Zhang M, Milot L, Khalvati F, Sugar L, Downes M, Baig SM, et al. Value of Increasing Biopsy Cores per Target with Cognitive MRI-targeted Transrectal US Prostate Biopsy. *Radiology*. 2019;291(1):83-9.
116. Yim JH, Kim CK, Kim JH. Clinically insignificant prostate cancer suitable for active surveillance according to Prostate Cancer Research International: Active surveillance criteria: Utility of PI-RADS v2. *Journal of magnetic resonance imaging : JMRI*. 2018;47(4):1072-9.
117. Scialpi M, D'Andrea A, Martorana E, Malaspina CM, Aisa MC, Napoletano M, et al. Biparametric MRI of the prostate. *Turk J Urol*. 2017;43(4):401-9.
118. (ACR) ACoR. Prostate Imaging - Reporting and Data System. Version 2.12019 05.11.2019. Available from: <https://www.acr.org/-/media/ACR/Files/RADS/Pi-RADS/PIRADS-V2-1.pdf?la=en>.
119. Padhani AR, Turkbey B. Detecting Prostate Cancer with Deep Learning for MRI: A Small Step Forward. *Radiology*. 2019;293(3):618-9.
120. Wang J, Wu CJ, Bao ML, Zhang J, Wang XN, Zhang YD. Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer. *European radiology*. 2017;27(10):4082-90.
121. Zhang Y, Chen W, Yue X, Shen J, Gao C, Pang P, et al. Development of a Novel, Multi-Parametric, MRI-Based Radiomic Nomogram for Differentiating Between Clinically Significant and Insignificant Prostate Cancer. *Front Oncol*. 2020;10:888.
122. Woźnicki P, Westhoff N, Huber T, Riffel P, Froelich MF, Gresser E, et al. Multiparametric MRI for Prostate Cancer Characterization: Combined Use of Radiomics Model with PI-RADS and Clinical Parameters. *Cancers*. 2020;12(7):1767.
123. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol*. 2016;69(1):16-40.
124. Algohary A, Shiradkar R, Pahwa S, Purysko A, Verma S, Moses D, et al. Combination of Peri-Tumoral and Intra-Tumoral Radiomic Features on Bi-Parametric MRI Accurately Stratifies Prostate Cancer Risk: A Multi-Site Study. *Cancers (Basel)*. 2020;12(8).
125. Viswanath SE, Chirra PV, Yim MC, Rofsky NM, Purysko AS, Rosen MA, et al. Comparing radiomic classifiers and classifier ensembles for detection of peripheral zone prostate tumors on T2-weighted MRI: a multi-site study. *BMC medical imaging*. 2019;19(1):22.
126. Merisaari H, Taimen P, Shiradkar R, Ettala O, Pesola M, Saunavaara J, et al. Repeatability of radiomics and machine learning for DWI: Short-term repeatability study of 112 patients with prostate cancer. *Magn Reson Med*. 2019.
127. Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al. Metrology Standards for Quantitative Imaging Biomarkers. *Radiology*. 2015;277(3):813-25.
128. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA: a cancer journal for clinicians*. 2019;69(2):127-57.
129. Turkbey B, Choyke PL. Future Perspectives and Challenges of Prostate MR Imaging. *Radiol Clin North Am*. 2018;56(2):327-37.
130. McGarry SD, Bukowy JD, Iczkowski KA, Unteriner JG, Duvnjak P, Lowman AK, et al. Gleason Probability Maps: A Radiomics Tool for Mapping Prostate Cancer Likelihood in MRI Space. *Tomography (Ann Arbor, Mich)*. 2019;5(1):127-34.
131. Nagpal K, Foote D, Liu Y, Chen P-HC, Wulczyn E, Tan F, et al. Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *npj Digital Medicine*. 2019;2(1):48.
132. Sowalsky AG, Kissick HT, Gerrin SJ, Schaefer RJ, Xia Z, Russo JW, et al. Gleason Score 7 Prostate Cancers Emerge through Branched Evolution of Clonal Gleason Pattern 3 and 4. *Clin Cancer Res*. 2017;23(14):3823-33.

# Danksagung

Mein ganz besonderer Dank gilt

Prof. Dr. Jens Ricke, meinem Doktorvater und Direktor der Klinik und Poliklinik für Radiologie, für die Möglichkeit, meine Dissertation in diesem Fachbereich am Klinikum Großhadern durchführen zu können.

Prof. Dr. rer. nat. Michael Ingrisch, der mir als mein Betreuer für die physikalischen Aspekte vor allem im zweiten Abschnitt meiner Arbeit jederzeit hilfsbereit mit Rat und Tat zur Seite stand. Ohne seine tatkräftige Annahme und Unterstützung wäre diese Dissertation so nicht zustande gekommen.

Oberarzt Dr. med. Dominik Nörenberg, meinem Betreuer für die Aspekte der klinischen Radiologie, für die freundliche Überlassung des Themas und die wertvolle, schnelle Erreichbarkeit und Rückmeldung von Beginn an. Vor allem im ersten Abschnitt meiner Arbeit hatte er immer ein offenes Ohr und nahm sich kurzfristig für gemeinsame Besprechungen Zeit.

Dr. rer. nat. Balthasar Schachtner, für die geduldige Einführung in die Statistik mit R und die unzähligen Hilfestellungen und ausdauernden Erklärungen zu den physikalischen Details meines Themas. Bis auf die Segmentierung erfolgten die einzelnen Schritte der Radiomics-Analyse inklusive der statistischen Analysen in der Abteilung für Clinical Data Science der Klinik und Poliklinik für Radiologie des Klinikums der Universität München durch Dr. rer. nat. Balthasar Schachtner unter Leitung von Prof. Dr. rer. nat. Michael Ingrisch, die bei dieser Prostatakarzinomstudie sehr eng mit uns zusammenarbeiteten.

Dr. Eva Gresser, die mich, wann immer es sich einrichten ließ, in allen radiologischen Fragestellungen tatkräftig unterstützte.

Meinen Eltern, Dr. Elisabeth Schreier und Dr. Axel Schreier, beides passionierte Mediziner, die den Grundstein für meine Arbeit legten und mir die Leidenschaft der Medizin weitergaben. Ganz besonderen Dank an dieser Stelle an meine Mutter, die mir durchgehend liebevollen Beistand geleistet hat und sich die Zeit nahm, mich auch mit wertvollen fachlichen Tipps zu unterstützen.

Meinem Freund Till Hertwig, der mich durch alle Hoch- und Tiefphasen geduldig begleitet hat und sich auch konstruktiv mit seiner fachlichen IT-Expertise eingebracht hat.

# Eidesstattliche Versicherung

Ich erkläre hiermit des Eides statt,  
dass ich die vorliegende Dissertation mit dem Thema

## **Detektion von Imaging Biomakern zur Erstellung eines prädiktiven Radiomics-Profiles zur Charakterisierung des Prostatakarzinoms**

selbstständig verfasst, mich außer der angegeben keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 05.05.2022

Andrea Schreier

Ort, Datum

Unterschrift Doktorandin