

Universität Ulm  
Medizinische Fakultät

Querschnittsstudie MRT zum Staging von Rektum-  
karzinomen:  
Validierung mit histopathologischen  
Befunden

Dissertation zur Erlangung des Doktorgrades der Medizin  
der Medizinischen Fakultät der Universität Ulm

vorgelegt von

**David Schäfer**

aus Dachau

2020

Amtierender Dekan: Prof. Dr. Thomas Wirth  
1. Berichterstatter: Prof. Dr. Andrea Rieber-Brambs  
2. Berichterstatter: PD Dr. Frank Leithäuser  
Tag der Promotion: 18.02.2022

## Inhaltsverzeichnis

<b>1. Einleitung</b> .....	<b>1</b>
<b>2. Material und Methoden</b> .....	<b>4</b>
2.1. Art der Studie .....	4
2.2. Patientenkollektiv .....	4
2.3. Dateninhalt.....	6
2.4. MRT-Untersuchung.....	11
2.5. Statistische Auswertung .....	20
<b>3. Ergebnisse</b> .....	<b>24</b>
3.1. Demographische Daten.....	24
3.2. T-Stadien.....	25
3.3. CRM-Status.....	35
3.4. N-Status .....	40
3.5. EMVI .....	41
3.6. Zusammenfassung der wichtigsten Ergebnisse .....	42
<b>4. Diskussion</b> .....	<b>44</b>
4.1. Demographische Daten.....	44
4.2. T-Stadien.....	44
4.3. CRM-Status.....	52
4.4. N-Status .....	65
4.5. EMVI .....	72
4.6. Schlussfolgerungen aus der Diskussion.....	75
<b>5. Zusammenfassung</b> .....	<b>79</b>
<b>6. Literaturverzeichnis</b> .....	<b>81</b>
Danksagung .....	87
Lebenslauf.....	88

## Abkürzungen

A	Arterie
Anm. d.	Anmerkung des
Ca.	(als Substantiv) Karzinom
Cor.	Koronar (Schichtebene einer MRT-Untersuchung)
CRM	Circumferential resection margin (chirurgischer zirkumferenzieller Resektionsrand bei Rektumkarzinomen)
CT	Computertomographie
DTI	Diffusion tensor imaging (Bildgebungsverfahren der MRT)
DOR	Diagnostic odds ratio (Statistische Kennzahl)
EMVI	Extramurale venöse Tumorinvasion (Tumorinfiltration von Venen außerhalb der Muskularis propria)
Fa.	Firma
FOV	Field-of-view (Untersuchungsfeld einer MRT-Untersuchung)
Fs	Fett gesättigt
HR	High resolution (hochauflösend)
I.v.	Intravenös
KIS	Krankenhausinformationssystem
KM	Kontrastmittel
LK	Lymphknoten
M.	Musculus (Muskel)
MRT	Magnetresonanztomographie
NPW	Negativ prädiktiver Wert
PACS	Picture Archiving and Communication System (Radiologisches Bildspeicher- u. Bildverwaltungssystem)
PPW	Positiv prädiktiver Wert
RIS	Radiologie-Informationssystem
SPIR	Spectral Presaturation with Inversion Recovery (MRT-Untersuchungs-Technik mit Fettunterdrückung)
TME	Totale mesorektale Exzision (standardisiertes chirurgisches Operationsverfahren zur kontinenserhaltenden Behandlung von Krebserkrankungen)
TNM	Facetten-Klassifikation zum Staging von Tumorerkrankungen (T= Tumor N= Nodus/Lymphknoten, M= Metastasen)

TSE	Turbo-Spin-Echo (MRT-Untersuchungstechnik)
V	Vene
V.a.	Verdacht auf
VEGF	Vascular Endothelial Growth Factor (vaskulärer endothelialer Wachstumsfaktor)

## **1. Einleitung**

Ca. 60.000 Menschen in Deutschland erkranken jährlich an einer Krebserkrankung des Dickdarms (Kolon) bzw. des Mastdarms (Rektum), etwa jede 8. Krebserkrankung betrifft somit das Kolorektum. Darmkrebs ist in Deutschland die zweithäufigste Todesursache unter den Krebserkrankungen. Der Anteil der Rektumkarzinome an den kolorektalen Karzinomen beträgt etwas weniger als ein Drittel, jedoch ist das Rektum nach dem Kolon der zweithäufigste Entstehungsort eines bösartigen Tumors des Magen-Darmtraktes [24,26].

Die Prognose eines Patienten mit Rektumkarzinom ist neben einer möglichst frühzeitigen Erkennung und Behandlung der Erkrankung von vielen weiteren Faktoren abhängig. Neben der Bildung von Tochtergeschwülsten in anderen Organen (Fernmetastasen) bestimmt der Krankheitsrückfall am selben Ort (Lokalrezidiv) wesentlich den Heilungsverlauf. Seit Einführung standardisierter Operationstechniken wie der TME (totale mesorektale Exzision) konnte das Lokalrezidivrisiko bei Rektumkarzinomen auf ca. 10% gesenkt werden. Auch der Einsatz einer Strahlentherapie oder kombinierten Strahlen- und Chemotherapie, die präoperativ bei fortgeschrittenem Tumorwachstum eingesetzt wird (neoadjuvante Therapie), zählt zu den modernen Therapieverfahren, die die Chancen einer erfolgreichen operativen Entfernung des Tumors durch Tumorschrumpfung und/oder die Langzeitprognose weiter verbessern können, allerdings ist deren Einsatz mit einem Risiko möglicher Spätfolgen verbunden, u.a. sexuelle Dysfunktion, Inkontinenz, Infertilität, Entstehung von Zweittumoren oder kardiovaskuläre Erkrankungen. Das einzige Therapieverfahren jedoch, mit dem Darmkrebs potenziell vollständig geheilt werden kann, ist nach wie vor die primäre Operation. Sie stellt daher die wichtigste Behandlungsmethode dar [25,43].

Bei fortgeschrittenem Tumorwachstum besteht die Schwierigkeit der diagnostischen Entscheidung eben genau darin, eine möglichst risikoarme und gleichzeitig nebenwirkungsarme Therapieform zu wählen, im Idealfall die primär kurative Tumorsektion. Hierbei ist es in der prätherapeutischen Ausbreitungsdiagnostik von maßgeblicher Bedeutung, diejenigen Tumorstadien zu identifizieren, die noch einer schonenden primären Operation zugeführt werden können, um den Einsatz von Radio- und/oder Chemotherapie möglichst zu minimieren, ohne dabei jedoch die Langzeitrisiken für ein Lokalrezidiv oder eine Fernmetastasierung signifikant zu erhöhen. Bei fortgeschrittenen Tumorstadien gewinnt in der Ausbreitungsdiagnostik die Evaluation des

Tumorwachstums in Bezug auf die mesorektale Faszie und der damit verbundenen möglichen Therapieoptionen zunehmend an Bedeutung, insbesondere in der Kontrolle des Risikos eines möglichen Lokalrezidivs. Für Fragestellungen, die eine Darstellung der mesorektalen Faszie und die Relation des Tumorwachstums zu ihr erfordern, weist die MRT aktuell die höchste Sensitivität auf und wird insbesondere deswegen von vielen Experten für das lokale Staging des fortgeschrittenen Rektumkarzinoms bevorzugt und in aktuellen Leitlinien der medizinischen Fachgesellschaften empfohlen [28,43,52].

Im Klinikum Neuperlach, als Haus der Maximalversorgung mit onkologischem Schwerpunkt und zertifiziertem Darmzentrum hat sich die MRT neben anderen diagnostischen Optionen wie z.B. dem endoskopischen Ultraschall, als standardisiertes Verfahren zum präoperativen Staging kolorektaler Karzinome etabliert, da sie ein breites Spektrum an diagnostischen Möglichkeiten bietet, die für multimodale Therapieentscheidungen hoch relevant sind. Über die Klassifizierung der Tumorausbreitung nach dem TNM-System hinaus wird die anatomische Darstellung des Lokalbefundes und dessen relevanter Risikostrukturen mittels MRT z.B. im Rahmen von radiologischen Falldemonstrationen aufgrund des hohen intrinsischen Weichteilkontrastes von den Operateuren für die Planung chirurgischer Resektionsebenen operativer Verfahren wie der TME sehr geschätzt. Neben der Evaluation der Tumorausbreitung in Bezug auf die mesorektale Faszie, deren hoher Stellenwert in großen prospektiven Multicenter-Studien wie der MERCURY-Studie oder der OCUM-Studie, an letzterer auch das Klinikum Neuperlach teilgenommen hat, nachgewiesen werden konnte, können in der MRT-Diagnostik auch weitere Hochrisikofaktoren wie z.B. eine extramurale venöse Tumorausbreitung identifiziert und beurteilt werden, die nach der Kontrolle des Lokalrezidivrisikos zunehmende prognostische und therapeutische Bedeutung hinsichtlich des Risikos einer metachronen Fernmetastasierung erlangen [4,5,6,13,43,50,52].

In Kombination moderner diagnostischer und chirurgischer Verfahren und interdisziplinärer Zusammenarbeit, wie sie im Klinikum Neuperlach in onkologischen Konferenzen realisiert wird, ist für Patienten ein zunehmend individuelles, risikoadaptiertes Therapieschema möglich, um ein möglichst gutes Outcome zu erzielen. Hierbei leistet die multiparametrische MRT-Diagnostik einen wichtigen Beitrag und nimmt unter allen diagnostischen Optionen eine herausragende Stellung ein. Entscheidend ist neben den technisch-diagnostischen Möglichkeiten jedoch nicht zuletzt die Expertise vor Ort [28].

Für eine Standortbestimmung der radiologischen Diagnostik von Rektumkarzinomen im Klinikum Neuperlach wurden daher als Qualitätssicherungsmaßnahme in der vorgelegten Studie die präoperativen MRT-Staging-Ergebnisse primär operativ therapierter Patienten mit Rektumkarzinom des Zeitraums 01.01.2008-31.12.2015 mit den postoperativen Ergebnissen der Pathologie als Goldstandard retrospektiv verglichen. Untersucht wurden die T-Stadien, der N-Status sowie das extramurale Tumorwachstum mit CRM-Status, Tumorinvasionstiefe in das perirektale Fettgewebe sowie eine etwaige extramurale venöse Tumorinvasion (EMVI).

Bezüglich der T-Stadien sowie der Bedrohung der mesorektalen Faszie (CRM-Status) erfolgte neben der Gesamtbetrachtung des Patientenkollektivs noch eine weitere Untergliederung abhängig von der Höhe der Tumorlokalisation mit der Fragestellung, ob und ggf. inwiefern sich die Tumorphöhe auf die Übereinstimmung der Ergebnisse mit der Pathologie auswirkte.

Folgende Fragestellungen sollen beantwortet werden:

1. Wie hoch ist die Genauigkeit des MRT-Stagings im Vergleich mit der Pathologie als Goldstandard (Gütekriterien)?
2. Wie ist die Übereinstimmung (Reliabilität) der Ergebnisse im Bereich der prognostisch besonders wichtigen T-Kategorie sowie des CRM-Status einzuschätzen? Hat die Tumorphöhe diesbzgl. Einfluss auf die Ergebnisse genommen?
3. Wie ist die diagnostische Performance im Vergleich mit der Literatur einzuschätzen?
4. War in der Analyse der Ergebnisse und der Literatur Potenzial für Verbesserungen für die Standarddiagnostik zu finden?

## **2. Material und Methoden**

### **2.1. Art der Studie**

In der hier vorliegenden unizentrischen Studie wurden retrospektiv 238 Patienten in die Studie eingeschlossen. Es wurden MRT-Untersuchungen ausgewertet, die im Zeitraum 01.01.2008 bis 31.12.2015 durchgeführt worden waren und mit den entsprechenden Datensätzen der Pathologie abgeglichen.

### **2.2. Patientenkollektiv**

Es wurden die Daten von Patienten untersucht, die sich in dem Zeitraum 01.01.2008 bis 31.12.2015 im Städtischen Klinikum Neuperlach einer MRT-Untersuchung des Unterbauchs unterzogen hatten. In die Studie eingeschlossen wurden Patienten mit histologisch gesichertem Rektumkarzinom, was aus den im KIS dokumentierten hausinternen Pathologiebefunden, den im KIS dokumentierten Befunden der pathologischen Konferenz, oder eventuell vorliegenden externen Befundberichten hervorging.

Ein Teilkollektiv der Studie waren Patienten aus der OCUM-Studie. Für alle Patienten wurden in der MR-Bildanalyse die Kriterien der OCUM-Studie angewandt (siehe 2.4.4.). Ausgeschlossen wurden Patienten, die vor der Operation eine Radio- und/oder Chemotherapie zur Reduktion der Tumormasse erhalten hatten, da aufgrund des veränderten Lokalbefundes somit ein unbeeinflusster Vergleich zwischen den Ergebnissen der MRT-Untersuchung und den postoperativ erhobenen Befunden der pathologischen Aufarbeitung nicht möglich gewesen wäre.

Ausgeschlossen wurden weiterhin Patienten, für die nach der erfolgten MRT-Staging-Untersuchung und der Indikationsstellung zur primären Operation keine hausinternen histopathologischen Ergebnisse evaluiert werden konnten, z.B. bei Ablehnung der Therapie oder Therapie in einem anderweitigen Klinikum oder einer sonstigen Behandlungseinrichtung.

## **2.2.1. Patientengesamtliste**

### **2.2.1.1. Daten aus dem RIS**

In dem radiologischen Informations- und Bildarchivierungssystem Centricity RIS-i 4.2 Plus der Fa. GE Healthcare wurde zunächst über die Filterfunktion eine Gesamtliste der Patienten erstellt, die eine MRT-Untersuchung des Abdomens erhalten hatten unter Eingabe des Suchwortes „MR Unterbauch“ sowie unter Einschränkung des Untersuchungszeitraumes auf das Intervall 01.01.2008-31.12.2015. Aus dieser Gesamtliste wurden nach manueller Durchsicht der MRT-Befunde diejenigen Patienten herausgefiltert, die eine Staging-Untersuchung bei V.a. kolorektales Karzinom erhalten hatten.

### **2.2.1.2. Datenabgleich mit dem KIS**

Basierend auf der Gesamtliste wurde anschließend im hausinternen KIS mittels des Softwaresystems der Fa. SAP in der Suchmaske „Dokumentenliste“ unter Eingabe des Namens und Geburtsdatums des Patienten der gesamte Datensatz des Patienten aufgerufen und überprüft, ob für den jeweiligen Patienten ein postoperativer histopathologischer Befund vorlag. Nur Patienten mit histologisch gesichertem Rektumkarzinom der Stadien T1-T4 wurden in die Studie aufgenommen, entsprechend die Gesamtliste aus dem RIS diesbzgl. angepasst, in der zunächst z.B. noch Patienten mit Sigmakarzinom enthalten waren.

Ebenfalls aus dem KIS konnte aus den chirurgischen Befundberichten das Operationsdatum ermittelt werden.

Nach Abgleich der Daten aus dem KIS und dem RIS ging die Patientengesamtliste mit den enthaltenen Basisdaten Name, Vorname, Geschlecht, Geburtsdatum und Operationsdatum in die Urliste ein.

## **2.3. Dateninhalt**

### **2.3.1. Demographische Daten**

#### **2.3.1.1. Alter**

Das Alter des Patienten zum Operationszeitpunkt wurde mithilfe der Formel =DATEDIF(x;z;"y") in Excel bestimmt, wobei „x“ dem im KIS hinterlegtem Operationsdatum, „z“ dem Geburtsdatum des Patienten entsprach, und das Ergebnis in Jahren „y“ aufgenommen.

#### **2.3.1.2. Geschlecht**

Des Weiteren wurde das Geschlecht des Patienten dokumentiert und hieraus der prozentuale Anteil von weiblichen und männlichen Patienten am Patientenkollektiv bestimmt.

#### **2.3.1.3. Zeitraum zwischen MRT-Staging und Operation**

Der Zeitraum zwischen dem Operationsdatum aus dem chirurgischen Befundbericht im KIS und der vorangegangenen MRT-Staging-Untersuchung wurde mittels der Formel (DATEDIF;x;y;"d") in Excel berechnet, wobei „x“ dem Tag der MR-Untersuchung, „y“ dem Operationsdatum entsprach, und das Ergebnis in Tagen „d“ angegeben.

### **2.3.2. Untersuchungsbezogene Daten der Pathologie**

#### **2.3.2.1. T-Stadien**

Aus der TNM-Klassifikation wurden die T-Stadien der kritischen pathologischen Beurteilung, eingeteilt in T1-T4, übernommen. Zusätzlich erfolgte nachträglich eine Subklassifikation der T3-Stadien in T3a/b sowie T3c/d.

**Tab. 1:** Tumor-Stadien der TNM-Klassifikation (TNM = Tumor/Lymphknoten/Metastasen) [27]

T-Tumor	Eigenschaften
T1	Tumor hat sich in die dünne Bindegewebsschicht unterhalb der Darmschleimhaut (Submukosa) ausgebreitet
T2	Tumor hat sich in die Muskelschicht unterhalb der Schleimhaut (Muskularis propria) ausgebreitet
T3	Tumor hat sich über die Muskelschicht unterhalb der Schleimhaut hinaus in die Subserosa oder das Fettgewebe ausgebreitet
T4	Tumor hat andere Organe und Strukturen oder das Bauchfell befallen

#### 2.3.2.2. N-Status

Ebenfalls wurden die N-Stadien aus der TNM-Klassifikation der pathologischen Befunde in die Auswertung übernommen. Die N-Stadien der Kategorien N1: N1a, N1b sowie N2: N2a, N2b wurden als Status „N-positiv“ in die Urliste übernommen, Unterkategorien wurden nicht explizit ausgewertet. N0-Stadien sowie das Stadium N1c, (letzteres wurde nur bei einem Patienten vergeben), das Tumorsatelliten in der Subserosa, des Mesenteriums, extraperitoneal parakolisch oder perirektal beschreibt, jedoch keine befallenen Lymphknoten, entsprechend als „N-negativ“ in die Auswertung aufgenommen. Ausgeschlossen von der Vergleichsbeurteilung wurde das Stadium Nx (4 Patienten), das inkomplette Informationen bezüglich des nodalen Status beschreibt.

**Tab. 2:** Lymphknoten-Stadien der TNM-Klassifikation (TNM = Tumor/Lymphknoten/Metastasen) [27]

N-Lymphknoten	Eigenschaften
Nx	Regionäre Lymphknoten können nicht beurteilt werden
N0	Kein Befall regionärer Lymphknoten
N1	Befall von 1-3 regionären Lymphknoten
N1a	1 Lymphknoten befallen
N1b	2-3 Lymphknoten befallen
N1c	Tumorknötchen im Fettgewebe unterhalb der Schleimhaut ohne regionäre Lymphknoten
N2	Befall von 4 oder mehr regionären Lymphknoten
N2a	4-6 Lymphknoten befallen
N2b	7 oder mehr Lymphknoten befallen

### 2.3.2.3. Extramurale Tumorinvasion

#### 2.3.2.3.1. Invasionstiefe in das perirektale Fettgewebe

Bei Tumoren, die die Muskularis propria überschritten, wurde die in den pathologischen Befundberichten beschriebene Tumorausdehnung in das perirektale Fettgewebe in mm übernommen. Die Invasionstiefe in mm wurde für T3-Stadien ausgewertet.

#### 2.3.2.3.2. EMVI

Venöse Gefäßeinbrüche, die in der Beschreibung der Tumorausdehnung im perirektalen Fettgewebe zu finden waren, wurden entsprechend der Auswertungskategorie EMVI als „EMVI-positiv“ gewertet.

#### 2.3.2.3.3. CRM-Status

Ausgewertet wurde das Tumorwachstum hinsichtlich einer Infiltration oder Bedrohung der mesorektalen Faszie oder dem zirkumferenziellen Resektionsrand in der Kategorie „CRM-Status“. Als Status „CRM positiv“ wurde festgelegt, wenn dies entweder aus dem pathologischen Befundbericht wörtlich hervorging, oder das Tumorwachstum lt. Pathologiebefund direkt an den zirkumferenziellen chirurgischen Resektionsrand heranreichte, an diesem randbildend war, oder von diesem  $\leq 1$  mm entfernt blieb.

Bei allen anderen Patienten, deren Tumorwachstum  $>1$  mm vom zirkumferenziellen Resektionsrand entfernt blieb, wurde der CRM-Status als „CRM-negativ“ gewertet.

### 2.3.3. Untersuchungsbezogene Daten des präoperativen MRT-Stagings

Ausgewertet wurden die T-Stadien, der N-Status sowie eine extramurale Tumoringvasion. Die hierfür zu Grunde gelegten Kriterien der MRT-Bildanalyse sind unter 2.4.4 im Detail beschrieben.

Die MRT-Staging-Untersuchungen wurden jeweils an einem zertifizierten Befundungsmonitor im RIS analysiert und i.d.R. von einem erfahrenen Oberarzt verantwortlich gezeichnet (Befunder). Befundbeurteilungen wurden primär von insgesamt 12 Befundern erstellt.

Bei Unsicherheit bezüglich des T-Stadiums in der Befundbeurteilung des Originärbefundes wurden die betreffenden Untersuchungen retrospektiv im Jahr 2019 durch einen zweiten Befunder reevaluiert - einer erfahrenen Oberärztin, die regelmäßig an Spezial-Schulungen und Fortbildungen zum Thema MRT-Staging des Rektumkarzinoms teilnimmt- und sich hierdurch letztlich auf ein definitives T-Stadium festgelegt, um eine Vergleichsbeurteilung mit der T-Klassifikation der Pathologie durchführen zu können. Dies war bei 26 Patienten der Fall.

Bei der Tumoreindringtiefe in das perirektale Fettgewebe der Stadien  $> T2$  wurden in den originären Befundbeurteilungen bei insgesamt 11 Patienten entweder keine Angaben oder lediglich Intervallangaben (z.B.  $< 5$  mm) gemacht. Auch diese Untersuchungen wurden durch den zweiten Befunder s.o. reevaluiert, eine definitive Angabe in mm ermittelt und entsprechend gewertet. Auch wurden CRM-positive Befunde der anatomischen Zone A (siehe 2.4.4.1.1.) reevaluiert, die diesbzgl.

abweichenden Ergebnisse wurden jedoch lediglich diskutiert, ohne die ursprüngliche Datenbasis zu verändern.

#### **2.3.3.1. T-Stadien**

Die T-Stadien wurden analog der TNM-Klassifikation der Pathologie in T1 bis T4 eingeteilt. Zusätzlich erfolgte eine Subklassifikation der T3-Stadien in die Kategorien T3a/b sowie T3c/d.

#### **2.3.3.2. N-Status**

Analog zu der Auswertung der N-Stadien der Pathologie gingen Untersuchungen, in welchen keine suspekten Lymphknoten befundet wurden, mit dem Status „N-negativ“ in die Auswertung ein, Untersuchungen mit  $\geq 1$  suspekten Lymphknoten entsprechend mit dem Status „N-positiv“. Die diesbzgl. Angaben wurden aus den Befundbeurteilungen sowie den Eintragungen aus dem im RIS integrierten standardisierten Erhebungsbogen (Formatvorlage) entnommen.

#### **2.3.3.3. Extramurale Tumorinvasion**

##### **2.3.3.3.1. Invasionstiefe in das perirektale Fettgewebe**

Bei Untersuchungen, in welchen das Tumorwachstum über die Muskularis propria hinausging, wurden die Eintragungen der Infiltrationstiefe in mm dem standardisierten Erhebungsbogen entnommen. Die Invasionstiefe in mm wurde für T3-Stadien ausgewertet.

##### **2.3.3.3.2. EMVI**

Die Angaben bzgl. einer befundlichen EMVI (extramurale venöse Invasion) wurden dem standardisierten Erhebungsbogen (Formatvorlage) im RIS entnommen.

#### 2.3.3.3.3. CRM-Status

Der CRM-Status wurde aus den Befundbeurteilungen im RIS entnommen und eingeteilt in „CRM-positiv“ bzw. „CRM-negativ“ ausgewertet.

#### 2.3.3.4. Anatomische Zone

Die Höhenlokalisierung des Tumorwachstums im oberen, mittleren und unteren Rektumdrittel wurde anhand von anatomischen Landmarken der MRT-Bildanalyse in 3 Zonen (A-C) eingeteilt, die unter 2.4.4. mit den zugrundeliegenden Kriterien der Bildanalyse detailliert beschrieben sind.

### 2.4. MRT-Untersuchung

Die MRT-Untersuchungen wurden an einem 1,5 Tesla Scanner „Intera“ der Firma Philips Healthcare, Niederlande, durchgeführt.

#### 2.4.1. Patientenvorbereitung

Nach ausführlichem Gespräch mit dem Patienten, Ausschluss von Kontraindikationen hinsichtlich einer MRT-Untersuchung sowie Einwilligung des Patienten in die Untersuchung und die Auswertung seiner Ergebnisse im Rahmen dieser Studie, wurde die Untersuchung folgendermaßen durchgeführt: Die Positionierung des Patienten erfolgte in Rückenlage. Als Empfangsspule wurde eine Phased-array-Oberflächenspule eingesetzt. Es wurden sowohl T2-gewichtete, als auch T1-gewichtete Untersuchungsreihen, letztere vor und nach intravenöser Kontrastmittelgabe von 0,1 mmol/kg Gd-DOTA (Dotarem) angefertigt, sofern keine Kontraindikationen für eine Kontrastmittelapplikation seitens des Patienten, wie eine schwere Niereninsuffizienz oder allergische Reaktionen bei ggf. vorangegangenen KM-gestützten MRT-Untersuchungen, vorlagen.

Zusätzlich wurden zur Verminderung von Bewegungsartefakten durch Darmperistaltik 20 mg Butylscopolamin i.v. verabreicht, sofern keine Kontraindikationen seitens des Patienten, wie Engwinkelglaukom, schwere Herz-Rhythmus-Störungen oder Prostatahypertrophie, vorlagen. Alternativ wurde Glucagon verabreicht, sofern keine Diabeteserkrankung bekannt war.

Ansonsten galten die üblichen Kontraindikationen einer MRT-Untersuchung.

### 2.4.2. Untersuchungsprotokoll

Auf eine rektale Füllung wurde verzichtet.

Zur Beurteilung des Lokalbefundes wurden folgende Sequenzen des Beckens angefertigt:

**T2 TSE sagittal**, hochauflösend (**HR**, max. 3 mm Schichtdicke) mit vollständiger Abdeckung des Tumorbereichs.

**T2 TSE HR** kontinuierlich (kein gap), perpendicular (senkrecht) zu Tumurlängsachse eingekippt.

T2 TSE nativ (Standard) transversal des gesamten Beckens.

T2 TSE koronar, anguliert auf die entspr. Rektumachse, bei Tumoren des unteren Drittels auch in der Sphinkterebene obligat.

T1 TSE transversal nativ, fakultativ.

T1 SPIR transversal/koronar fakultativ nach i.v. KM-Applikation (Dotarem).

Richtgrößen für die Ortsauflösung bei HR-Sequenzen (Voxelgröße):

T2 TSE transversal:

FOV 200 x 200 Matrix 328 x 512 Voxel 0,6 x 0,4 x 4,0 cm

T2 TSE sagittal/koronar:

FOV 200 x 220 Matrix 328 x 512 Voxel 0,7 x 0,4 x 4,0 cm

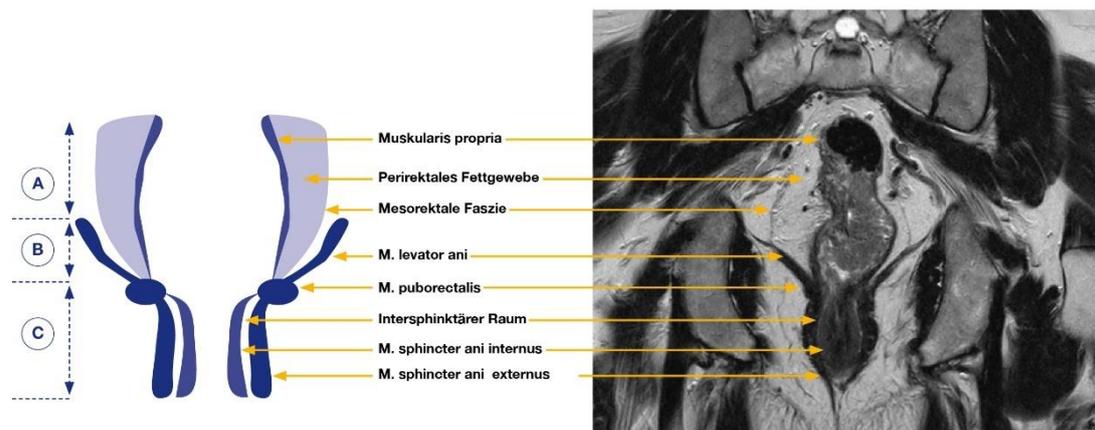
### 2.4.3. Rechtfertigende Indikation

Die rechtfertigende Indikation erging aus dem klinischen Kontext der Abklärung bei V.a. Rektumkarzinom, z.B. im Rahmen von suspekten Befunden aus Vorsorgeuntersuchungen oder Befunden aus anderen Untersuchungsmodalitäten wie der CT und wurde entsprechend von einem Radiologen gestellt.

## 2.4.4. MR-Bildanalyse, Kriterien und Beispiele

### 2.4.4.1. Definition der anatomische Zonen A-C

Die Einordnung der Tumorausdehnung in hoch sitzende, mittel- und tief sitzende Tumoren erfolgte anhand einer Befundvorlage im RIS (Formatvorlage) in definierte Zonen A bis C, die sich jeweils an Leitstrukturen der MRT-Anatomie orientieren. Hierfür wurde das am weitesten aboral gelegene Tumorsignal in Relation zu den genannten Leitstrukturen gesetzt und erfolgte im Wesentlichen in der Analyse der koronaren Bildserien.



**Abb. 1:** Anatomische Zonen und Leitstrukturen. Rechts: Beispiel eines Rektumkarzinoms der Kernspintomographie in der aboralen Ausdehnung in Relation zur Sphinkterebene. T2W-Sequenz koronar. Links: Schema (Eigenerstellung). Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

#### 2.4.4.1.1. Zone A, hoch

Die Zone A beschreibt das Tumorwachstum oberhalb des Ursprungs des M. levator ani:

Die oberen 2/3 des Rektums sind von viszeralem Peritoneum umgeben (intraperitoneal), anterolateral in den oberen Abschnitten, anterior in den tieferen Abschnitten [8]. Der tiefste Punkt der Insertion der peritonealen Umschlagfalte, auf Höhe der „Kohlrauschfalte“, erscheint in den axialen Schichtbildern oft „schwalbenschwanzartig“ (seagull sign) als dünne, hypointense Linie in den T2-gewichteten Sequenzen [27,8], in den sagittalen Bildserien meist zwischen der Harnblasenhinterwand und der Rektumvorderwand auf Höhe Oberrand der Samenbläschen, bzw. zwischen dem Übergang Corpus/Cervix uteri und ist bei etwaiger Infiltration relevant für die

Klassifikation eines T4-Stadiums [8]. Der unterhalb dieser Ebene gelegene Abschnitt der Rektumampulle liegt also extraperitoneal.

Auf dieser Höhe oder oral der peritonealen Umschlagfalte verläuft die mesorektale Faszie, die als feine hypodense Linie in den T2-gewichteten Sequenzen erkennbar ist, nicht vollständig zirkumferenziell, sondern ummantelt nur die posterioren und posterolateralen Anteile des perirektalen Fettgewebes [27].



**Abb. 2:** Peritoneale Umschlagfalte in der Kernspintomographie. Axiale T2W-Sequenz. „Seagull sign“, Bild aus PACS (Picture Archiving and Communication System). Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

#### 2.4.4.1.2. Zone B, mittlere Zone

Tumorwachstum vom Ursprung des M. levator ani nach aboral maximal bis zur Levatorinsertion:

Der M. levator ani läuft hier in der koronaren Ebene trichterförmig auf den anorektalen Übergang zu, bevor er auf Höhe der Levatorinsertion nahtlos in den tiefer gelegenen Abschnitten in den schlingenförmigen M. puborectalis bzw. in Verlängerung im Analkanal in den M. sphincter ani externus übergeht, begleitet von einem sich zunehmend verschmälernden Saum an perirektalem, T2 hyperintensiven Fettgewebe (Mesorektum), zirkumferenziell eingehüllt von der mesorektalen Faszie [8,44].

Auf Höhe des anorektalen Übergangs, der durch den Oberrand des M. puborectalis markiert ist, ist kein mesorektales Fettgewebe mehr erkennbar. Hier (ca. Höhe SWK 4), verschmilzt die mesorektale Faszie *dorsalseitig* der Rektumwand mit der Fascia pelvis parietalis, welche kranial hiervon mit dem zwischenliegenden, mit

Fettgewebe ausgefüllten retrorektalen Spaltraum, insbesondere in den sagittalen Bildserien von ihr unterscheidbar ist.

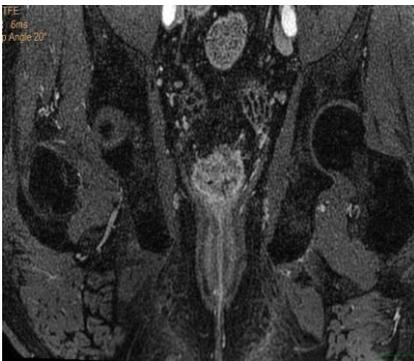
*Kaudal* verschmilzt sie mit der Fascia diaphragmatis pelvis superior, der Faszienbedeckung des M. levator ani, und endet am Oberrand des Sphinkterkomplexes. Tief anterior verschmilzt die mesorektale Faszie mit dem Septum rectogenitale. In den jeweiligen Verschmelzungszonen sind die Faszien in der MRT nicht mehr voneinander zu differenzieren [8].

#### 2.4.4.1.3. Zone C, tief, unteres Rektumdrittel

Tumorwachstum beginnend auf Höhe des Oberrandes des M. puborectalis nach aboral bis in die Sphinkterebene:

In den unterhalb des M. puborectalis gelegenen Abschnitten des Analkanals liegt der Sphinkterapparat mit dem M. sphincter ani externus und -internus, zwischen welchen in den T2-gewichteten Sequenzen ein schmaler hyperintenser Fettstreifen als Intersphinktärebene erkennbar ist. Der M. sphincter ani internus entspricht anatomisch der verbreiterten inneren Ringmuskelschicht der Muskularis propria und wird an seinem Oberrand von der Puborektalschlinge des M. puborectalis umgeben. In der Intersphinktärebene können neben dem Fettstreifen auch eingelagerte Muskelfasern des sog. „longitudinal anal muscle“ (LAM) erkennbar sein.

Der M. sphincter ani internus, bestehend aus glatter Muskulatur, reichert homogen und kräftig Kontrastmittel an, der M. sphincter ani externus, bestehend aus quergestreifter Muskulatur, ist hingegen nur geringfügig anreichernd [8].



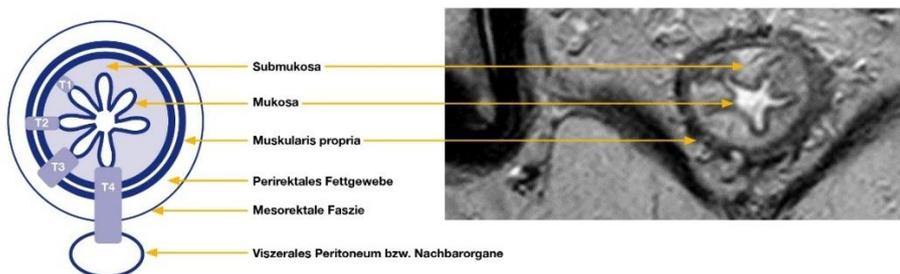
**Abb. 3:** Darstellung der Sphinkterebene in der Kernspintomographie. T1-Sequenz fettgesättigt nach Kontrastmittel koronar. Bildmaterial aus PACS (Picture Archiving and Communication System). Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

#### 2.4.4.2. Kriterien transmurales Tumorwachstum

Das transmurale Tumorwachstum beschreibt die radiale Tumorausdehnung entlang der Darmwand und wurde auf den Bildserien beurteilt, die senkrecht auf die Tumoralängsachse eingekippt waren. Hierfür waren die hochauflösenden T2-gewichteten Sequenzen relevant.

Die Darmwand ist von innen (Lumen) nach außen (muskuläre Darmwand) betrachtet untergliedert in die Mukosa, die Submukosa, sowie die Muskularis propria (bestehend aus der innen liegenden zirkulären und der äußeren Längsmuskelschicht), die jeweils in hochauflösenden T2-Sequenzen diskriminiert werden können.

Mit Ausnahme der muzinösen Tumoren, die auch signal angehoben oder inhomogen erscheinen können, weisen alle Rektumkarzinome eine gegenüber der Muskularis propria und Submukosa intermediäre Signalintensität auf, die Submukosa stellt sich den Tumorformationen gegenüber hyperintens dar, die Muskularis propria ist als hypointense Linie zwischen dem nach extramural angrenzenden hyperintens perirektalen Fettgewebe und der nach innen angrenzenden Submukosa erkennbar [5]. Die T-Stadien richteten sich entsprechend der TNM-Klassifikation aus.



**Abb. 4:** Schematische Darstellung der Tumor-Stadien nach der TNM-Klassifikation (TNM = Tumor/Lymphknoten/Metastasen). Links: Schema (Eigenerstellung). Rechts: Rektumwand in der Kernspintomographie. T2W-Sequenz axial. Bildmaterial aus PACS (Picture Archiving and Communication System). Patientenkollektiv primär operativ behandelter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

#### 2.4.4.3. Kriterien T-Stadien

##### 2.4.4.3.1. T1-Stadium

Ein T1-Stadium war zu vergeben, wenn in der Submukosa entweder ein zu ihr relativ gesehen niedrigeres Signal erkennbar oder die Submukosa vollständig durch abnormales Tumorsignal ersetzt war, das sich nicht weiter in die außen angrenzende Muskularis propria ausdehnte [11].

#### 2.4.4.3.2. T2-Stadium

Ein T2-Stadium war zu vergeben, wenn ein intermediäres Signal innerhalb der Muskularis propria erkennbar war, bzw. nach außen hin maximal der äußere Rand der Muskularis propria durch Tumorformationen intermediären Signals ersetzt war, sich jedoch kein Tumorsignal außerhalb der Muskelschicht im perirektalen Fettgewebe nachweisen ließ [11].

#### 2.4.4.3.3. T3-Stadium

Ein T3-Stadium war zu vergeben, wenn ein breitbasiges kontinuierliches Vorwölben oder einzelne noduläre Projektionen intermediären Tumorsignals außerhalb der Muskularis propria erkennbar waren (Tumordeposits, „Satellitenknoten“). *Nicht* jedoch rein spikuläre Formationen, die als desmoplastische Reaktion und nicht als Tumorsignal zu werten waren [11].

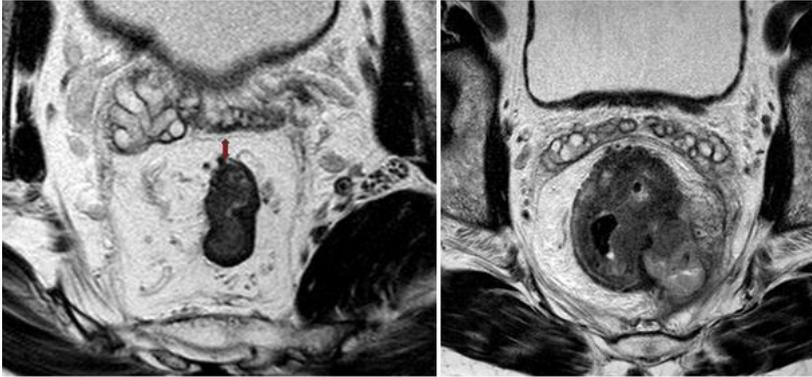
#### 2.4.4.3.4. T4-Stadium

Ein T4-Stadium war zu vergeben, wenn sich das Tumorsignal kontinuierlich in angrenzende Organe oder Strukturen ausdehnte (z.B. Beckenboden, Vagina, Prostata, Blase, Samenblase, Anm. d. Autors) oder das viszerale Peritoneum in kontinuierlicher Ausdehnung erreichte. Die Perforation der peritonealen Umschlagfalte war neben einer kontinuierlichen Infiltration zu werten bei nodulärer Ausbreitung von Tumorsignal über die im unbeteiligten Fall ebenmäßig konturierte Umschlagfalte hinaus [11].

#### 2.4.4.4. Kriterien CRM-Status

Die Bedrohung oder Infiltration der mesorektalen Faszie (CRM-Status „positiv“) war zu werten, wenn jegliche Tumorformation, also entweder ein Tumorsignal in kontinuierlicher Ausdehnung, in nodulärer Projektion (Tumordeposit, „Satellitenknoten“), oder in Form eines suspekten Lymphknotens (siehe 2.4.4.5.), direkt ohne messbaren Abstand an die mesorektale Faszie angrenzte, durch diese infiltrierend kontinuierlich hindurchzog oder an diese in Abstand von 1 mm oder weniger heranreichte, was mittels Linealfunktion im RIS in der kürzesten Distanz zu messen war [11].

In allen anderen Fällen, einer T1- oder T2-Läsion, war entsprechend die kürzeste Distanz zwischen der äußeren Rektumwandbegrenzung, also der äußeren Längsmuskelschicht der Muskularis propria, und der mesorektalen Faszie zu messen und analog zu werten.



**Abb. 5:** Beispiel: Bestimmung der Tumorausdehnung in Relation zur mesorektalen Faszie (CRM-Status. CRM = circumferential resection margin). T2W-Sequenz axial. Links: CRM-negativ, rechts: CRM-positiv (Tumorinfiltration der mesorektalen Faszie), Bildmaterial aus PACS (Picture Archiving and Communication System). Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

#### 2.4.4.5. Kriterien N-Status

Die Beurteilung der lokoregionären Lymphknoten erfolgte anhand von morphologischen Kriterien, nicht anhand von metrischen Kriterien (z.B. Kurzachsen-Durchmesser). Hierbei waren Lymphknoten mit irregulärer Außenkontur und/oder gemischtem, inhomogenem T2-Signal als Metastasen zu werten [12].



**Abb. 6:** Beispiel eines suspekten Lymphknotens in der Kernspintomographie. T2W-Sequenz koronar. Links oben: vergrößerter Bildausschnitt, Bildmaterial aus PACS (Picture Archiving and Communication System). Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

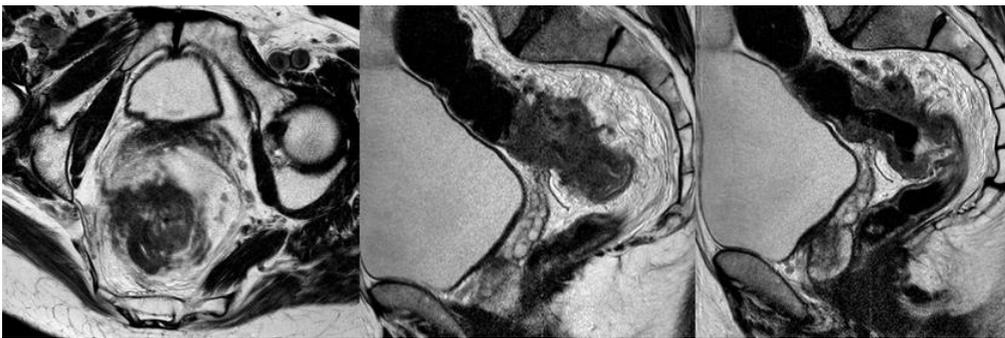
#### 2.4.4.6. Kriterien Tumorinvasion in das perirektale Fettgewebe

Wenn sich das Tumorsignal über die Muskularis propria hinaus ausdehnte, war mittels Linealfunktion im RIS seine maximale Ausdehnung in das angrenzende perirektale Fettgewebe, gemessen vom äußeren Rand der Längsmuskelschicht der Muskularis propria, zu bestimmen [11].

#### 2.4.4.7. Kriterien EMVI

Die EMVI beschreibt die Tumorinfiltration in außerhalb der Muskularis propria gelegene Venen im perirektalen Fettgewebe.

Die EMVI war definiert als geschlängelte Ausdehnung von Tumorsignal in einer vasculären Struktur (Gefäß). Ein Gefäß wurde definiert als in angrenzenden Schichten kontinuierlich verfolgbare tubuläre Struktur mit flow-void in den T2-gewichteten Sequenzen [11].



**Abb. 7:** Beispiel einer extramuralen venösen Tumorinvasion in der Kernspintomographie. Links: T2W-Sequenz axial. Mitte, rechts: T2W-Sequenz sagittal. Bildmaterial aus PACS (Picture Archiving and Communication System). Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

#### 2.4.4.8. Sonstige Daten der MRT-Befundung

Folgende allgemeine Daten wurden aus den Befundbeurteilungen sowie dem standardisierten Erhebungsbogen entnommen:

1. Freizeichnender Arzt (Befunder)
2. Bildqualität:

Nach subjektivem Empfinden des Befunders unterteilt in:  
Optimal / Suboptimal / Deutlich eingeschränkt

## 2.5. Statistische Auswertung

Der Vergleich der Ergebnisse der MRT und Pathologie wurde nach den folgenden statistischen Definitionen unter Modifikation für die gegebene Fragestellung bewertet:

**Sensitivität** (in Prozent):

$$\text{Sensitivität} = \text{Richtig-positive} / (\text{richtig-positive} + \text{falsch-negative})$$

**Spezifität** (in Prozent):

$$\text{Spezifität} = \text{Richtig-negative} / (\text{richtig-negative} + \text{falsch-positive})$$

**Positiv prädiktiver Wert** (in Prozent):

$$\text{PPW} = \text{Richtig-positive} / (\text{richtig-positive} + \text{falsch-positive})$$

**Negativ prädiktiver Wert** (in Prozent):

$$\text{NPW} = \text{Richtig-negative} / (\text{richtig-negative} + \text{falsch-negative})$$

Die Berechnung erfolgte jeweils anhand von Konvergenztabelle nach der gängigen Praxis medizinischer Studien [23].

### Reliabilität

Die Reliabilität oder Übereinstimmung der MRT-Ergebnisse im Vergleich mit der Pathologie im Bereich der T-Stadien sowie des CRM-Status wurde mittels Cohen's Kappa-Statistik evaluiert. Der Grund für die Auswahl dieses Verfahren war, dass alle Übereinstimmungen und Abweichungen aus Konvergenztabelle in einer Maßzahl zusammengefasst werden können (und nicht nur Teile hiervon, wie z.B. bei der Sensitivität). Hierbei wurden sowohl einzelne Kappa-Werte als auch ihre 95%-Konfidenzintervalle mit zugehörigen Z-Werten sowie der Standardabweichung berechnet [36].

## 2.5.1. Demographische Daten

### 2.5.1.1. Alter

Das Durchschnittsalter der Patienten zum Operationszeitpunkt wurde als arithmetischer Mittelwert berechnet.

### 2.5.1.2. Zeitraum zwischen MRT-Staging und Operation

Der Zeitraum zwischen MRT-Staging und Operation wurde mittels Interquartilsanalyse ausgewertet.

## 2.5.2. T-Stadien

Gesamtkollektiv:

Für das Gesamtkollektiv wurde jeweils die Sensitivität und Spezifität der Stadien T1-T4 berechnet.

Die Reliabilität (Übereinstimmung) der Ergebnisse des T-Stagings für das Gesamtkollektiv wurde sowohl stadienunabhängig (Overall) als auch jeweils für die Stadien T1-T4 mittels Cohen's Kappa-Statistik berechnet.

Teilkollektive der anatomischen Zonen A-C:

Für die Teilkollektive der anatomischen Zonen A-C wurde für die Stadien T1-T4 jeweils die Sensitivität sowie die Spezifität berechnet.

Die Reliabilität (Übereinstimmung) wurde sowohl stadienunabhängig (Overall) als auch für die Stadien T1-T4 jeweils für die anatomischen Zonen A-C bestimmt.

Einfluss der Tumorphöhe auf die Reliabilität:

Die Fragestellung, ob sich die anatomische Zone auf das Ergebnis der Reliabilität auswirkte, wurde mittels Konfidenzintervallanalyse der Kappa-Statistik evaluiert. Hierbei wurde das 95%-Konfidenzintervall der Overall-Reliabilität des Gesamtkollektivs den 95%-Konfidenzintervallen der Overall-Reliabilität der Teilkollektive der anatomischen Zonen A-C gegenübergestellt.

Es wurde kein Einfluss auf eine Metrik nach Tumorphöhe festgestellt, wenn sich das jeweilige 95%-Konfidenzintervall der zugehörigen Kappa-Werte überlappte.

Tumorinvasion perirektales Fettgewebe T3-Stadien:

Die Tumorinvasion in das perirektale Fettgewebe wurde für die T3-Stadien ausgewertet. Hierbei wurde die Invasionstiefe in mm aus den Datensätzen der Pathologie sowie den Befundberichten der MRT jeweils in zwei Klassen zusammengefasst: Invasionstiefe  $\leq 5$  mm sowie Invasionstiefe  $> 5$  mm. Anschließend wurden für die MRT die Sensitivität, die Spezifität, der PPW sowie der NPW für die Erkennung einer Invasionstiefe  $> 5$  mm der Pathologie bestimmt.

### **2.5.3. CRM-Status**

Gesamtkollektiv:

Für den Vergleich des CRM-Status der MRT mit der Pathologie wurde für das Gesamtkollektiv die Sensitivität, die Spezifität, der PPW sowie der NPW berechnet.

Die Reliabilität (Übereinstimmung) der Ergebnisse mit der Pathologie wurde mittels Cohen's Kappa-Statistik evaluiert.

Teilkollektive der anatomischen Zonen A-C:

Für die Teilkollektive der anatomischen Zonen A-C wurde jeweils die Sensitivität sowie die Spezifität berechnet.

Die Reliabilität (Übereinstimmung) wurde mittels Cohen's Kappa-Statistik evaluiert.

Einfluss der Tumorhöhe auf die Reliabilität:

Die Fragestellung, ob sich die anatomische Zone auf das Ergebnis der Reliabilität auswirkte, wurde mittels Konfidenzintervallanalyse der Kappa-Statistik evaluiert. Hierbei wurde das 95%-Konfidenzintervall des Kappa-Wertes des Gesamtkollektivs den 95%-Konfidenzintervallen der jeweiligen Kappa-Werte der anatomischen Zonen A-C gegenübergestellt.

Es wurde kein Einfluss auf eine Metrik nach Tumorhöhe festgestellt, wenn sich das jeweilige 95%-Konfidenzintervall überlappte.

#### **2.5.4. N-Status**

Für den Vergleich des N-Status wurde die Sensitivität, die Spezifität, der PPW sowie der NPW für das Gesamtkollektiv berechnet.

#### **2.5.5. EMVI**

Aufgrund der Datenlage konnte der EMVI-Status nicht statistisch ausgewertet werden.

Ethikkommission:

Die Erlaubnis der Auswertung der Daten wurde im Rahmen der OCUM-Studie eingeholt.

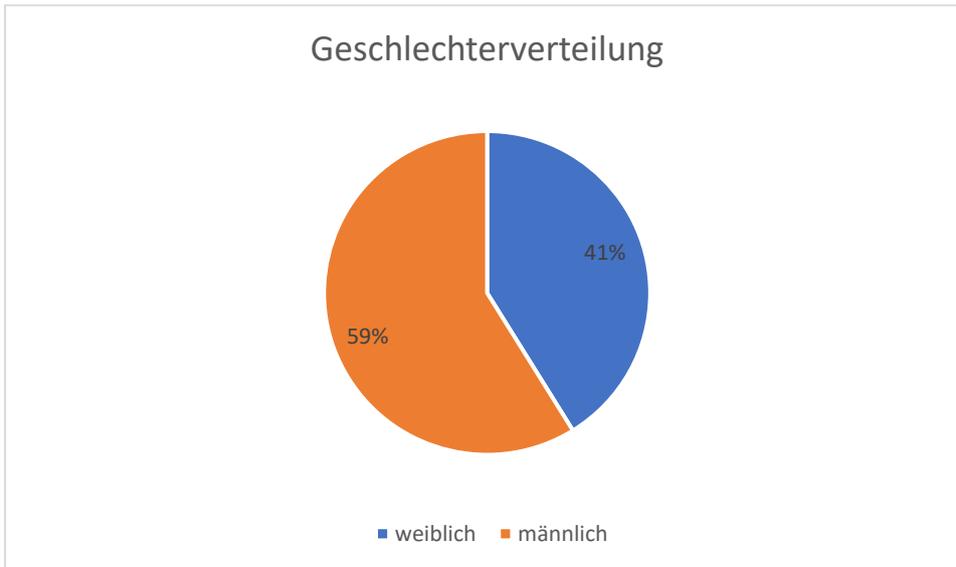
Aktenzeichen 837.203.07 (5738) Landesärztekammer Rheinland Pfalz.

### 3. Ergebnisse

#### 3.1. Demographische Daten

Das Patientendurchschnittsalter zum Operationszeitpunkt betrug 66 Jahre.

Der Anteil weiblicher und männlicher Patienten ist in folgender Abbildung ersichtlich:



**Abb. 8:** Geschlechterverteilung. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Bei den insgesamt 238 untersuchten Patienten bestand eine leichte Prädominanz des männlichen Geschlechts von 140 Patienten gegenüber 98 weiblichen Patienten.

Hinsichtlich des Zeitraums zwischen MRT-Staging und Operationsdatum lagen die mittleren 50% der Datensätze zwischen 7 (unteres Quartil) und 16 Tagen (oberes Quartil).

Daraus ergibt sich ein mittlerer Quartilsabstand von 9 Tagen.

Die Grenze für einen Ausreißer wurde mit dem 1,5-fachen des Interquartilsabstands oberhalb des oberen Quartils definiert.

Der Operationszeitraum von  $\geq 30$  Tagen galt somit als Ausreißer.

Hierauf entfielen 16 Datensätze.

## 3.2. T-Stadien

Der Vergleich der T-Stadien konnte für das Gesamtkollektiv (238 Patienten) evaluiert werden.

Die Ergebnisse der Stadien T1-T4 wurden sowohl für das Gesamtkollektiv als auch jeweils für die Teilkollektive der anatomische Zonen A-C berechnet und hierbei jeweils in einer Konvergenztabelle zusammengefasst, um ein etwaiges Over- bzw. Understaging sichtbar zu machen.

Die Stadien T1, T2 sowie T4 wurden hierbei ohne Berücksichtigung etwaiger Unterkategorien (wie z.B. Stadium T4a) ausgewertet. Für die T3-Stadien erfolgte eine separate die Einteilung in die Stadien T3a/b sowie T3c/d entsprechend der Infiltrationstiefe in das perirektale Fettgewebe (siehe 3.2.3.).

### 3.2.1. Gütekriterien des T-Stadienvergleichs

Die Verteilung der Stadien T1-T4 wurden für das Gesamtkollektiv als auch jeweils für die Teilkollektive der anatomische Zonen A-C in einer Konvergenztabelle zusammengefasst, um ein etwaiges Over- bzw. Understaging sichtbar zu machen. Die Gütekriterien errechneten sich jeweils aus der entsprechenden Konvergenztabelle.

#### 3.2.1.1. Ergebnisse des Gesamtkollektivs

##### 3.2.1.1.1. T-Stadienverteilung

Die grau hinterlegten Datenfelder entsprechen Übereinstimmungen mit den Ergebnissen der Pathologie. Die Datenfelder oberhalb des grau hinterlegten weisen jeweils ein Understaging, die Datenfelder unterhalb des grau hinterlegten weisen jeweils ein Overstaging der MRT im Vergleich mit der Pathologie aus.

So wurde z.B. in der T1-Kategorie in 14 von 24 Fällen (58%) der histopathologischen T1-Stadien in der MRT ein Stadium > T1 gesehen (Overstaging).

In der T2-Kategorie hingegen resultierte z.B. in 5 von 71 Fällen (7%) ein Understaging der MRT usw.

Für das Gesamtkollektiv erfolgte ein korrektes Staging in 63% (150/238 Fälle). Das Over- bzw. Understaging war im Gesamtkollektiv mit jeweils ca. 18% (jeweils 44/238 Fälle) exakt gleich.

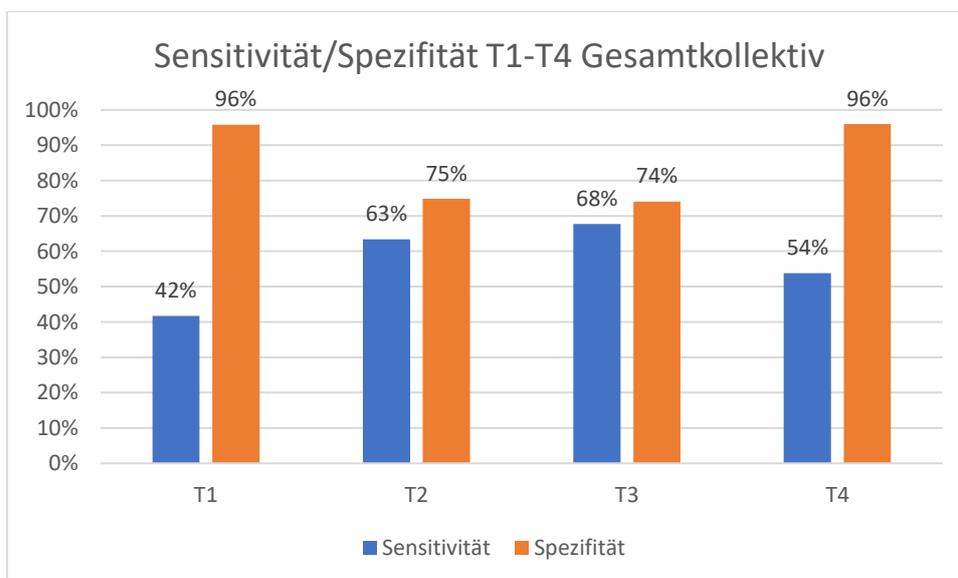
**Tab. 3:** Tumor-Stadienverteilung (T-Stadien) T1-T4 des Gesamtkollektivs. Gegenüberstellung der Ergebnisse der Kernspintomographie (MRT) und der Pathologie. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

MRT	Pathologie					n MRT
	T1	T2	T3	T4		
T1	10	5	4	0		19
T2	13	45	29	0		87
T3	1	21	88	6		116
T4	0	0	9	7		16
n Pathologie	24	71	130	13		238

In der T1-Kategorie dominierte das Overstaging zu T2. In der T4-Kategorie fielen die Übereinstimmungen nur geringgradig höher aus als das Understaging zu T3. In der T2-Kategorie war das Overstaging deutlich stärker ausgeprägt als das Understaging, in der T3-Kategorie hingegen trat das Understaging weitaus häufiger auf als das Overstaging.

#### 3.2.1.1.2. Sensitivität/Spezifität

Die errechneten Kennzahlen resultieren aus dem vollständigen Interstadienvergleich der Konvergenztabelle aus 3.2.1.1.1. (z.B. Sensitivität T1 = T1 vs. T2/T3/T4. Sensitivität T2 = T2 vs. T1/T3/T4 usw.). Die Ergebnisse sind in folgender Abbildung zusammengefasst:



**Abb. 9:** Sensitivität/Spezifität der Tumorstadien (T-Stadien) T1-T4 des Gesamtkollektivs. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Die Spezifität fiel in jeder Unterkategorie höher aus als die Sensitivität und war für T1- und T4-Stadien am höchsten.

Die Sensitivität war für T3-Stadien am höchsten sowie für T1-Stadien am geringsten.

### 3.2.1.2. Zonenabhängige Ergebnisse

#### 3.2.1.2.1. T-Stadienverteilung

Wie für das Gesamtkollektiv (siehe 3.2.1.1.1.) wurden auch die Ergebnisse der Teilkollektive der anatomischen Zone A-C in einer Konvergenztabelle zusammengefasst.

**Tab. 4:** Tumor-Stadienverteilung (T-Stadien) T1-T4 der anatomischen Zonen A-C. Gegenüberstellung der Ergebnisse der Kernspintomographie (MRT) und der Pathologie. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Zone A						
MRT	Pathologie					n MRT
	T1	T2	T3	T4	n Pathologie	
T1	7	2	2	0	11	
T2	6	19	16	0	41	
T3	1	10	47	3	61	
T4	0	0	5	4	9	
n Pathologie	14	31	70	7	122	

Zone B						
MRT	Pathologie					n MRT
	T1	T2	T3	T4	n Pathologie	
T1	3	3	2	0	8	
T2	5	21	10	0	36	
T3	0	10	37	1	48	
T4	0	0	2	2	4	
n Pathologie	8	34	51	3	96	

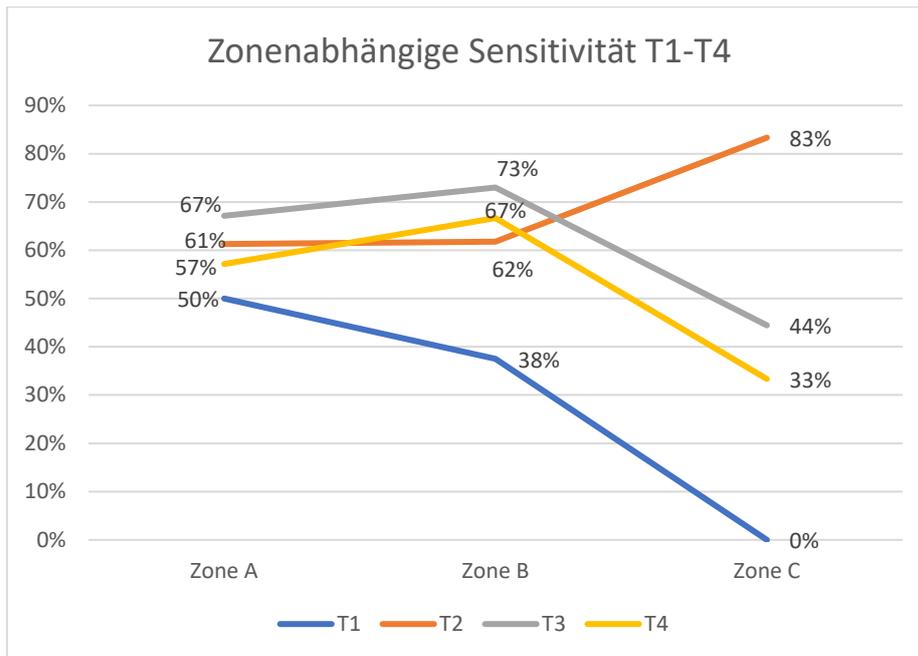
  

Zone C						
MRT	Pathologie					n MRT
	T1	T2	T3	T4	n Pathologie	
T1	0	0	0	0	0	
T2	2	5	3	0	10	
T3	0	1	4	2	7	
T4	0	0	2	1	3	
n Pathologie	2	6	9	3	20	

### 3.2.1.2.2. Sensitivität/Spezifität

#### 3.2.1.2.2.1. Zonenabhängige Sensitivität T1-T4

In folgendem Diagramm ist ersichtlich, wie die Sensitivität des jeweiligen T-Stadiums in den verschiedenen anatomischen Zonen ausfiel.



**Abb. 10:** Sensitivität der Tumorstadien (T-Stadien) T1-T4 der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Zusammenfassend war die geringe Sensitivität der T1-Stadien auffallend, insbesondere in Zone C mit 0%. In Zone C fanden sich zwei T1-Tumoren, die beide in der MRT nicht als T1 erkannt wurden. Das T1-Stadium wurde darüber hinaus in jeder Zone weniger sensitiv erkannt als alle übrigen T-Stadien.

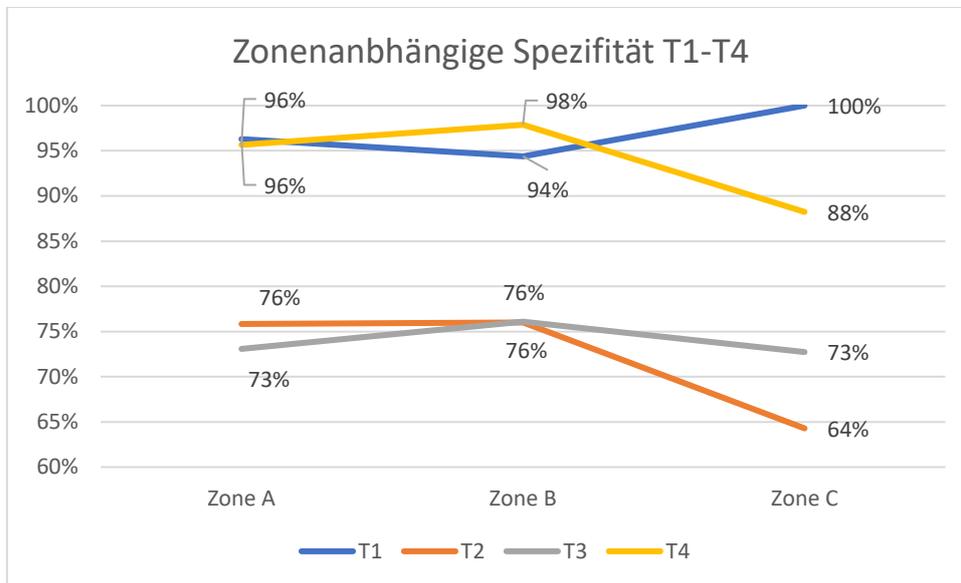
T3- und T4- Stadien wurden jeweils in Zone B am sensitivsten erkannt.

Im Gegensatz hierzu fiel die Sensitivität der T2-Stadien in Zone C am besten aus.

Mit Ausnahme des T2-Stadiums fiel die Sensitivität in Zone C am geringsten aus.

### 3.2.1.2.2.2. Zonenabhängige Spezifität T1-T4

In folgendem Diagramm ist ersichtlich, wie die Spezifität des jeweiligen T-Stadiums in den verschiedenen anatomischen Zonen ausfiel.



**Abb. 11:** Spezifität der Tumorstadien (T-Stadien) T1-T4 der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Die Schwankungen der Spezifität der einzelnen T-Stadien waren zwischen Zone A und B vergleichsweise gering.

Wie im Gesamtkollektiv fiel die Spezifität der T1- und T4-Stadien in jeder anatomischen Zone höher aus als die der übrigen T-Stadien.

Mit Ausnahme der T1-Stadien fiel die Spezifität in Zone C jeweils am geringsten aus.

### 3.2.2. Reliabilität der T-Stadien

In dieser Studie sollte untersucht werden, inwieweit die Ergebnisse des T-Stagings mit den Ergebnissen der Pathologie übereinstimmten und ob sich die Höhe des Tumorsitzes auf die Übereinstimmung mit der Pathologie auswirkte. Dieser Zusammenhang wurde mittels Cohen's Kappa-Statistik evaluiert. Der Kappa-Wert fasst alle Übereinstimmungen und Abweichungen in einer Maßzahl zusammen und beschreibt in dieser Studie die Reliabilität der MRT-Ergebnisse im Vergleich mit der Pathologie. Ein Kappa-Wert von 1 definiert die maximal erreichbare

Übereinstimmung, ein Kappa-Wert  $< 0$  weist auf eine Übereinstimmung hin, die kleiner ist als eine zufällige Übereinstimmung.

Landis und Koch schlagen zur Interpretation der Kappa-Werte vor [32]:

- $\kappa < 0$ : Schlechte Übereinstimmung („poor agreement“)
- $0 < \kappa < 0,20$ : Geringe („slight“) Übereinstimmung“
- $\kappa = 0,21-0,40$ : Ausreichende („fair“) Übereinstimmung“
- $\kappa = 0,41-0,60$ : Moderate („moderate“) Übereinstimmung“
- $\kappa = 0,61-0,80$ : Beachtliche („substantial“) Übereinstimmung“
- $\kappa = 0,81-1,00$ : (Fast) vollkommene („almost perfect“) Übereinstimmung“

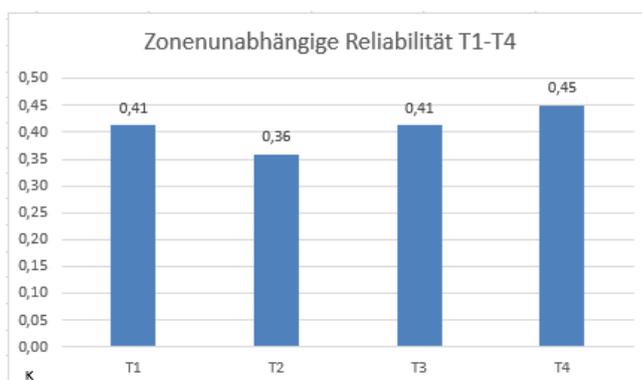
### 3.2.2.1. Ergebnisse des Gesamtkollektivs

#### 3.2.2.1.1. Overall-Reliabilität des Gesamtkollektivs

Für das gesamte T-Staging errechnete sich ein Overall-Kappa-Wert von 0,40, der die kumulativen Übereinstimmungen und Abweichungen sämtlicher Ergebnisse des T-Stagings für das Gesamtkollektiv zusammenfasst, also unabhängig von der Stadienverteilung *und* unabhängig von der anatomischen Zone.

#### 3.2.2.1.2. Reliabilität T1-T4 des Gesamtkollektivs

In folgender Feinauswertung ist die Übereinstimmung der einzelnen Stadien für das Gesamtkollektiv ersichtlich, unabhängig von der anatomischen Zone.



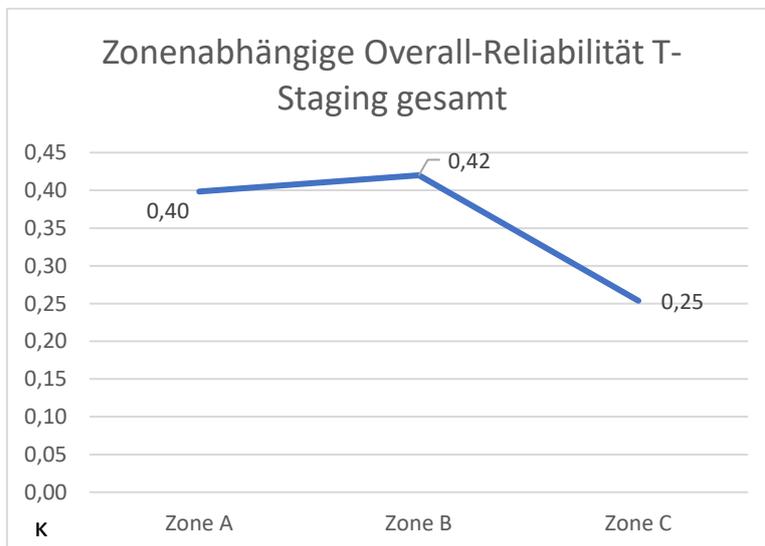
**Abb. 12:** Zonenunabhängige Reliabilität der Tumorstadien (T-Stadien) T1-T4 des Gesamtkollektivs. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Es zeigte sich eine ausreichende Übereinstimmung in der Bewertung der T2-Stadien. Die Übereinstimmung der Stadien T1, T3 und T4 lag jeweils auf einem höheren, einem moderaten Niveau.

### 3.2.2.2. Zonenabhängige Ergebnisse

#### 3.2.2.2.1. Zonenabhängige Overall-Reliabilität

Die Overall-Reliabilität des MRT T-Stagings, rein abhängig von der anatomischen Zone, unabhängig von der Stadienverteilung ist in folgender Abbildung zusammengefasst:



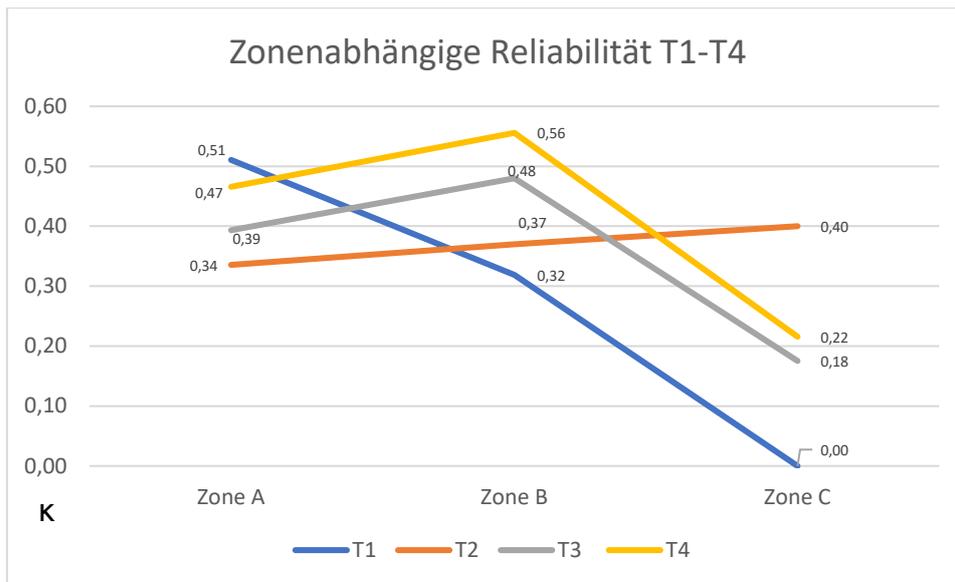
**Abb. 13:** Overall-Reliabilität des Tumor-Stagings (T-Staging) gesamt der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Die Gesamtübereinstimmung im T-Staging der MRT war in Zone B am besten und lag hier auf einem moderaten Niveau.

In Zone A und C lag die Übereinstimmung jeweils auf einem ausreichenden Niveau.

#### 3.2.2.2.2. Zonenabhängige Reliabilität T1-T4

In folgender Feinauswertung ist die Übereinstimmung der jeweiligen Unterkategorie des T-Stagings, abhängig von der anatomischen Zone A-C ersichtlich:



**Abb. 14:** Reliabilität der Tumorstadien (T-Stadien) T1-T4 der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Es zeigte sich kein einheitlicher Trend in den T-Unterkategorien:

Die Übereinstimmung der T2-Stadien lag in jeder Zone auf einem ausreichenden Niveau.

Die Übereinstimmung der T1-Stadien nahm kontinuierlich ab und fiel von Zone zu Zone um je ein Niveau von moderat über ausreichend auf gering ab, je weiter aboral der Tumorsitz lag.

Die Übereinstimmung der T3-Stadien war in Zone A ausreichend, in Zone B mit einem moderaten Niveau am besten, fiel jedoch in Zone C nurmehr gering aus.

Die Übereinstimmung der T4-Stadien war in Zone A und B moderat, in Zone C jedoch nurmehr ausreichend.

Zusammenfassend fiel mit Ausnahme der T2-Stadien die Übereinstimmung in Zone C am geringsten aus.

### 3.2.2.3. Einfluss der Tumorphöhe auf die Reliabilität der T-Stadien

Ob die Tumorphöhe Einfluss auf die Übereinstimmung der T-Stadien mit der Pathologie nahm, konnte in der vorliegenden Studie aufgrund der Fallzahlen sinnvoll nur für die Overall-Reliabilität des T-Stagings vorgenommen werden, also ohne Berücksichtigung der Unterkategorien T1-T4. Dies wurde mittels Analyse der Konfidenzintervalle der erreichten Kappa-Werte evaluiert. Ein einzelner Kappa-Wert einer Studie ist hierfür nicht geeignet. Die 95%-Konfidenzintervalle beschreiben in der

Kappa-Statistik die erwartete Streuung der Kappa-Werte, wenn die Studie mehrfach in gleicher Weise wiederholt werden würde.

Die Maßzahlen der zugrundeliegenden Kappa-Statistik können der folgenden Tabelle entnommen werden:

**Tab.5:** Maßzahlen der Kappa-Statistik der Tumor-Stadien (T-Stadien) Overall des Gesamtkollektivs sowie der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

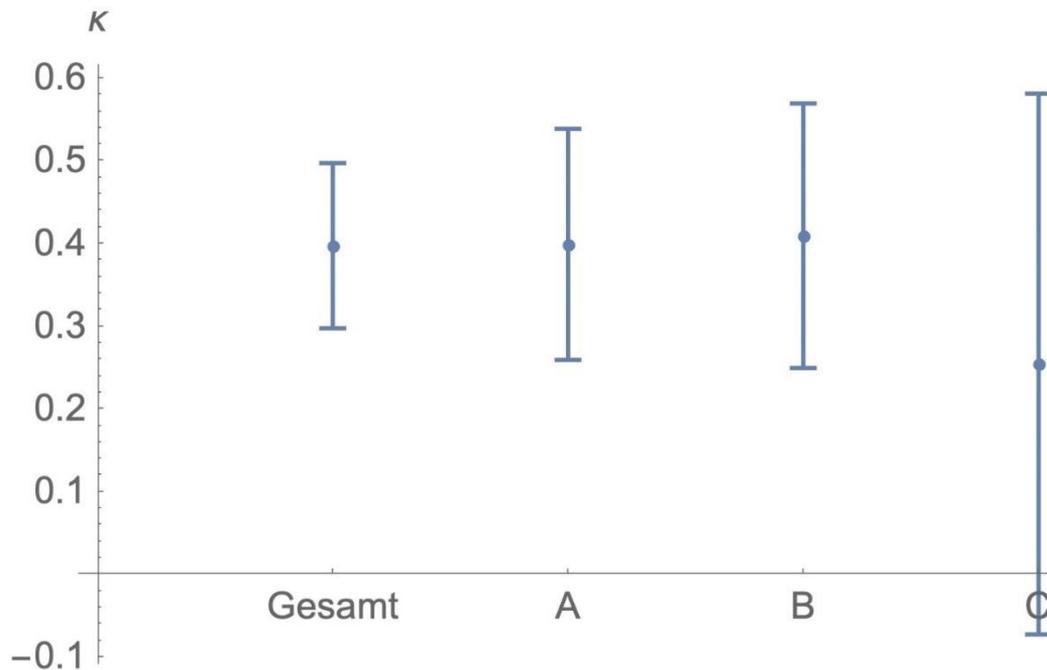
	Kappa	95%-Konfidenzintervall		Z-Wert	$\sigma$
Gesamt	0,40	0,30	0,50	7,77	0,05
Zone A	0,40	0,26	0,54	5,59	0,07
Zone B	0,42	0,26	0,58	5,15	0,08
Zone C	0,25	-0,07	0,58	1,52	0,17

Neben den jeweiligen Kappa-Werten für das Gesamtkollektiv sowie für die Teilkollektive der anatomischen Zonen A-C mit ihren zugehörigen 95%-Konfidenzintervallen (KI) ist auch der jeweilige Z-Wert angegeben. Der Z-Wert beschreibt, wie signifikant der jeweilige Kappa-Wert ist, mathematisch ausgedrückt, um wie viele Standardabweichungen ( $\sigma$ ) der gemessene Kappa-Wert dieser Studie bei mehrfacher Wiederholung der Studie von „ $\kappa = 0$ “ entfernt läge, wobei „ $\kappa = 0$ “ ausdrückt, dass die gemessenen Übereinstimmungen und Abweichungen einer zufälligen Übereinstimmung entsprechen können. Je höher also der Z-Wert, desto signifikanter ist der entsprechende Kappa-Wert zu beurteilen.

Für die vorgelegte Studie galt in der Interpretation der Konfidenzintervalle:

Es wurde kein Einfluss der anatomischen Zone auf die Übereinstimmung mit der Pathologie festgestellt, wenn sich das 95%-Konfidenzintervall zweier Kappa-Werte überlappte.

Die Ergebnisse der Studie sind in folgender Abbildung ersichtlich:



**Abb. 15:** Konfidenzintervallanalyse der Kappa-Statistik der Tumor-Stadien (T-Stadien) Overall des Gesamtkollektivs sowie der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Es zeigte sich eine starke Überlappung der 95%-Konfidenzintervalle für den jeweiligen Kappa-Wert (Punktmarkierung im Intervall) sowohl für das Gesamtkollektiv („Gesamt“) im Vergleich mit den anatomischen Zonen A-C, als auch für jedes Konfidenzintervall der anatomischen Zonen A-C untereinander.

Es ist auch erkennbar, dass im Konfidenzintervall der Zone C „ $\kappa = 0$ “ enthalten war.

### 3.2.3. Tumorinvasion perirektales Fettgewebe T3-Stadien

Anhand der Invasionstiefe in das perirektale Fettgewebe kann nach den ESMO-Leitlinien eine Subklassifikation in die Stadien T3 a-d vorgenommen werden, wobei einem Stadium T3a einer Tumorerfiltration  $< 1$  mm in das perirektale Fettgewebe entspricht, einem Stadium T3b die Infiltration von 1-5 mm, einem Stadium T3c die Infiltration von 5-15 mm sowie einem Stadium T3d die Infiltration  $> 15$  mm [20].

Diese Feinauswertung konnte aufgrund der geringen Fallzahlen nicht sinnvoll vorgenommen werden. Es waren insgesamt nur 68 Datensätze mit entsprechenden Daten auswertbar. Hiervon fanden sich lediglich vier Datensätze der Pathologie mit verwertbaren Angaben entsprechend eines Stadiums T3a bzw. lediglich zwei

Datensätze mit verwertbaren Angaben eines Stadiums T3d. Daher wurden die Daten zu zwei größeren Klassen zusammengefasst: Invasionstiefe  $\leq 5$  mm (41 Datensätze), was einem Stadium T3a/b entspricht, sowie Invasionstiefe  $> 5$  mm (27 Datensätze), was einem Stadium T3c/d entspricht. Die Ergebnisse wurden in einer 4-Felder-Tafel zusammengefasst. Die grau hinterlegten Felder entsprechen den Übereinstimmungen mit der Pathologie.

**Tab. 6:** Tumorinvasion in das perirektale Fettgewebe der T3-Tumor-Stadien (T-Stadien), gemessen in mm. Gegenüberstellung der Ergebnisse der Kernspintomographie (MRT) und der Pathologie. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Feinauswertung T3-Stadien				
MRT	Pathologie			n MRT
		$> 5\text{mm}$	$\leq 5\text{ mm}$	
$> 5\text{ mm}$		14	9	23
$\leq 5\text{ mm}$		13	32	45
n Pathologie		27	41	68

Aus dieser Darstellung lassen sich Kennzahlen in der Evaluation eines Stadiums T3c/d berechnen. Die Sensitivität betrug 52%, die Spezifität 78%, der PPW 61% sowie der NPW 71%.

### 3.3. CRM-Status

Von den insgesamt 238 in dieser Studie untersuchten Patienten konnte bei 228 Patienten ein vergleichbarer CRM-Status der Pathologie ermittelt werden. 10 Datensätze gingen nicht in die Bewertung ein. Hierbei handelte es sich z.B. um Patienten, die nicht im Rahmen einer TME, sondern z.B. mittels Lokalexzision therapiert wurden. Auch gab es Datensätze, in welchen aufgrund der Höhe des Tumorsitzes keine Relation des Tumorwachstums zur mesorektalen Faszie bestimmt werden konnte (Tumorsitz weit oberhalb der peritonealen Umschlagfalte bzw. Tumorsitz auf Höhe der Sphinkterebene) oder aus einem anderen Grund kein vergleichbarer CRM-Status ausgewertet werden konnte.

### 3.3.1. Gütekriterien des CRM-Status

#### 3.3.1.1. Ergebnisse des Gesamtkollektivs

Die Ergebnisse sind in der folgenden 4-Felder-Tafel zusammengefasst, aus der sich die jeweiligen Kennzahlen berechneten. Die grau hinterlegten Datenfelder weisen Übereinstimmungen mit der Pathologie aus.

**Tab.7:** Tumorausdehnung in Relation zur mesorektalen Faszie (CRM-Status. CRM = circumferential resection margin) des Gesamtkollektivs. Gegenüberstellung der Ergebnisse der Kernspintomographie (MRT) und der Pathologie. „Positiv“ beschreibt den Fall einer Bedrohung oder Infiltration der mesorektalen Faszie durch eine Tumorformation. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

CRM-Status gesamt				
MRT	Pathologie			n MRT
		positiv	negativ	
	positiv	7	35	42
	negativ	2	184	186
	n Pathologie	9	219	228

Im CRM-Status des Gesamtkollektivs konnte eine Sensitivität von 78%, eine Spezifität von 84%, ein PPW von 17% sowie ein NPW von 99% erreicht werden. Auffallend waren der hohe NPW sowie der niedrige PPW. Die Spezifität fiel höher aus als die Sensitivität.

#### 3.3.1.2. Zonenabhängige Ergebnisse

Wie für das Gesamtkollektiv wurden auch die Ergebnisse der Teilkollektive der anatomischen Zone A-C in einer Konvergenztabelle zusammengefasst.

**Tab. 8:** Tumorausdehnung in Relation zur mesorektalen Faszie (CRM-Status. CRM = circumferential resection margin) der anatomischen Zonen A-C. Gegenüberstellung der Ergebnisse der Kernspintomographie (MRT) und der Pathologie. „Positiv“ beschreibt den Fall einer Bedrohung oder Infiltration der mesorektalen Faszie durch eine Tumorformation. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Zone A				
MRT	Pathologie			
		positiv	negativ	n MRT
	positiv	1	21	22
	negativ	2	93	95
n Pathologie	3	114	117	

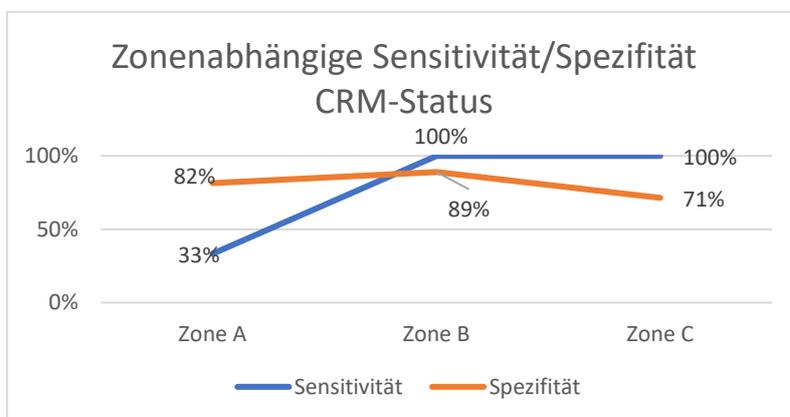
  

Zone B				
MRT	Pathologie			
		positiv	negativ	n MRT
	positiv	2	10	12
	negativ	0	81	81
n Pathologie	2	91	93	

Zone C				
MRT	Pathologie			
		positiv	negativ	n MRT
	positiv	4	4	8
	negativ	0	10	10
n Pathologie	4	14	18	

Sensitivität und Spezifität wurden abhängig von der anatomischen Zone A-C für den CRM-Status gesondert ausgewertet. Die Ergebnisse sind in folgender Abbildung zusammengefasst.



**Abb. 16:** Sensitivität/Spezifität der Tumorausdehnung in Relation zur mesorektalen Faszie (CRM-Status. CRM = circumferential resection margin) der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Auffallend ist die geringe Sensitivität in Zone A sowie die hohen Ergebnisse in Zone B und C. In Zone A waren sämtliche falsch-negative Fälle enthalten, also Patienten, die im MRT-Staging „CRM negativ“ bewertet wurden und in der pathologischen Aufarbeitung einen positiven CRM hatten (2 Patienten).

Die übrigen 7 CRM-positiven Fälle der Pathologie (Zone B und C) wurden im präoperativen MRT-Staging als richtig-positiv erkannt.

### 3.3.2. Reliabilität des CRM-Status

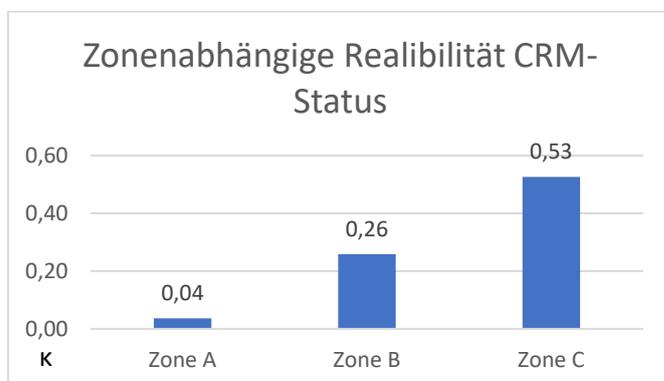
In dieser Studie sollte untersucht werden, inwieweit die Ergebnisse des CRM-Status mit den Ergebnissen der Pathologie übereinstimmen und ob sich die Höhe des Tumorsitzes auf die Übereinstimmung im T-Stadienvergleich mit der Pathologie auswirkte. Dieser Zusammenhang wurde mittels Cohen's Kappa-Statistik evaluiert. Zur Interpretation der Kappa-Werte siehe 3.2.2.

#### 3.3.2.1. Ergebnisse des Gesamtkollektivs

Für das Gesamtkollektiv errechnete sich ein Kappa-Gesamt von 0,22 der MRT im Vergleich mit der Pathologie.

#### 3.3.2.2. Zonenabhängige Ergebnisse

In folgender Feinauswertung ist die Übereinstimmung der Ergebnisse im CRM-Status mit der Pathologie in der jeweiligen anatomischen Zone A-C ersichtlich:



**Abb. 17:** Reliabilität der Tumorausdehnung in Relation zur mesorektalen Faszie (CRM-Status. CRM = circumferential resection margin) der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Auffallend ist die geringe Übereinstimmung in Zone A, die sehr weit unter der ausreichenden Übereinstimmung der Zone B lag. Die Übereinstimmung war in Zone C mit einem moderaten Niveau am besten.

Zusammenfassend lag die Übereinstimmung in jeder anatomischen Zone auf einem unterschiedlichen Niveau und verbesserte sich mit sinkender Tumorphöhe.

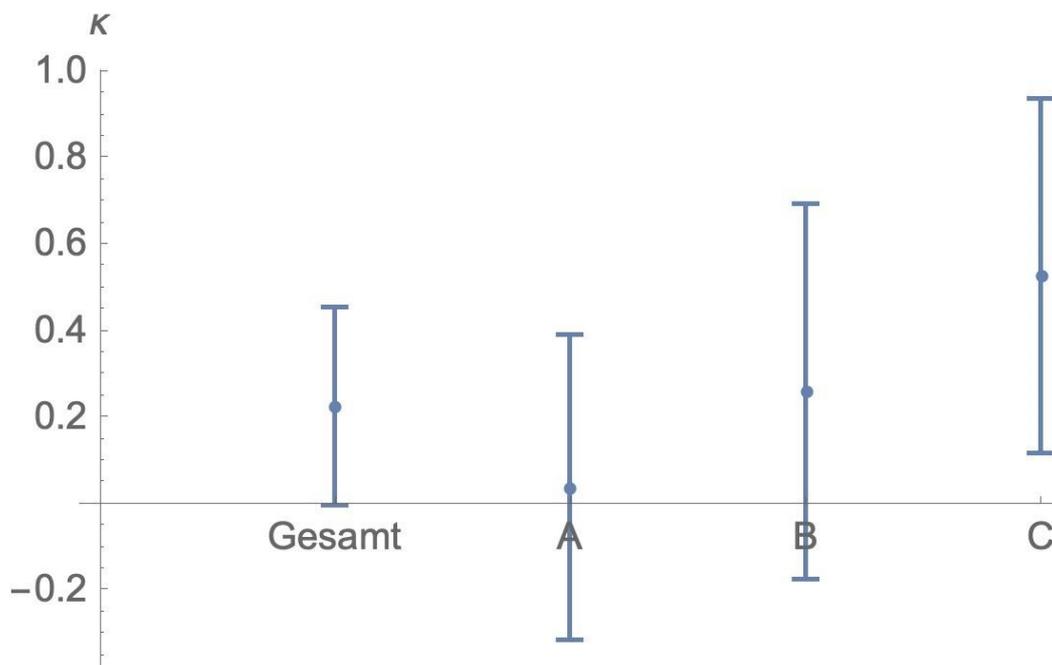
### 3.3.2.3. Einfluss der Tumorphöhe auf die Reliabilität des CRM-Status

Ob die Tumorphöhe Einfluss auf die Übereinstimmung mit der Pathologie nahm, wurde mittels Analyse von 95%-Konfidenzintervallen der jeweiligen Kappa-Werte überprüft. Zur Interpretation und Erläuterung der Maßzahlen siehe 3.2.2.3. Die Maßzahlen der Kappa-Statistik können der folgenden Tabelle entnommen werden:

**Tab. 9:** Maßzahlen der Kappa-Statistik der Tumorausdehnung in Relation zur mesorektalen Faszie (CRM-Status. CRM = circumferential resection margin) des Gesamtkollektivs sowie der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

	Kappa	95%-Konfidenzintervall		Z-Wert	$\sigma$
Gesamt	0,22	-4,78 e <sup>-3</sup>	0,45	1,92	0,12
Zone A	0,04	-0,32	0,39	0,20	0,18
Zone B	0,26	-0,18	0,69	1,17	0,22
Zone C	0,53	0,12	0,94	2,52	0,21

Die Ergebnisse der Konfidenzintervallanalyse sind in folgender Abbildung ersichtlich:



**Abb. 18:** 95%-Konfidenzintervallanalyse der Kappa-Auswertung zur Tumorausdehnung in Relation zur mesorektalen Faszie (CRM-Status. CRM = circumferential resection margin) des Gesamtkollektivs sowie der anatomischen Zonen A-C. Patientenkollektiv primär operativ therapierter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

Zusammenfassend zeigte sich eine starke Überlappung der Konfidenzintervalle für den jeweiligen Kappa-Wert (Punktmarkierung im Intervall) sowohl für das Gesamtkollektiv („Gesamt“) im Vergleich mit den anatomischen Zonen A-C, als auch für jedes Konfidenzintervall der anatomischen Zonen A-C untereinander. Auch war in den Konfidenzintervallen mit Ausnahme der Zone C jeweils „ $\kappa = 0$ “ enthalten.

### 3.4. N-Status

Insgesamt 6 Patientendatensätze gingen nicht in die Auswertung ein. Hierbei handelte es sich um Patienten mit dem Nodalstatus „x“ der Pathologie bzw. um Patienten, die mittels Lokalexzision therapiert worden, für die sich kein entsprechender N-Status der MRT vergleichen ließ.

Die Ergebnisse sind in der folgenden 4-Felder-Tafel zusammengefasst, aus der sich die jeweiligen Kennzahlen berechneten. Die grau hinterlegten Datenfelder weisen Übereinstimmungen mit der Pathologie aus.

**Tab. 10:** Verteilung des Lymphknotenstatus (N-Status) des Gesamtkollektivs. Gegenüberstellung der Ergebnisse der Kernspintomographie (MRT) und der Pathologie. „Positiv“ beschreibt den Fall eines tumorsuspekten Lymphknotens. Patientenkollektiv primär operativ behandelter Patienten mit Rektumkarzinom im Klinikum Neuperlach 2008-2015.

N-Status				
MRT	Pathologie			n MRT
		N-positiv	N-negativ	
	N-positiv	72	81	153
	N-negativ	13	66	79
	n Pathologie	85	147	232

Es konnte eine Sensitivität von 85%, eine Spezifität von 45%, ein PPW von 47%, sowie ein NPW von 84% erreicht werden. Auffallend waren die geringe Spezifität sowie der niedrige PPW. Die Sensitivität fiel deutlich höher aus als die Spezifität.

### 3.5. EMVI

In 28 Datensätzen der MRT war in den Befundbeurteilungen der Hinweis auf eine EMVI dokumentiert. In 18 Fällen hiervon wurde sich definitiv auf das Vorliegen einer EMVI festgelegt (64%), in 10 Fällen (36%) war die Diagnose als nicht sicher dokumentiert (in 7 Fällen fand sich der Hinweis „fragliche EMVI“, in 3 Fällen fand sich der Hinweis „V.a. EMVI“).

In 16 Datensätzen der Pathologie fanden sich in den Befundbeschreibungen „venöse Tumorinfiltrate“ oder „venöse Gefäßeinbrüche“. Nur in 3 Datensätzen fand sich übereinstimmend in den Befundbeurteilungen der MRT der Hinweis auf das Vorliegen einer EMVI und gleichzeitig der Hinweis auf eine venöse Tumorinfiltration in den Befundbeschreibungen der Pathologie. Es ist festzustellen, dass die Diskrepanz der Befunde hoch war. Aufgrund der Datenlage konnte der Status „EMVI-positiv“ aus den pathologischen Befundberichten retrospektiv nicht sicher gewertet und daher nicht statistisch ausgewertet werden.

### 3.6. Zusammenfassung der wichtigsten Ergebnisse

- Im T-Stadienvergleich fiel die Spezifität für jedes Stadium T1-T4 (96%, 75%, 74% und 96%) jeweils höher aus als die Sensitivität (42%, 63%, 68% und 54%), wobei die T3-Stadien am sensitivsten erkannt wurden. Die Spezifität hingegen war für T1- und T4-Stadien am höchsten. In der Feinauswertung der anatomischen Zonen fand sich in Zone C (mit Ausnahme der Sensitivität der T2-Stadien sowie der Spezifität der T1-Stadien) tendenziell die niedrigste Güte in der Evaluation der T-Stadien.
- In der Reliabilität lagen die einzelnen Kappa-Werte des T-Staging gesamt (Overall) sowie in der zonenunabhängigen Feinauswertung der Stadien T1-T4 teilweise auf unterschiedlichen Niveaus, jedoch ohne größere Heterogenität im Sinne auffälliger Ausreißer.
- In der Feinauswertung der anatomischen Zonen A-C war die Übereinstimmung mit der Pathologie in Betrachtung der einzelnen Kappa-Werte sowohl für das T-Staging gesamt (Overall) als auch für die Auswertung der einzelnen T-Stadien T1-T4 (mit Ausnahme der T2-Stadien) in Zone C tendenziell am geringsten.
- In der Konfidenzintervallanalyse der T-Stadien konnte kein Einfluss auf eine Metrik nach Tumorhöhe festgestellt werden, wobei die Ergebnisse in Zone C statistisch nicht signifikant waren.
- In der Feinauswertung der T3-Stadien hinsichtlich einer höhergradigen tumorösen Invasionstiefe in das perirektale Fettgewebe  $> 5$  mm konnte ein moderater negativer Vorhersagewert von 71% sowie eine etwas höhere Spezifität von 78% erreicht werden. Die Sensitivität von 52% sowie der positive Vorhersagewert von 61% hingegen fielen vergleichsweise gering aus.
- Im Vergleich des CRM-Status war die Evaluation eines negativen CRM von wesentlich höherer Güte, als die eines positiven CRM: Die Spezifität von 84% und insbesondere der negative Vorhersagewert von 99% waren auf einem hohen, bzw. einem sehr hohen Niveau. Die Sensitivität fiel mit 78% etwas geringer aus als die Spezifität, der positive Vorhersagewert von 17% fiel im Vergleich zum negativen Vorhersagewert jedoch sehr gering aus.  
Mit der Definition eines positiven CRM-Status der MRT bei Abstand einer Tumorformation  $\leq 1$  mm zur mesorektalen Faszie konnten zusammenfassend in der Prädiktion eines negativen CRM gute Ergebnisse erzielt werden.

- In der Feinauswertung der anatomischen Zonen A-C war in Zone A die Sensitivität mit 33% am geringsten ausgeprägt, in Zone B und C war die Sensitivität mit jeweils 100% maximal.
- In der Übereinstimmung des CRM-Status mit der Pathologie war der Kappa-Wert von 0,04 in Zone A sehr gering und ein auffälliger Ausreißer der Daten in der Einzelwertbetrachtung.
- In der Konfidenzintervallanalyse des CRM-Status der Reliabilität konnte kein Einfluss auf eine Metrik nach Tumorchöhe festgestellt werden, da sich die Konfidenzintervalle in der Feinauswertung der anatomischen Zonen A-C sowie dem Gesamtkollektiv stark überschneiden. Hierbei waren jedoch lediglich die Ergebnisse der anatomischen Zone C statistisch signifikant.
- Im Vergleich des N-Status der MRT mit der Pathologie konnte eine hohe Sensitivität von 85% erreicht werden. Der negative Vorhersagewert fiel mit 84% ebenfalls hoch aus. Die Spezifität von 45% sowie der positive Vorhersagewert von 47% fielen jedoch gering aus. Die Sensitivität fiel deutlich höher aus als die Spezifität.

Mit den angewandten Malignitätskriterien für suspekten Lymphknoten, nämlich dem inhomogenen T2-Signal und/oder einer irregulären Außenkontur konnten zusammenfassend gute Ergebnisse erzielt werden.

- Die Evaluation einer EMVI im Vergleich mit der Pathologie konnte aufgrund der Datenlage des retrospektiven Studiendesigns nicht sinnvoll statistisch ausgewertet werden.

## **4. Diskussion**

### **4.1. Demographische Daten**

Das Durchschnittsalter der Patienten zum Operationszeitpunkt betrug 66 Jahre. Es zeigte sich eine leichte Prädominanz männlicher Patienten ggü. weiblichen Patienten mit einem Anteil von 59% bzw. 41% im Kollektiv.

Im Jahr 2016 war die Rate an Neuerkrankungen kolorektaler Karzinome in Deutschland mit 55% bzw. 41% ähnlich verteilt, ebenfalls mit einer leichten Prädominanz des männlichen Geschlechts. Das mittlere Erkrankungsalter lag in Deutschland im Jahr 2016 geschlechterübergreifend bei >70 Jahren und war damit ähnlich, wenn auch etwas höher als im Patientenkollektiv der vorgelegten Studie [24].

In den mittleren 50% der Datensätzen betrug der Zeitraum zwischen MRT-Staging und Operation 7 bis 16 Tage. Als Ausreißer von diesem Zeitraum war ein Zeitraum > 30 Tage zu sehen. Hierauf entfielen ca. 7% der Patientendatensätze. Gründe hierfür waren z.B. eine individuelle Therapieentscheidung der jeweiligen Patienten gewesen oder das Vorliegen einer anderen Erkrankung, die therapeutisch zunächst im Vordergrund stand.

### **4.2. T-Stadien**

#### **4.2.1. Gütekriterien des T-Stadienvergleichs**

##### **4.2.1.1. Vergleich der Ergebnisse mit der Literatur**

Eine Metaanalyse von Zhang et al. (2016) fasste die Ergebnisse von 35 Studien bzgl. des präoperativen T-Stagings der MRT nach dem TNM-System der Jahre 2000-2016 zusammen.

Es errechnete sich eine gepoolte Sensitivität der Stadien T1-T4 von:

58%, 62%, 80% sowie 71% [56].

Für das Patientenkollektiv aus Neuperlach errechnete sich analog eine Sensitivität von: 42%, 63%, 68% sowie 54%.

Übereinstimmend mit dieser Metaanalyse war im Patientenkollektiv Neuperlach die Sensitivität der T1-Stadien am geringsten und die der T3-Stadien am höchsten, auch fiel die Spezifität wie in der Metaanalyse jeweils höher aus als die Sensitivität. Mit Ausnahme der T2-Stadien, die nahezu deckungsgleiche Ergebnisse in der Sensitivität zeigten, lagen die Ergebnisse im Studienkollektiv aus Neuperlach deutlich unter den Ergebnissen der Metaanalyse:

Die Sensitivität der T1-, T3- sowie T4-Stadien lag außerhalb der jeweiligen 95% Konfidenzintervalle der Metaanalyse (Range T1: 49%-66%, Range T3: 78%-82%, Range T4: 63%-79%) [56].

Die errechnete Spezifität der Metaanalyse betrug jeweils: 97%, 81%, 74% sowie 97% [56]. Die Spezifität des Patientenkollektivs aus Neuperlach betrug jeweils 96%, 75%, 74% sowie 96%.

Übereinstimmend mit dieser Metaanalyse fand sich die höchste Spezifität bei den T1- und T4-Stadien.

Die Spezifität der Stadien T1, T3 und T4 lag mit einem Unterschied von max. 1% auf dem gleichen Niveau (die Spezifität der T3-Stadien war sogar exakt gleich). Die Spezifität der T2-Stadien jedoch war mit 75% deutlich geringer ausgeprägt und lag außerhalb des 95%-Konfidenzintervalls der Metaanalyse von 79%-84% [56].

#### 4.2.1.2. **Bewertung der Ergebnisse**

Zusammenfassend ist festzustellen, dass neben tendenziellen Gemeinsamkeiten in der diagnostischen Performance vieler Studien, nämlich der gleichlautend höchsten Sensitivität für T3-Stadien und der höchsten Spezifität für T1- und T4-Stadien, die Sensitivität in der Feinauswertung mit Ausnahme der T2-Stadien vergleichsweise gering ausfiel. Die Spezifität lag – mit Ausnahme der T2-Stadien – auf einem vergleichbaren Niveau mit anderen Studien.

Somit sind die Abweichungen der Sensitivität im T-Staging näher zu untersuchen. Es ist bekannt, dass die Stadien T1 und T2 sowie frühe T3-Stadien in der MRT nicht mit hoher Genauigkeit differenziert werden können, wobei die Unterscheidung in der MRT auch nicht hochrelevant ist [5,48,15]. Die Erkennung eines T1-Stadiums kann ggf. für die Entscheidung einer Therapie mittels Lokalexzision relevant sein, hier kommt jedoch dem Verfahren der Endosonographie eine bedeutendere Rolle

zu als dem der MRT [28]. Daher ist eine Abweichungsanalyse für diese beiden Kategorien eher unbedeutend.

Die vergleichsweise geringe Sensitivität der klinisch bedeutsameren T3- und T4-Stadien dieser Studie im Vergleich mit der Literatur ist im Kontext des jeweils erfolgten Over- bzw. Understaging zu sehen, das Over- bzw. Understaging der MRT histologisch gesicherter T3-Stadien betrug kumuliert 32% (42/130 Fälle), das Understaging der T4-Stadien betrug 46% (6/13 Fälle).

Innerhalb der T3-Kategorie dominierte das Understaging zu T2 und erklärt 78% der Abweichungen in der T3-Kategorie (33/42 Fälle inkorrekt Staging). In 26 dieser Fälle war in den Datensätzen der Pathologie eine Invasionstiefe des perirektalen Fettgewebes angegeben worden. In 16 Fällen hiervon betrug die Invasionstiefe  $\leq$  2 mm. Mehrheitlich handelte es sich also bzgl. des Understaging der MRT um Fälle, in welchen ein T2-Stadium nicht sicher von einem frühen T3-Stadium unterschieden werden konnte, was prognostisch nicht sehr relevant ist.

In der Betrachtung der Teilkollektive der anatomischen Zone A-C fiel die Sensitivität der T3-Stadien in Zone C mit 44% am geringsten aus und war auffallend geringer als die Sensitivität des Gesamtkollektivs, jedoch fand sich hier ein Over- bzw. Understaging in lediglich 5 Fällen, was nur 11% des gesamten Over- bzw. Understaging der T3-Kategorie erklärt und somit die relativ niedrige Sensitivität der anatomischen Zone C für das Gesamtergebnis keine große Rolle gespielt haben dürfte. Eine bekannte diagnostische Schwierigkeit ist weiterhin die Unterscheidung eines realen Tumorsignals von einer desmoplastischen Reaktion [4]. Diese Schwierigkeit ist in der klinischen Standarddiagnostik nach wie vor nicht sicher zu lösen [47]. Eine Fehlinterpretation eines realen Tumorsignals hin zu einer desmoplastischen Reaktion könnte jedoch einen gewissen Anteil des Understaging erklären, auch wenn dieser retrospektiv nicht zu ermitteln ist.

In der vorgelegten Studie wurden die Kriterien zur Unterscheidung eines Tumorsignal von einer desmoplastischen Reaktion nach Brown et al. (2003) angewandt [11]. In dieser Studie konnte in der Evaluation eines T3 Stadiums rechnerisch zwar einer Sensitivität von 86% ermittelt werden. Allerdings handelte es sich hier um eine prospektive Studie, so dass die Ergebnisse nicht direkt vergleichbar sind. Allerdings dominierte auch hier das Understaging zu T2 und war – ähnlich wie im Patientenkollektiv aus Neuperlach – für 75% der Fälle eines inkorrekt Staging in der T3-

Kategorie verantwortlich [11]. Der Einfluss einer desmoplastischen Reaktion als Grund für ein mögliches Understaging wurde hier zwar nicht diskutiert, könnte aber auch hier eine Rolle gespielt haben.

Bzgl. der allgemeinen Problematik der MRT in der Unterscheidung von T1- und T2-Stadien sowie T2-Stadien von frühen T3-Stadien und der Evaluation einer Fibrose oder desmoplastischen Reaktion wurden in einer jüngeren Studie aus 2019 vielversprechende Ergebnisse aus der DTI-Bildgebung der MRT veröffentlicht, in der die Erkennung der Muskularis propria deutlich besser gelang als in den herkömmlichen T2-Sequenzen. Auch die Unterscheidung von Tumorsignal von einer Fibrose oder desmoplastischen Reaktion gelang deutlich besser als in den diagnostisch üblichen T2-Sequenzen [40]. Allerdings handelte es sich hier um eine experimentelle Studie mit einem 11,2-Tesla-Scanner, diese Feldstärken sind in modernen Tomographien für klinische Diagnostik in jüngerer Zukunft noch nicht zu erwarten.

Für die vergleichsweise geringe Sensitivität der Erkennung eines T4-Stadiums im Patientenkollektiv aus Neuperlach im Vergleich mit der Literatur hat wohl die Fehlinterpretation der mesorektalen Faszie mit der peritonealen Umschlagfalte die entscheidende Rolle gespielt, wie sie in der Abweichungsanalyse der Reliabilität bzgl. des CRM-Status in Zone A gefunden wurde (siehe 4.3.2.2.). Diese Fehlinterpretation erklärt das Understaging zu T3 in 3 der 6 Fälle und damit 50% der gesamten Abweichungen in der Evaluation eines T4-Stadiums.

#### 4.2.1.3. **Bewertung der zonenabhängigen Ergebnisse des T-Stagings**

In den zonenabhängigen Ergebnissen der Sensitivität und Spezifität fand sich kein einheitliches Bild, es gab wenige zonenabhängige Auffälligkeiten:

Zusammenfassend war in der Zone C (mit Ausnahme der T2-Stadien) die Sensitivität tendenziell am geringsten.

Mit Ausnahme der Spezifität der T1-Stadien, die in Zone C um 6% moderat anstieg, war die Spezifität der übrigen T-Stadien in Zone C eher geringer als in Zone A und B.

Somit ist festzustellen, dass sich die Zone C tendenziell negativ sowohl auf die Sensitivität als auch auf die Spezifität im T-Staging auswirkte, die Zonen A und B zeigten tendenziell bessere Ergebnisse.

Es gilt zu bedenken, dass in Zone C die Fallzahl der Patienten mit 20 Datensätzen jedoch sehr klein war und daher die Ergebnisse hier größeren Schwankungen unterliegen können als in Zone A und B mit 122 bzw. 96 Datensätzen.

#### **4.2.2. Reliabilität des T-Stadienvergleichs**

In dieser Studie sollte die Übereinstimmung des T-Stadienvergleichs mit der Pathologie untersucht werden, dieser Zusammenhang wurde mittels Kappa-Statistik evaluiert.

##### **4.2.2.1. Reliabilität des Gesamtkollektivs**

Für das gesamte, stadienunabhängige T-Staging lag die Übereinstimmung mit der Pathologie auf einem ausreichenden Niveau.

Es fand sich ein ebenfalls ausreichendes Niveau der T2-Stadien.

In den Kategorien T1, T3 sowie T4 fand sich einheitlich ein moderates Niveau der Übereinstimmung, das sowohl höher war als das der T2-Stadien als auch des stadienunabhängigen gesamten T-Stagings.

Die Unterschiede der einzelnen Kappa-Werte waren nicht ausgeprägt, es fanden sich zumindest keine stark auffälligen Abweichungen der Übereinstimmung im Stadienvergleich T1-T4 des Gesamtkollektivs, die einer näheren Analyse bedurft hätten.

##### **4.2.2.2. Zonenabhängige Reliabilität**

Die stadienunabhängige Übereinstimmung für das gesamte T-Stagings (Overall) war in Zone B am besten und lag hier auf einem moderaten Niveau.

In Zone A und C lag die Übereinstimmung jeweils auf einem ausreichenden Niveau, die Übereinstimmung in der Zone C war hierbei insgesamt tendenziell am geringsten ausgeprägt. Es fanden sich jedoch lediglich 20 Datensätze in der Zone C. Somit können die Ergebnisse hier großen Schwankungen unterliegen.

Es zeigte sich kein einheitlicher Trend in den T-Unterkategorien:

Zusammenfassend fanden sich mit Ausnahme der T2-Stadien, in welchen die Übereinstimmung je auf einem einheitlichen Niveau lag, in den übrigen Stadien

mindestens zwei, maximal sogar drei unterschiedliche Niveaus abhängig von der anatomischen Zone, wenn auch mit unterschiedlichen Trends. Mit Ausnahme der T2-Stadien war die Übereinstimmung in Zone C jeweils am geringsten. Auffällig war lediglich die geringe Übereinstimmung der T1-Stadien in Zone C mit einem Kappa-Wert von 0, der ausdrückt, dass sich die Übereinstimmung nicht wesentlich von einer zufälligen Übereinstimmung unterscheidet. Auch dies bedarf jedoch keiner näheren Analyse, die Fallzahlen der T1-Stadien in Zone C (2 Datensätze), waren faktisch zu gering, als dass die Kappa-Statistik hier aufschlussreich gewesen wäre.

Insgesamt lieferte die Feinauswertung der T-Unterkategorien abhängig von der anatomischen Zone keine Ergebnisse, die einer näheren Analyse bedurft hätten, die Fallzahlen in der feinen Aufgliederung waren hier überwiegend zu gering. Die Analyse wurde dennoch durchgeführt um ggf. Hinweise zu finden, ob es starke Auffälligkeiten im Kollektiv gab, in welchen die Dichte der Fallzahlen relativ gesehen hoch war (also z.B. T2- und T3-Stadien in Zone A und B). Hier fanden sich jedoch ebenfalls keine drastischen Abweichungen der Kappa-Werte untereinander, oder Kappa-Werte nahe 0, die einer näheren Analyse bedurft hätten.

#### **4.2.2.3. Einfluss der Tumorphöhe auf die Reliabilität des T-Stadienvergleichs**

Die Beurteilung, ob die Tumorphöhe entsprechend der anatomischen Zonen A-C Einfluss auf die Übereinstimmung im T-Staging mit der Pathologie genommen hat, wurde mittels Konfidenzintervallanalyse der Kappa-Statistik evaluiert. Aufgrund der starken Überlappung der Konfidenzintervalle muss gefolgert werden, dass die Tumorphöhe keinen Einfluss auf die Reliabilität im T-Stadienvergleich mit der Pathologie genommen hat.

Hierbei waren die Ergebnisse in Zone C nicht signifikant, wohl aufgrund der geringen Fallzahlen.

### 4.2.3. Tumorinvasion perirektales Fettgewebe T3-Stadien

#### 4.2.3.1. Bewertung der Ergebnisse

Auf Basis der Einteilung in 2 Klassen mit einem Cutoff von 5 mm wurde in der Evaluation einer höhergradigen Invasionstiefe von  $> 5$  mm in diesem Kollektiv im Vergleich mit der Pathologie eine Sensitivität von 52%, eine Spezifität von 78%, ein PPW von 61% sowie ein NPW von 71% erreicht. Ein Vergleich mit der Literatur fand sich für diese Einteilung nicht.

Bezüglich erreichbarer Genauigkeiten der MRT-Messungen konnte die MERCURY-Study in ihrem Kollektiv zeigen, dass Messungen der MRT und Histopathologie innerhalb 0,5 mm Abweichung als gleichwertig anzusehen sind (bei einer Range von -0.49 mm bis 0.40 mm 95%-Konfidenzintervall) [16]. Eine Auswertung in dieser Genauigkeit konnte anhand der vorliegenden Daten für das Kollektiv aus Neuperlach nicht vorgenommen werden. Einerseits handelt es sich im Patientenkollektiv aus Neuperlach (im Gegensatz zur MERCURY-Studie) um eine retrospektive Studie, so dass nicht überprüft werden konnte, ob die Messungen der Pathologie und der MRT auf gleicher Tumorhöhe erfolgten, womit eine 1:1-Korrelation nicht möglich war. Andererseits wurden die Abstufungen der Invasionstiefe auf Seiten der MRT-Befunde nur in einem Datensatz auf 0,5 mm genau angegeben, auf Seiten der Pathologie analog nur in drei Datensätzen. In den übrigen Datensätzen fanden sich auf Seiten der Pathologie und MRT jeweils 1 mm-Abstufungen, in der MRT zusätzlich noch Intervallangaben, wie z.B.  $< 5$  mm.

Somit war eine Modalität für die Vergleichbarkeit zu finden, die sich auf die vorhandenen Datensätze anwenden ließ.

In der Literatur fand sich die Empfehlung, den Cutoff von 5 mm für eine optimierte Bildgebungsstrategie in der Routinediagnostik zu verwenden und die T3-Stadien in T3a/b sowie T3c/d einzuteilen [5].

Weiterhin konnte eine Studie mit 188 primär operativ therapierten Patienten histologisch gesicherter T3-Stadien zeigen, dass sich die Übereinstimmung zwischen den Ergebnissen zweier verschiedener Befunder in dieser Klasseneinteilung gegenüber der Feinauswertung jeder separaten Unterkategorie verbessert, besser reproduzierbar ist und somit die MRT-Befunde konsistenter werden [41].

Diese Einteilung war für das Patientenkollektiv aus Neuperlach möglich und klinisch sinnvoll und könnte in dieser Form in die standardisierte Befundvorlage im RIS integriert werden.

Bezüglich der Abweichungen war festzustellen, dass in 46 der 68 untersuchten Datensätze die Abweichung der gemessenen Infiltrationstiefe der MRT-Befunde von der Pathologie mindestens 2 mm betrug (ca. 68% der Fälle).

Es ist möglich, dass sich diese größeren Differenzen in einer zu stark unterschiedlichen Kippung zwischen den MRT-Sequenzen (die ja senkrecht auf die Tumurlängsachse eingekippt wurden, siehe 2.4.2.) und den histologischen Schnitten und den damit verbundenen unterschiedlichen Partialvolumeneffekten erklären.

In einer prospektiven Abweichungsanalyse von kleineren Fallzahlen (z.B. 20 Datensätze) könnten diese Unterschiede näher untersucht werden. Dies wäre sicherlich sinnvoll, jedoch Gegenstand einer separaten Studie.

#### 4.2.3.2. Stellenwert des MRT-Staging in Diagnostik und Therapie

T3-Stadien bilden aufgrund der unterschiedlichen Ausdehnung des Mesorektums die heterogenste Gruppe im T-Staging [46]. Gemäß den ESMO-Leitlinien kann, je nach Infiltrationstiefe des Tumors in das perirektale Fettgewebe, eine Subklassifikation der T3-Stadien in T3a-d erfolgen [20].

Der Hintergrund ist, dass die Invasionstiefe in das perirektale Fettgewebe in T3-Stadien ein wichtiger prognostischer Faktor ist, aus der sich ein individuelles Therapieschema ableiten kann. So konnte eine Studie mit hoher Patientenfallzahl (853 primär operativ therapierte Patienten mit histopathologischem T3-Stadium) zeigen, dass sich sowohl die Lokalrezidivraten als auch die 5-Jahres-Überlebensraten gegenüber einer Tumordinfiltration  $\leq 5$  mm signifikant unterscheiden. Die Lokalrezidivraten betragen 10,4% (T3a/b) bzw. 26,3% (T3c/d). Auch die Überlebensraten von 85,4% bzw. 54,1% unterschieden sich stark [37].

Die prognostische Relevanz einer Infiltration  $> 5$  mm fand sich auch in vielen anderen Studien [5,48,46,1]. Das separate Erkennen eines T3a-Stadiums scheint jedenfalls hinsichtlich des Outcomes nicht so relevant zu sein, da sich dieses nicht signifikant von einem T2-Stadium unterscheidet. Selbiges gilt auch für das Stadium T3b, sofern keine Bedrohung der mesorektalen Faszie vorliegt [5].

In der aktuellen S3-Leitlinie „Kolorektales Karzinom“ wird diesen Unterschieden in individuellen Therapieempfehlungen bereits Rechnung getragen. Für UICC Stadien II/III im mittleren Drittel findet sich hier als Ausnahmefall die Empfehlung einer primär operativen (und nicht zusätzlich neoadjuvanten) Therapie für T3a/b-Tumore im MRT-Staging, sofern kein bildgebender Verdacht auf Lymphknotenmetastasen oder eine EMVI besteht [28].

### **4.3. CRM-Status**

#### **4.3.1. Gütekriterien des CRM-Status**

In der Beurteilung des CRM-Status betrug die Sensitivität der vorliegenden Studie 78%, die Spezifität 84%, der NPW hohe 99%, der PPW jedoch nur 17%.

##### **4.3.1.1. Vergleich der Ergebnisse mit der Literatur**

###### **4.3.1.1.1. Beurteilung eines negativen CRM**

Es gibt mehrere große Studien, die den CRM-Status des MRT-Staging untersucht haben und zeigen konnten, dass CRM-Negativität mittels präoperativem MRT-Staging mit einer hohen Genauigkeit vorhergesagt werden kann. Dies ist an der Spezifität und am negativ prädiktiven Wert (NPW) ablesbar.

In der OCUM-Studie, einer großen, multizentrischen prospektiven Beobachtungsstudie mit 642 Patienten, wurde ein negativer CRM-Status der MRT bei 232 der 254 primär operierten Patienten gesehen (91,3%), der NPW betrug dort 98,3% [43].

Einen ähnlich hohen NPW zeigte auch die MERCURY-Studie mit 94% [52].

Im Patientenkollektiv aus Neuperlach wurde ein negativer CRM-Status der MRT bei 186 von 228 untersuchten, primär operierten Patienten gesehen (81,5%). Der NPW betrug 98,9%. Die Vorhersagegenauigkeit eines negativen CRM war somit in der Standarddiagnostik sehr gut zu treffen und ähnlich hoch im Vergleich mit relevanten Studien.

Eine Metaanalyse von Zhang et al. (2016) fasste die Ergebnisse von 35 Studien zur CRM-Evaluation zusammen. Etwaige Inhomogenitäten der unterschiedlichen

Studien, insbesondere der Einfluss unterschiedlicher Kriterien für die CRM-Positivität (Tumorabstand  $\leq 1$  mm bis zu 5 mm und andere) wurden in dieser Metaanalyse bereits statistisch berücksichtigt und der entsprechende Einfluss auf die Ergebnisse eliminiert.

Die Metaanalyse errechnete bei der Evaluation des CRM-Status von retrospektiven Studien mit primär operativ therapierten Patienten eine gepoolte Spezifität von 82% (bei einer Range von 76-87% im 95%-Konfidenzintervall) [56].

Im Patientenkollektiv Neuperlach konnte eine Spezifität von 84% erreicht werden. Große prospektive Studien, wie z.B. die MERCURY-Studie, erreichten zwar höhere Ergebnisse: Die Spezifität der MERCURY-Studie betrug im Patientenkollektiv der primär operierten Patienten 91% [52]. Bei prospektiven Studien wurde in der Metaanalyse von Zhang et al. (2016) jedoch tendenziell eine höhere Spezifität und generell eine leicht bessere Performance der MRT bei der CRM-Evaluation gefunden, die Spezifität für prospektive Studien wurde mit 96% angegeben (bei einer Range von 95-97% im 95% Konfidenzintervall) [56].

Das bessere Abschneiden prospektiver Studien könnte darin begründet liegen, dass in diesen Studien die diagnostischen Prozesse und die interdisziplinäre Zusammenarbeit idealerweise auf das Studienziel hin optimiert und aufeinander abgestimmt sind. Prozessoptimierung auf ein Studienziel kann im Setting einer prospektiven Studie sicherlich gut verwirklicht werden.

Oft gibt es sowohl auf Seiten der Radiologie als auch auf Seiten der Pathologie Studienbetreuer, die entsprechend fachlich besonders gut qualifiziert und geschult sind, was z.B. die Konsistenz der Befunde verbessern kann. Das Ziel dieser Studie war jedoch die Standarddiagnostik abzubilden, in der die Herangehensweise eine andere ist, ein Querschnitt der Standarddiagnostik kann die Abweichung in den Ergebnissen damit sicherlich teilweise erklären.

#### 4.3.1.1.2. **Beurteilung eines positiven CRM**

Die Beurteilung eines positiven CRM drückt sich in den Kenngrößen der Sensitivität sowie des positiv prädiktiven Werts aus.

Ma, Xiaoji et al. (2015) errechneten für 7 große Studien eine Bandbreite der Ergebnisse des PPW von niedrigen 20% bis maximal hohen 89%. Für ihr eigenes Studienkollektiv (203 primär operativ therapierte Patienten) errechneten sie einen PPW von niedrigen 21%. Für die MERCURY-Studie gaben die Autoren einen ebenfalls

nur mäßigen PPW von 53% an [38]. Im Patientenkollektiv von Neuperlach konnte lediglich ein sehr niedriger PPW von 17% erreicht werden.

Die Sensitivität der MRT bei der Evaluation des CRM-Status fällt in vielen Studien geringer aus als die Spezifität. Entgegen der Spezifität von 82% wurde in der bereits unter (4.3.1.1.1.) zitierten Metaanalyse bei retrospektiven Studien eine etwas geringere Sensitivität von 79% angegeben (bei einer Range von 66%-88% im 95%-Konfidenzintervall). Die Sensitivität des Studienkollektivs aus Neuperlach von knapp 78% (77,8%) lag auf einem vergleichbaren Niveau und fiel ebenfalls geringer aus als die Spezifität.

Für prospektive Studien errechnete die Metaanalyse eine gepoolte Sensitivität von 77% (bei einer Range von 70% - 84% im 95%-Konfidenzintervall) [56].

Ma, Xiaoji et al. (2015) errechneten für die prospektive MERCURY-Studie eine Sensitivität von 64%, die sogar sehr deutlich unter der erreichten Spezifität lag [38].

#### 4.3.1.2. Einflussfaktoren auf die Ergebnisse des CRM-Status

##### 4.3.1.2.1. Wahl des Cutoffs

Die Evaluation des CRM der MRT beruht auf einer Abstandsmessung. Entgegen den TNM-Kriterien für das T-Stadium sind die Kriterien für den CRM-Status der MRT nicht einheitlich definiert, insbesondere wird der Mindestabstand einer Tumormarkierung zur mesorektalen Faszia bzw. zum potenziellen zirkumferenziellen Resektionsrand in verschiedenen Studien unterschiedlich gehandhabt und teils kontrovers diskutiert. Der Cutoff beeinflusst die Ergebnisse der CRM-Evaluation der MRT jedoch stark:

In einer Metaanalyse aus dem Jahr 2014, die 14 Studien mit insgesamt 1.600 Patienten auswertete, ist dieser Effekt anschaulich erkennbar.

Je nachdem, ob als Kriterium ein Abstand  $\leq 1$  mm,  $\leq 2$  mm oder  $\leq 5$  mm gewählt wurde, betrug die gepoolte Sensitivität 76%, 79% bzw. 92%. Die Spezifität betrug jeweils 88%, 66% bzw. 48% [54]. Dies zeigt, dass ein höherer Cutoff sich zwar zu Gunsten der Sensitivität auswirken kann, dies jedoch unter Inkaufnahme einer sich zunehmend verringernden Spezifität. Die Studie kam zu dem Schluss, dass der Cutoff  $\leq 1$  mm zu präferieren ist. Neben der höchsten DOR dieses Cutoffs kann ein höherer Cutoff wie z.B. 5 mm das Risiko einer Übertherapie erhöhen [54].

Der Cutoff von  $\leq 1$  mm ist in der aktuellen Literatur am gebräuchlichsten und stellt damit die Beurteilungsgenauigkeit eines negativen CRM in den Vordergrund. Der Cutoff für das Patientenkollektiv aus Neuperlach wurde ebenfalls mit  $\leq 1$  mm gewählt und führte konsekutiv zu einer relativ hohen Spezifität.

#### 4.3.1.2.2. **Patientenkollektiv**

Die vorliegende Studie betrachtet lediglich primär operativ therapierte Patienten. Das untersuchte Kollektiv hat damit per se charakteristische Merkmale:

Diejenigen Patienten, die einen eindeutig positiven CRM-Status der MRT hatten, wurden nahezu ausschließlich neoadjuvant therapiert und bildeten somit ein Ausschlusskriterium für diese Studie.

Auf Seiten der Pathologie führte dies zu einer geringen Prävalenz histopathologisch CRM-positiver Patienten (9 Patienten) im Kollektiv.

Aus radiologischer Sicht verblieben im Kollektiv einerseits tendenziell diagnostische Grenzfälle für CRM-Positivität der MRT. Andererseits sind in diesem Kollektiv die Anzahl der Fälle mit eindeutig negativem CRM-Status tendenziell als hoch anzusehen. Die genannten Besonderheiten wirken sich auf den PPW und den NPW diametral aus.

Der PPW wird in Studien selten explizit ausgewiesen und entsprechend diskutiert. In dem Kollektiv primär operierter Patienten wirken sich diagnostische Grenzfälle und eine geringe Prävalenz jedoch *ungünstig* auf den PPW aus:

Diagnostische Grenzfälle erhöhen einerseits potenziell die Wahrscheinlichkeit *falsch*-positiver Ergebnisse, andererseits ist eine geringe Prävalenz gleichbedeutend mit einer geringen Fallzahl maximal potenziell *richtig*-positiv zu detektierender Fälle. Diese Kombination wirkt sich gleichlautend ungünstig auf den PPW aus, somit ist ein eher geringer PPW in diesem Kollektiv und ähnlichen Kollektiven nicht verwunderlich.

Umgekehrt wirken sich Fälle mit eindeutig negativem CRM-Status der MRT einerseits auf die Wahrscheinlichkeit *richtig*-negativer Ergebnisse positiv aus.

Die geringe Prävalenz CRM-positiver Patienten der Pathologie ist andererseits auch gleichbedeutend mit einer potenziell geringen Anzahl der maximal möglich zu detektierenden *falsch*-negativen Fälle. Diese Kombination wirkt sich daher, genau umgekehrt zum PPW, gleichlautend potenziell günstig auf den NPW aus.

Das unbeeinflusste Potenzial der MRT-Diagnostik kann somit aufgrund der Auswahl des Patientenkollektivs nicht gänzlich abgebildet werden. Jedoch kann auch in einem Kollektiv, das sowohl primär operativ therapierte Patienten als auch neoadjuvant therapierte Patienten enthält, das unbeeinflusste Potenzial der MRT-Diagnostik ebenfalls nicht genau abgebildet werden, da therapeutische Effekte wie Tumorschrumpfung oder Narbenbildung den Lokalbefund verändern und damit die Vergleichbarkeit zwischen MRT-Befunden und Histopathologie erschweren und die Ergebnisse dementsprechend beeinflussen.

In dieser Studie sollte jedoch das MRT-Staging zumindest unbeeinflusst von therapeutischen Effekten untersucht werden, um die maximal mögliche Vergleichbarkeit mit den Ergebnissen der Pathologie zu ermöglichen.

#### 4.3.1.2.3. Kriterien für CRM-Positivität

Die Kriterien zur Bewertung eines positiven CRM-Status in der vorliegenden Studie waren insofern weit gefasst, als dass neben einer extramuralen Tumorformation (also Tumorstadien  $\geq T3$ ) bzw. eines suspekten Lymphknotens (was in verschiedenen Studien unterschiedlich gehandhabt wird) bereits MR-befundliche T2- und sogar T1-Stadien, deren tumortragende Darmwand  $\leq 1$  mm von der mesorektalen Faszie entfernt blieb, als „CRM-positiv“ zu werten waren (siehe 2.4.4.4.).

In einer alternativen Berechnung wurde der Einfluss der Darmwandnähe von MR-befundlichen T1- und T2-Stadien und zusätzlich die des Lymphknotenstatus (suspekte Lymphknoten in  $\leq 1$  mm Entfernung zur mesorektalen Faszie) auf die CRM-Positivität des MRT-Stagings im Patientenkollektiv aus Neuperlach eliminiert (8 Patienten). Diese enger gefasste Betrachtungsweise hätte in jedem Datensatz zu einer Reduktion der falsch-positiven Ergebnisse geführt. Konsekutiv hätte sich ein um 4% gesteigerter PPW für das Gesamtkollektiv auf weiterhin niedrige 21%, ein unwesentlich um 0,2% leicht verbesserter NPW von 99,0%, jedoch eine nochmals um 4% auf 88% verbesserte Spezifität bei unveränderter Sensitivität errechnet.

### 4.3.2. Reliabilität des CRM-Status

In dieser Studie sollte die Übereinstimmung der Ergebnisse des CRM-Status mit der Pathologie untersucht werden, dieser Zusammenhang wurde mittels Kappa-Statistik evaluiert.

#### 4.3.2.1. Reliabilität des Gesamtkollektivs

Es ergab sich ein zonenunabhängiger Overall-Kappa für das Gesamtkollektiv von 0,22, der als ausreichende Übereinstimmung zu werten ist.

#### 4.3.2.2. Zonenabhängige Reliabilität

In der Feinauswertung der anatomischen Zonen A-C lag die Übereinstimmung in jeder Zone auf einem unterschiedlichen Niveau. Es fand sich ein ausreichendes Niveau der Zone B, das auf dem gleichen Niveau des Gesamtkollektivs lag.

Die Übereinstimmung im CRM-Status war in Zone C mit einem moderaten Niveau am besten und lag sogar über dem Niveau des Gesamtkollektivs. In den Ergebnissen der Zone C können jedoch die geringen Fallzahlen von lediglich 18 Datensätzen eine Rolle gespielt und zu ungenauen Ergebnissen geführt haben.

Die Übereinstimmung der Zone A fiel nur gering aus. Die Übereinstimmung war mit einem Kappa-Wert von nur 0,04 auffallend deutlich geringer als in den übrigen anatomischen Zonen. In Zone A waren die Fallzahlen mit 117 Datensätzen am höchsten. Hier ist nicht davon auszugehen, dass die Fallzahlen zu ungenauen Ergebnissen geführt haben und der erreichte Kappa-Wert ist daher mit Skepsis zu betrachten.

Um nach Gründen für die starke Abweichung von den übrigen Ergebnissen zu suchen, wurden die Datensätze der Zone A von einer zweiten Befunderin, einer erfahrenen Oberärztin, die regelmäßig an Schulungen und Fortbildungen zum MRT-Staging des Rektumkarzinoms teilnimmt, reevaluiert. Die Befunderin wäre, ohne Kenntnis der Ergebnisse der Pathologie, aufgrund ihrer Messungen in drei Fällen zu einem abweichenden CRM-Status gelangt. Anstatt eines positiven CRM-Status wäre stattdessen ein negativer CRM-Status vergeben worden. Dies hätte zwar mit den Ergebnissen der Pathologie übereingestimmt und damit die Übereinstimmung in Zone A verbessert, hierbei handelt es sich jedoch um eine interindividuelle

Befundabweichung zwischen zwei Befundern, die in der Standarddiagnostik erfahrungsgemäß des Öfteren vorkommt.

In drei Befunden jedoch fand sich ein positiver CRM-Status aufgrund einer Bedrohung/Infiltration der vermeintlichen mesorektalen Faszie, der in der Nachbetrachtung als falsch-positiv zu werten war. In allen drei Fällen handelte es sich um anterior gelegene Tumorentitäten.

Es stellte sich heraus, dass in diesen drei Fällen das viszerale Peritoneum irrtümlicherweise für die mesorektale Faszie gehalten wurde, da auf der entsprechenden Tumorphöhe anteriorseitig keine mesorektale Faszie mehr existiert. Hierbei handelte es sich also um eine systematische Fehlbefundung, die untersucherunabhängig vermieden werden kann. Diese drei falsch-positiven Befunde erklären zumindest 14% der falsch-positiven Befunde des Patientenkollektivs in Zone A sowie 8,5% der falsch-positiven Befunde der CRM-Evaluation des Gesamtkollektivs.

In mehreren Studien wurde die CRM-Evaluation bei anterioren Tumorentitäten ausgewertet. Ma, Xiaoji et al. (2015) untersuchten die Beziehung zwischen dem CRM-Status der MRT und der Pathologie und berichteten über eine Häufung falsch-positiver Befunde bei anterioren Tumorentitäten. In ihrem Kollektiv wurde nur bei 3 von 24 Patienten anterior gelegener Tumoren der positive CRM-Status der MRT in der Pathologie bestätigt (12,5%) [38].

Schwierigkeiten bei der Diagnostik anteriorer Tumore können anatomisch begründet sein:

Aboral der peritonealen Umschlagsfalte grenzt der anteriore Anteil der mesorektalen Faszie eng an die Rückwand der Vagina bzw. der Samenblasen und Prostata an (und bildet bei letzteren ein Band dichten Bindegewebes, die „Denonvillier'sche Faszie“) [42]. Auch ist die mesorektale Faszie auf dieser Höhe anteriorseitig nicht leicht zu identifizieren, da sie relativ dünn und durchscheinend ist. Auch einen Beckenkonkavität kann die Befundung erschweren [31].

In der Studie von Ma, Xiaoji et al. (2015) wurde diskutiert, dass auf dieser anatomischen Höhe nach präoperativ positivem CRM-Status der MRT (im Gegensatz zu posterioren/lateralen Tumoren) erweiterte Resektionen (wie z.B. Resektion der Samenbläschen oder der Rückwand von Vagina oder Prostata) besser möglich seien, um postoperativ doch noch einen negativen zirkumferenziellen Resektionsrand zu erreichen, was die Abweichung der Ergebnisse ihres Patientenkollektivs von den Ergebnissen der Pathologie teilweise erklären könne [38]. Dieser Zusammenhang

könnte zwar auch bei den Kennzahlen der Ergebnisse des Patientenkollektivs aus Neuperlach im CRM-Status eine Rolle gespielt haben (insbesondere dem niedrigen PPW, siehe 3.3.1.1.), beschreibt jedoch tiefer gelegenen Tumorentitäten. Die o.g. Fehlbefunde sind also durch diesen Zusammenhang nicht erklärt.

Eine prospektive Studie, die ebenfalls explizit Faktoren untersuchte, die zu einer inakkuraten Vorhersage im CRM-Status an 66 Patienten untersuchte, kam zu dem Schluss, dass eine anteriore Tumorlokalisation sowie ein anterior dünner ausgeprägtes perirektales Fettgewebe mit einer inkorrekten Vorhersage im CRM-Status assoziiert waren. Darüber hinaus waren ein T3- Stadium und Lymphknotenbefall mit einer Fehleinschätzung im CRM-Status assoziiert. Eine mögliche Verwechslung der mesorektalen Faszie mit dem Peritoneum anteriorseitig wurde auch hier nicht explizit diskutiert, es fand sich jedoch der Hinweis, dass Tumorentitäten oberhalb der peritonealen Umschlagfalte häufig gefunden wurden und die Messungen zur Bestimmung des CRM-Status in diesen Fällen nicht verlässlich gewesen wären (und daher Patienten mit hoch sitzenden Rektumkarzinomen als Ausschlusskriterium der Studie gegolten hätten) [31].

Zumindest dieser Hinweis könnte die Ergebnisse des Teilkollektivs hoch sitzender Tumoren der vorgelegten Studie Neuperlach teilweise erklären, in welchem ein sehr niedriger Kappa-Wert bei hoch sitzenden Tumoren der Zone A Ausdruck einer geringen Übereinstimmung mit den Ergebnissen der Pathologie war, die Verlässlichkeit der Ergebnisse also bei hoch gelegenen Tumorentitäten in der vorgelegenen Studie ebenfalls gering war.

Ein hoher Tumorsitz oberhalb der peritonealen Umschlagfalte als Ursache falsch-positiver Ergebnisse im CRM-Status wurde in den beiden zitierten Studien zwar nicht expliziert diskutiert. Es ist jedoch zumindest anzunehmen, dass dieser Pitfall (die Verwechslungsmöglichkeit der mesorektalen Faszie mit dem viszeralem Peritoneum) in der CRM-Evaluation auch bei anderen Patientenkollektiven in der Routinediagnostik eine potenzielle Fehlerquelle bei hoch anterior gelegenen Tumorentitäten bilden kann.

#### **4.3.3. Einfluss der Tumorphöhe auf die Reliabilität des CRM-Status**

Die Beurteilung, ob die Tumorphöhe entsprechend der anatomischen Zonen A-C Einfluss auf die Übereinstimmung im CRM-Status mit der Pathologie genommen hat,

wurde mittels Konfidenzintervallanalyse der Kappa-Statistik evaluiert. Aufgrund der starken Überlappung der Konfidenzintervalle muss die Frage, ob die Tumorphöhe Einfluss auf die Reliabilität im CRM-Status genommen hat, verneint werden.

Weiterhin konnte eine zufällige Übereinstimmung basierend auf der Verteilung im Gesamtkollektiv sowie in den anatomischen Zonen A und -B nicht ausgeschlossen werden, die Kappa-Werte der Konfidenzintervallanalyse waren also in diesen Teilkollektiven (mehr oder weniger) nicht signifikant. Lediglich die Ergebnisse der Zone C waren statistisch signifikant.

Bei der Interpretation der Konfidenzintervallanalyse gilt es jedoch zu bedenken, dass das Patientenkollektiv primär operativ therapierter Patienten per se charakteristische Merkmale besitzt, insbesondere sind die Fallzahlen CRM-positiver Patienten in der pathologischen Aufarbeitung in dieser und vergleichbaren Studien relativ gering und betragen im Gesamtkollektiv dieser Studie entsprechend lediglich 9 Fälle (siehe auch 4.3.1.2.2.). Der Grund hierfür ist, dass eindeutig CRM-positive Fälle der MRT nicht primär operiert wurden und deshalb ein Ausschlusskriterium für die Studie bildeten. Es ist davon auszugehen, dass diese kleinen Fallzahlen CRM-positiver Patienten starken Einfluss auf die Ergebnisse nahmen: Die Aussagekraft der Kappa-Statistik ist stark abhängig vom Vorhandensein positiver Fälle im Kollektiv, die mit einer bestimmten Sensitivität erkannt werden. Aufgrund der inhärenten mathematischen Definitionen und interdependenten Abhängigkeiten wesentlicher Parameter, die der Konfidenzintervallanalyse mittels Kappa-Statistik zugrunde liegen, würden sich sowohl die Kappa-Werte selbst, die z-Werte,  $\sigma$ -Werte (und konsequent auch die Konfidenzintervalle) schon beim theoretischen Hinzutreten einiger weiterer positiver Fälle, die mit einer fixen Sensitivität erkannt werden, in diesem Fallzahlenbereich noch stark verändern, was rein statistisch für die Aussagekraft der Kappa-Statistik wünschenswert gewesen wäre [36].

#### **4.3.4. Bewertung der Ergebnisse**

Es lässt sich feststellen, dass sich der negative Vorhersagewert (NPW) im Patientenkollektiv aus Neuperlach in der CRM-Evaluation im Vergleich mit der Literatur auf einem hohen Niveau befand. Auch bei der Spezifität konnte (unter

Berücksichtigung der Limitation des retrospektiven Studiendesigns) im Vergleich mit der Literatur ein gutes Ergebnis erreicht werden.

Die Ergebnisse in der Evaluation eines positiven CRM im Vergleich mit der Pathologie fielen niedriger aus. Die Sensitivität lag im Vergleich mit der Literatur in einem mittleren Bereich und war niedriger als die Spezifität. Diese Tendenz deckte sich ebenfalls mit den Ergebnissen anderer Studien. Die niedrigere Sensitivität hängt hierbei wohl maßgeblich mit dem gewählten Cutoff für CRM-Positivität zusammen. Der positive Vorhersagewert fiel sehr gering aus, auch dies deckte sich mit den Ergebnissen anderer Studien vergleichbarer Kollektive.

Der tendenziell niedrige PPW und hohe NPW in Kollektiven primär operativ therapierter Patienten dieser und anderer Studien erklärt sich wohl maßgeblich aus der Zusammensetzung des Kollektivs und dessen charakteristischen Merkmalen (siehe 4.3.1.2.2).

Es gab in der CRM-Evaluation des Studienkollektivs Potenzial zur Reduktion falsch-positiver Ergebnisse:

Die diagnostische Unsicherheit bzgl. des MRT-Nodalstagings wird unter 4.4. diskutiert. Suspekte Lymphknoten sollten aufgrund der hohen diagnostischen Unsicherheit im präoperativen MRT-Staging zukünftig nicht mehr als Kriterium für CRM-Positivität herangezogen werden. Dies hätte in der Evaluation des CRM-Status potenziell zu einer Verbesserung der Spezifität beigetragen und in der Konsequenz das Risiko einer etwaigen Übertherapie, die mit einem falsch-positivem CRM-Status aufgrund einer Lymphknotenfehlbeurteilung einhergehen kann, vermindert. Auch in anderen Studienkollektiven war das Nodalstaging mit einer Fehlbeurteilung des CRM assoziiert.

Die CRM-Positivität aufgrund des Kriteriums der Darmwandnähe zur mesorektalen Faszie bei T1- und T2-Stadien im MRT-Staging führte in diesem Patientenkollektiv ausschließlich zu falsch-positiven Ergebnissen und sollte ebenfalls zumindest überdacht und in weiteren Auswertungen reevaluiert werden.

Eine Alternative wäre, die Darmwandnähe zur mesorektalen Faszie bei T1- und T2-Stadien gesondert zu kennzeichnen, z.B. anhand einer Änderung der standardisierten Formatvorlage mit entsprechendem Hinweissfeld, jedoch nicht als Kriterium eines positiven CRM-Status.

Dies hätte den Vorteil, dass die entsprechende Information nicht verloren ginge und in ihrer Funktion „als Frühwarnsystem“, erhalten bliebe. Die Information der Darmwandnähe zur mesorektalen Faszie könnte für die Operateure dahingehend hilfreich sein, in der jeweiligen anatomischen Höhe die beste entsprechende Resektionsebene zu finden, bzw. Resektionsebenen zu identifizieren, in welchen aufgrund schwieriger anatomischer Verhältnisse in der chirurgischen Exploration besondere Achtsamkeit gefragt ist. Dieser Hinweis könnte z.B. im Rahmen radiologischer Fall-Demonstrationen erfolgen.

Gleichzeitig könnten falsch-positive Befunde im CRM-Status vermieden werden, die ggf. therapeutisch relevant sind.

Basierend auf den auffälligen Einzelergebnissen der Reliabilität im CRM-Status fand sich eine systematische Fehlerquelle in der CRM-Evaluation im Teilkollektiv der oberhalb der peritonealen Umschlagfalte gelegenen hoch sitzenden Tumoren, die auf einer Verwechslungsmöglichkeit der mesorektalen Faszie mit dem viszeralen Peritoneum beruhte.

Um diesen Pitfall zu vermeiden, sollte die Lagebeziehung des Tumors zur peritonealen Umschlagfalte bei hoch sitzenden Tumoren zukünftig dokumentiert werden, am besten anhand einer geänderten Befundvorlage mit entsprechendem Auswahl-feld. Eine standardisierte Evaluation der Tumorrelation zur peritonealen Umschlagfalte und die damit verbundene Bewusstmachung der Tumorphöhe könnte sicherlich eine Fehlinterpretation der Tumorrelation zur mesorektalen Faszie besser vermeiden und damit falsch-positive Ergebnisse im CRM-Status weiter minimieren.

In dieser Studie sollte untersucht werden, ob die Tumorphöhe Einfluss auf die Übereinstimmung im CRM-Status mit der Pathologie nahm. Diese Bewertung konnte in der Konfidenzintervallanalyse aufgrund der geringen Fallzahlen CRM-positiver Patienten im Kollektiv nicht vorgenommen werden.

#### **4.3.5. Stellenwert des MRT-Staging in Diagnostik und Therapie**

Die CRM-Evaluation ist für die Auswahl eines optimalen multimodalen Therapieschemas unentbehrlich, auch wenn sie nicht im TNM-Schema enthalten ist.

Die mesorektale Faszie fungiert als Barriere für die lokale Tumorausbreitung und markiert den chirurgischen zirkumferenziellen Resektionsrand. Der Tumorabstand

zum zirkumferenziellen Resektionsrand ist hierbei der wichtigste Risikofaktor für das Auftreten eines Lokalrezidivs [4].

Die Beurteilung der CRM-Negativität ist in Studien essenziell für die Rechtfertigung rein operativer Therapieerfahren. Die MRT kann den potenziellen Mindestabstand eines Tumors zum zirkumferenziellen Resektionsrand innerhalb eines Millimeters Distanz mit hoher Genauigkeit bestimmen, was in der Histopathologie als Schrittrandnegativität gilt [5].

In einer Metaanalyse, die 21 Studien von primär operativ therapierten Patienten mit Rektumkarzinomen zusammenfasste, war die MRT in der CRM-Evaluation sogar spezifischer als in der T-Kategorie des TNM-Systems (in der Sensitivität fanden sich hier jedoch keine signifikanten Unterschiede) [2].

Die MERCURY Study Group konnte zeigen, dass die MRT im präoperativen Staging die beste verfügbare Methode ist, einen potenziell erreichbaren tumorfreien zirkumferenziellen Resektionsrand vorherzusagen [10].

Patienten mit positivem CRM-Status der MRT haben durch neoadjuvante Radiochemotherapie zumindest eine erhöhte Chance, postoperativ einen freien zirkumferenziellen Resektionsrand zu erreichen [21].

Auch wenn ein positiver CRM-Status des MRT nicht sicher geeignet scheint, einen positiven postoperativen zirkumferenziellen Resektionsrand vorherzusagen (siehe 4.3.1.1.2), so hat ein positiver CRM-Status der MRT dennoch prognostische Bedeutung. Die MERCURY Study Group kam in einer multivariaten Analyse zu dem Schluss, dass die MRT-gestützte CRM-Evaluation sogar die einzige präoperative Staging-Methode ist, die signifikant für die Beurteilung der Gesamtüberlebensraten, tumorfreien Überlebensraten sowie für das Auftreten eines Lokalrezidivs ist und diesbezüglich den TNM-Parametern überlegen [52].

Die Lokalrezidivrate, tumorfreie Überlebensrate sowie Gesamtüberlebensrate der Patienten mit negativem CRM-Status der MRT betragen in ihrem Studienkollektiv im 5-Jahres follow up 7%, 67% bzw. 62%, wohingegen sich bei positivem CRM-Status Raten von 20%, 47% bzw. 42% fanden [52]. Der Unterschied im onkologischen Outcome ist hier deutlich zu erkennen. Für das Patientenkollektiv aus Neuperlach liegen diesbezüglich keine Daten vor, dies war nicht Gegenstand der Studie.

Die OCUM-Studie, eine große, multizentrische prospektive Beobachtungsstudie rückte ebenfalls die CRM-Evaluation der MRT in den Fokus: Hier wurde eine

risikostratifizierte Indikation zur neoadjuvanten Radiochemotherapie lediglich für Rektumkarzinome  $\geq 6$  cm ab Anokutanlinie nur bei CRM+ Konstellation der MRT gestellt, unabhängig von der T- und N-Kategorie (Niedrigrisikogruppe). Tumorentitäten im unteren Drittel ( $< 6$  cm) wurden ab einem T3-Stadium im präoperativen MRT-Staging neoadjuvant therapiert (Hochrisikogruppe).

Im Patientenkollektiv der OCUM-Studie wurde bei 428 Patienten in der Nachbeobachtung im 5-Jahres-Follow-Up eine Lokalrezidivrate von 2,7% gefunden. Die Rezidivraten waren hierbei in der Patientengruppe, die mittels alleiniger Operation therapiert wurden (254 Patienten) und der zusätzlich neoadjuvant mit Radiochemotherapie behandelten Patientengruppe (174 Patienten), sogar identisch [43].

Die OCUM-Studie kam daher zu dem Schluss, dass eine Beschränkung der neoadjuvanten Therapie auf Hochrisikogruppen ein gutes Outcome für den Patienten ermöglicht.

Diese beiden großen Studien konnten somit den hohen Stellenwert des MRT CRM-Staging besonders gut darlegen, was auch in zahlreichen weiteren Studien bestätigt wurde.

In der aktuellen S3-Leitlinie wird ein negativer CRM-Status der MRT (basierend auf dem Abstand des Primärtumors bzw. eines bildgebend suspekten Lymphknotens von der mesorektalen Faszie) außerhalb von Studien noch nicht als ausschließliches Entscheidungskriterium für eine primäre Operation empfohlen.

Die Leitlinie stellt heraus, dass es noch weiterer, qualitätsgesicherter, prospektiver Studien bedarf, auch wenn sie dem MRT-CRM-Status sowie dem histopathologischen CRM bereits eine hohe prognostische Bedeutung für die lokale Kontrolle sowie dem krankheitsfreien Überleben und Gesamtüberleben beimisst [28].

#### **4.4. N-Status**

Alle in der Studie enthaltenen Patienten, die im Vergleich des N-Status betrachtet wurden, wurden im Rahmen einer TME inklusive der mesorektalen Lymphknoten reseziert.

Im Vergleich mit der Pathologie konnte eine hohe Sensitivität von 85% sowie eine Spezifität von 45% bei insgesamt 232 untersuchten Patienten erreicht werden.

##### **4.4.1. Ergebnisse im Vergleich mit der Literatur**

In der Literatur variieren die Angaben zur Sensitivität und Spezifität des MRT-Lymphknotenstaging stark.

Mehrere Studien geben die Sensitivität mit 60-70%, die Spezifität mit 75-85% an [7,14,19,49]. Eine große Metaanalyse über die Evaluation von Lymphknotenmetastasen bei Rektumkarzinomen aus dem Jahr 2015 fasste 55 Studien mit insgesamt 2845 Patienten zusammen und gibt die Sensitivität in der MRT-Diagnostik bei Scannern  $\leq 1,5$  Tesla mit einer Range von 68% bis 80% an, die Spezifität mit einer Range von 71% bis 81%, jeweils im 95%-Konfidenzintervall [33]. Das Patientenkollektiv aus Neuperlach wurde ebenfalls mit einem 1,5-Tesla-Scanner untersucht.

Hieran gemessen konnte in unserem Patientenkollektiv eine hohe Sensitivität, jedoch nur eine geringe Spezifität erreicht werden.

Eine Metaanalyse bildet aufgrund der großen Patienten- und Studienzahlen sicherlich eine gute Näherung über potenziell erreichbare Genauigkeiten ab. Es ist jedoch in der Kumulation nicht ersichtlich, wie die einzelnen Studien zu den jeweiligen Ergebnissen kamen, obwohl dies für die kritische Beurteilung der Ergebnisse essenziell ist. Unter anderem nehmen Malignitätskriterien, technische Limitationen, untersucherabhängige Interpretationen und auch das Studiendesign erheblichen Einfluss auf die Ergebnisse.

##### **4.4.2. Diskussion der Malignitätskriterien**

Abweichend z.B. vom T-Stadium, in dem die Kriterien der Pathologie in der transmuralen Tumorausdehnung 1:1 auf die MRT-Bildgebung übertragen und verglichen werden können, gibt es aufgrund der unterschiedlichen Untersuchungsmodalität keine übereinstimmende Definition im N-Stadium und in der Literatur keinen

einheitlichen Konsens über die Definition eines pathologischen Lymphknotens im MRT-Staging.

So werden in manchen Studien immer noch metrische Kriterien für die Beurteilung eines Lymphknotens definiert. Es wird z.B. ein Kurzachsendurchmesser von  $> 0,5$  cm als das Malignitätskriterium mit der höchsten Genauigkeit angegeben [35] oder in der Praxis als Cutoff für einen benignen Lymphknoten angewandt [22]. Die in der Praxis verwendeten Cutoffs liegen gewöhnlich im Bereich von 5-8 mm [27].

Die rationale Grundlage hierfür ist jedoch fraglich. So konnte z.B. eine Studie mit 196 untersuchten Patienten mit Rektumkarzinomen, von denen 98 nodal-positiv waren, zeigen, dass die befallenen Lymphknoten (pararektal sowie entlang der A. rectalis superior) in den Ergebnissen der Pathologie in 32% der Fälle 5 mm oder kleiner waren, in 8% der Fälle sogar  $\leq 2$  mm [17]. Wie sollte man also einen geeigneten Cutoff allein an metrischen Kriterien sinnvoll definieren können, wenn man damit beispielsweise – wie in diesem Patientenkollektiv gezeigt – knapp 1/3 falsch-negative Ergebnisse in Kauf nehmen würde?

Abgesehen von dem diagnostischen Dunkelfeld bei der Verwendung eines Cutoffs ist auch zu bedenken, dass es große Schnittmengen zwischen reaktiven, metastatischen und blanden Lymphknoten gibt, wie z.B. eine systematische histologische Aufarbeitung von über 12.000 Lymphknoten gezeigt hat [18].

Übereinstimmend fand eine prospektive Studie mit 1:1 Korrelation von Lymphknoten des MRT-Stagings mit der Pathologie ebenfalls große Schnittmengen zwischen blanden und pathologischen Lymphknoten, hier bei einer Bandbreite der gefundenen blanden Lymphknoten von 2 mm bis 10 mm sowie einer Bandbreite der pathologischen Lymphknoten von 3 mm bis 15 mm. Die Überschneidung der Bandbreiten ist in diesem Beispiel deutlich zu erkennen.

Damit wird im Umkehrschluss auch klar, dass metrische Kriterien wie ein Cutoff die Sensitivität und Spezifität stark beeinflussen, indem sie nicht nur zu falsch-negativen, sondern auch zu falsch-positiven Ergebnissen führen können.

In derselben prospektiven Studie konnte auch gezeigt werden, dass morphologische Kriterien mit einer Sensitivität von 85% und einer Spezifität von 95% die besten Ergebnisse brachten, wobei sich die Ergebnisse durch zusätzliche metrische Kriterien nicht verbessern ließen [12].

Letztlich muss man sich für ein Malignitätskriterium entscheiden, dass nicht deckungsgleich mit den Kriterien und Methoden der Pathologie ist und dennoch

großen Einfluss auf die Ergebnisse im Vergleich des N-Status und ggf. auf die Therapie nimmt. Im Patientenkollektiv aus Neuperlach wurden dieselben morphologischen Malignitätskriterien wie in der o.g. prospektiven Studie gewählt, nämlich 1.: Unregelmäßige Außenkontur des Lymphknotens und 2.: Inhomogenes T2-Signal. Dies erschien, unabhängig von den erreichten Ergebnissen, in Anbetracht der angeführten Überlegungen und Studienlage, schlüssig.

#### 4.4.3. Diskussion untersucherabhängiger Einflussfaktoren

In der vorgelegten Studie fand sich eine hohe Anzahl falsch-positiver Ergebnisse: 81 der 232 untersuchten Patientendatensätze wiesen einen falsch-positiven Nodalstatus auf (34%).

Es existieren bzgl. des Lymphknotenstatus in der MRT-Bildgebung allgemeine Pitfalls, die zu falsch-positiven Ergebnissen führen können, in der Bildbeurteilung insbesondere die Evaluation von Satellitenknoten:

Einerseits können *kontinuierlich* im Tumorwachstum verlaufende Satellitenknoten (Tumordeposits) einen befallenen Lymphknoten imitieren und mit ihm verwechselt werden. Um diese Art des Overstagings bzgl. des Lymphknotenstatus zu vermeiden, sollte die Ausbreitung von Tumordeposits in Kontinuität der Tumordinvasion beachtet und in mehreren Schichtebenen beurteilt werden, um diese beiden Entitäten bestmöglich voneinander differenzieren zu können.

Andererseits können auch *diskontinuierlich* vom primären Tumorwachstum entfernt projizierende Satellitenknoten mit einem befallenen Lymphknoten verwechselt werden, insbesondere dann, wenn kein oder kaum residuales Lymphknotengewebe mehr vorhanden ist. Diese Differenzierung stellt bereits in der mikroskopischen Betrachtung der pathologischen Aufarbeitung eine Herausforderung dar, entsprechend mehr noch in der makroskopischen Bildgebung [30].

Es ist durchaus wahrscheinlich, dass diese Tumordeposits in der Vergangenheit selbst in histopathologischen Befunden als Lymphknoten interpretiert wurden [34]. In der aktuellen Version des TNM-Stagings jedoch werden diese Tumordeposits in der Kategorie N1c separat berücksichtigt, irrtümlicherweise in der Kategorie „N“ der Lymphknoten. Analog kann sich auch eine EMVI in der MRT morphologisch als diskontinuierliches knotiges Tumordeposit manifestieren und die Unterscheidung mit einem befallenen Lymphknoten weiter erschweren [30].

Diese möglichen Fehlerquellen, die zu einem falsch-positiven Ergebnis im MRT-Staging führen können, waren im Einzelnen auf Basis der Befundbeurteilungen im Patientenkollektiv aus Neuperlach nicht zu ermitteln. Es ist jedoch davon auszugehen, dass sie sowohl in dieser als auch in anderen Studien zu falsch-positiven Ergebnissen führten, auch wenn die Gründe für falsch-positive Ergebnisse in den einzelnen Studien meist nicht explizit diskutiert werden. In der aktuellen Literatur gibt es vermehrt Hinweise darauf, dass die MRT im N-Status zu einem Overstaging tendiert [33,34,18], wie es sich auch im Patientenkollektiv aus Neuperlach gezeigt hat. Abgesehen von falsch-positiven Ergebnissen aufgrund der Verwechslungsmöglichkeit mit anderen Entitäten eröffnet jedoch auch die Anwendung morphologischer Malignitätskriterien (siehe 2.4.4.5.) Interpretationsspielraum und übt damit untersucherabhängigen Einfluss auf die Ergebnisse des MRT-Nodalstagings aus.

Unabhängig davon fanden sich in den Befundbeurteilungen des Patientenkollektivs aus Neuperlach keine Hinweise darauf, dass Lymphknoten anderer Lymphknotenstationen in die Bewertung einbezogen worden wären, die nach der TNM-Klassifikation nicht als lokoregionäre Lymphknoten gegolten und somit die Vergleichbarkeit aufgrund eines inkongruenten Definitionsbereichs weiter limitiert hätte.

#### **4.4.4. Technische Limitationen**

Es ist bekannt, dass mehr als 50% der tumorös befallenen Lymphknoten bei Rektumkarzinomen einen Durchmesser  $< 5$  mm Durchmesser aufweisen und die Aussagekraft der MRT-Bildgebung bzgl. der Beurteilung kleiner Lymphknoten limitiert ist [17,3]. Welchen Einfluss dies potenziell auf die Ergebnisse des MRT-Stagings nimmt, kann reell nur in prospektiven Studien ermittelt werden. Es gibt leider nur wenige prospektive Studien bzgl. des Lymphknotenstagings der MRT bei Rektumkarzinomen, die sich um eine tatsächliche 1:1 Korrelation der Lymphknoten aus der Bildgebung mit der Histologie bemühen. In der Literatur fand sich hierzu jedoch eine Studie, die sich der Problematik der Mikrometastasen angenommen und dies mit einem 3-Tesla-Scanner untersucht hat.

In dieser Studie konnte gezeigt werden, dass in 341 untersuchten Lymphknoten 120 enthaltene Lymphknoten von  $< 3$  mm Durchmesser in der MRT gar nicht erst detektiert wurden (mehr als ein Drittel), wobei 18 Lymphknoten hiervon (15%) bereits

Mikrometastasen enthielten, was eine durchaus signifikante Teilmenge darstellt [39].

Dies zeigt die technischen Limitationen hinsichtlich des Auflösungsvermögens der MRT auf, die einen Vergleich mit den Ergebnissen der Pathologie erschweren und eine Fehlerquelle bzgl. falsch-negativer Befunde darstellen, unabhängig von den gewählten Malignitätskriterien.

Es ist davon auszugehen, dass auch in dem Patientenkollektiv aus Neuperlach ein gewisser Anteil an befallenen Lymphknoten der pathologischen Aufarbeitung im MRT-Nodalstaging präoperativ nicht detektiert und evaluiert werden konnte und das Ergebnis dementsprechend beeinflusste.

#### **4.4.5. Limitationen des Studiendesigns**

Die größte Schwäche der vorgelegten Studie ist jedoch das retrospektive Studiendesign. Im Gegensatz zu den wenigen prospektiven Studien wurde der Lymphknotenstatus unseres Patientenkollektivs nicht in der 1:1 Korrelation, sondern retrospektiv auf Basis einer Patientenfallbetrachtung untersucht. Dies stellt sicherlich eine Hauptlimitation der Aussagekraft dieser Studie, jedoch auch sämtlicher anderer retrospektiver Studien des MRT-Nodalstagings dar.

Retrospektiv kann schlichtweg nicht ermittelt werden, welcher Lymphknoten im Einzelnen zu einem positiven LK-Status des MRT-Stagings geführt hat und ob dies den wahren Gegebenheiten entsprach. Die Gesamtschau der Lymphknoten sämtlicher retrospektiver Studien ist somit die große Black Box, die retrospektiv nicht ausgeleuchtet werden kann und in der sich sowohl falsch-positiv, als auch falsch-negativ befundene Lymphknoten finden, die jedoch dann in einen ausgewiesenen Lymphknotenstatus münden. Dieser Lymphknotenstatus kann sich dann – man muss zu dem Schluss kommen: Mehr oder weniger zufällig - mit dem Lymphknotenstatus der Pathologie decken.

Aufschlussreich ist in diesem Zusammenhang eine Unterauswertung einer prospektiven Studie [39]:

Entgegen der vergleichsweise geringen Sensitivität von nur 58%, die in der prospektiven 1:1 Korrelation von 341 untersuchten perirektalen Lymphknoten mit der Pathologie erreicht werden konnte, wertete diese Studie zusätzlich eine retrospektive Patientenfallbetrachtung aus (43 Patienten). Unter dieser Betrachtung schnellte

die Sensitivität um 22% auf 80% nach oben und lag hier wieder im oberen Bereich der Genauigkeit, die z.B. von der unter 4.4.1. genannten Metaanalyse angegeben wird.

Dies zeigt, abhängig von der Betrachtungsweise: Prospektiv versus retrospektiv, bereits innerhalb ein und derselben Studie eine deutliche Diskrepanz auf und es ist stark zu bezweifeln, dass eine retrospektive Fallbetrachtung die Realität besser abbilden könnte als eine prospektive Betrachtungsweise.

Nahezu sämtliche Studien zum MRT-Nodalstaging und deren ausgewiesene Ergebnisse sind jedoch retrospektiv erhoben worden und daher bzgl. der Aussagekraft sehr kritisch zu werten.

Dies trifft jedoch nicht nur auf Studien mit vergleichsweise kleinen Fallzahlen, wie der vorgelegten Studie zu: So erreichte eine retrospektive Studie mit einer hohen Patientenanzahl von 2.178 primär operativ therapierten Patienten beispielsweise eine Sensitivität von nur 38%, eine Spezifität von 87%, sowie einen PPW von nur 56% (Patientenkollektiv Neuperlach: PPW 47%). Das Vorhersagen eines positiven Nodalstatus war somit selbst hier nicht weit entfernt von einer Zufallswahrscheinlichkeit, wie sie sich auch beim Patientenkollektiv aus Neuperlach fand. Hinsichtlich des Gesagten kann man sich den Autoren der genannten Studie nur anschließen, die die Genauigkeit der Ergebnisse mit einer Art Münzwurf verglichen, verbunden mit der durchaus ernstzunehmenden These, dass eine zu starke Gewichtung des MRT-Lymphknotenstagings zu einer Überbehandlung von Patienten führen könne [9].

#### **4.4.6. Therapeutische Bedeutung des MRT-Lymphknotenstagings**

Obwohl die Wertigkeit des MRT-Lymphknotenstagings aufgrund der unter 4.4.1.-4.4.5. erörterten Argumente und Unwägbarkeiten als schwach betrachtet werden muss, ist es in der therapeutischen Bedeutung und Konsequenz relevant und stellt z.B. im Rahmen der Empfehlungen der aktuellen S3-Leitlinie, unabhängig vom T-Stadium, eine Indikation zur neoadjuvanten Therapie dar. Eine Ausnahme bilden in der Leitlinie lediglich cT1/T2-Stadien mit fraglichem Lymphknotenbefall in der präoperativen Diagnostik [28]. Damit wird den bereits diskutierten diagnostischen Problemen des klinischen Lymphknotenstagings teilweise Rechnung getragen. In der Hintergrunderklärung für die Leitlinienempfehlung findet man die Feststellung,

dass „(...) die Sensitivität und Spezifität der Beurteilung des Lymphknotenbefalls für alle Staging-Methoden limitiert ist (MRT/endorektaler Ultraschall/Becken-CT)“ und damit das „(...) potentielle ‚Overstaging‘ und die daraus resultierende ‚Überbehandlung‘ von Patienten gewertet (...)“ wird, „(...) bei denen fälschlicherweise ein lymphknotenpositiver Tumor (cN+) diagnostiziert wurde“. Daher „(...) wird bei cT1/2-Tumoren mit bildgebend fraglichen cN+ (formal also UICC-Stadium III) die primäre Operation als sinnvolle Option erachtet.“ [28]

Abweichend von der Leitlinienempfehlung gibt es jedoch klinische Studien, die sich auf das MRT-Staging des Rektumkarzinoms stützen und einen anderen Weg beschreiben. In diesem Zusammenhang sei beispielsweise die MERCURY-Studie erwähnt, die im Gegensatz zur aktuellen S3-Leitlinie eine restriktivere Auswahl von Patienten zur neoadjuvanten Therapie proklamiert: In dieser groß angelegten, multizentrischen prospektiven Studie zum MRT-gestützten präoperativem Staging des Rektumkarzinoms konnte an einer Patientengruppe von 122 Patienten (33%) der präoperativen Klassifizierung „gute Prognose“ (gemeint ist ein MRT-basiertes Staging mit negativem CRM-Status bei Tumorabstand > 1 mm zur mesorektalen Faszie, ohne Hinweis auf eine EMVI und bis maximal Stadium T3b, also einer Invasionstiefe in das perirektale Fettgewebe <=5 mm) gezeigt werden, dass diese Patientengruppe nach primärer Operation *unabhängig* vom MRT-Nodalstatus im Follow-Up ein gutes onkologisches Outcome hatte.

Die Gesamtüberlebensrate und tumorfreie Überlebensrate im 5-Jahres-Follow-Up betrug in der MERCURY-Studie 68% bzw. 85% bei einer Lokalrezidivrate von nur 3% [51].

Diese Ergebnisse zeigen, dass durch die Kombination mehrerer Kriterien gute therapeutische Ergebnisse für den Patienten erzielt werden können und die Gefahr der Übertherapie, die aus dem reinen Nodalstaging resultiert, vermieden werden kann. Diese Nodalstatus-unabhängigen Kriterien hätten analog auf 99 Patienten des untersuchten Patientenkollektivs aus Neuperlach zugetroffen (42% der untersuchten Patienten).

Zusammenfassend ist festzustellen, dass die Ergebnisse des MRT-Nodalstagings des Patientenkollektivs aus Neuperlach im Vergleich mit der Pathologie - wie in allen retrospektiven Studien - mit großen Ungenauigkeiten behaftet war und zu einem Overstaging tendierte. Der PPW lag sogar unter der Zufallswahrscheinlichkeit von 50% zur Vorhersage eines positiven Nodalstatus. Die Ergebnisse sind zu ungenau,

um davon abhängig ein Therapieschema für den Patienten festzulegen, da es im MRT-Staging geeignetere und auch mit der Pathologie besser vergleichbare Kriterien zur Selektion eines Therapieschemas mit dem bestmöglichen Outcome für den Patienten gibt. Bereits ein einziger falsch-positiv beurteilter Lymphknoten kann weitreichende Konsequenzen nach sich ziehen. Der Nodalstatus sollte in der aktuellen Routinediagnostik als unabhängiger Faktor für eine Therapieentscheidung daher außer Acht gelassen werden, um das große Risiko einer Übertherapie zu vermeiden. Die Spätfolgen einer präoperativen Radiochemotherapie können erheblich sein und umfassen u.a. sexuelle Dysfunktion, Inkontinenz, Infertilität, Entstehung von Zweittumoren sowie kardiovaskuläre Erkrankungen [43].

#### **4.5. EMVI**

Die Ergebnisse der MRT und der Pathologie konnten in der Beurteilung einer EMVI in dieser Studie nicht sinnvoll statistisch korreliert werden. Einerseits war die Diskrepanz der Befunde hoch, da nur in 3 übereinstimmenden Datensätzen zumindest der Hinweis auf das Vorliegen einer EMVI auf Seiten der MRT mit dem Vorliegen venöser Tumorinfiltrate der Befundbeschreibungen der Pathologie korrelierte. Aus den Befundbeschreibungen der Pathologie konnte andererseits auch nicht sicher entnommen werden, ob venöse Tumorinfiltrationen extramural oder intramural vorlagen.

Für eine statistische Korrelation mit der Pathologie war das vorliegende Datenmaterial dieser retrospektiven Studie weiterhin schon allein aufgrund der Tatsache nicht geeignet, dass keine Übereinstimmung der Kriterien einer EMVI zwischen MRT-Befunden und Befundberichten der Pathologie sichergestellt werden konnte. In den pathologischen Befundberichten fand sich keine vergleichbare Kategorie pEMVI. Diese findet man in der Literatur lediglich in (überwiegend prospektiven) Studien, die explizit eine Korrelation der Pathologie mit der MRT in der Evaluation einer EMVI als Studienziel hatten und die Kriterien im Vorfeld zwischen Radiologie und Pathologie abgesprochen waren, was in der vorgelegten retrospektiven Querschnittsstudie der MRT-Diagnostik nicht der Fall war.

Eine Übersichtsarbeit, die die Ergebnisse von 14 histopathologischen Studien zum Rektumkarzinom zusammenfasste, fand eine gepoolte Prävalenz einer EMVI bei

26% der Patienten in der pathologischen Aufarbeitung [15]. Im Patientenkollektiv aus Neuperlach fanden sich auf Seiten der Pathologie in 16 Datensätzen (7% des Studienkollektivs) venöse Tumordinfiltrate, der Hinweis auf eine EMVI war auf Seiten der MRT in 28 Datensätzen (12% des Studienkollektivs) zu finden. Die vergleichsweise geringe Prävalenz sowohl auf Seiten der Pathologie als auch auf Seiten der MRT liegt wohl darin begründet, dass die meisten T3- und T4-Karzinome neoadjuvant therapiert wurden und somit ein Ausschlusskriterium für diese Studie bildeten. Eine EMVI ist jedoch nur für T3- und T4-Karzinome definiert [47].

In den Befundbeurteilungen der MRT fand sich in der Evaluation einer EMVI in 36% der Fälle eine unsichere Diagnose („V.a. EMVI bzw. „fragliche EMVI). Dies stellte eine hohe Teilmenge in der Evaluation einer EMVI dieser Studie dar und es ist nicht sicher, wie dies klinisch gewertet werden kann. Auf der Suche nach Verbesserungen für die Standarddiagnostik in der Vergleichbarkeit mit histopathologischen Ergebnissen fand sich in der Literatur eine Arbeit, die eine strukturierte Befunderhebung mit guten Ergebnissen ermöglicht und auch die Befundkonsistenz verbessern kann. Smith et al. (2007) erarbeiteten einen EMVI-Score zur strukturierten Befunderhebung der MRT. Hierbei beschreibt ein EMVI-Score von „0“ den Fall, dass keine noduläre Tumorausbreitung durch die Muskularis propria vorliegt und keine direkt an die Tumorpenetration angrenzenden Gefäße nachweisbar sind. Der Score von „1“ beschreibt, dass minimale extramurale spikuläre bzw. noduläre Formationen vorliegen, jedoch nicht in direkter Umgebung vaskulärer Strukturen. Der Score von „2“ beschreibt, dass extramurale spikuläre Formationen in direkter Umgebung vaskulärer Strukturen mit regelrechtem Kaliber und ohne eindeutiges intraluminales Tumorsignal vorliegen. Der Score von „3“ beschreibt, dass in einem Gefäßlumen mit nur minimal erweiterter Gefäßkontur bzw. Gefäßkaliber intermediäre Signalintensität vorliegt. Der Score von „4“ beschreibt den Fall, dass Gefäße mit offensichtlich irregulärer Kontur oder nodulärer Auftreibung mit eindeutigem Tumorsignal vorliegen [47]:

Nach dieser Einteilung kann ein Score von 0–2 als EMVI-negativ, ein Punkte-Score von 3-4 als EMVI-positiv definiert werden.

Unter dieser Einteilung konnte die Studie in ihrem Kollektiv an 94 primär operativ therapierten Patienten im Vergleich mit der Pathologie eine Sensitivität und Spezifität von 62% bzw. 88%, einen PPW von 67% sowie einen NPW von 86% erreichen

[47]. Insbesondere die hohe Spezifität und der NPW sprechen für die Verwendung dieses Scoring-Systems, das sich in der Literatur vielfach durchgesetzt hat.

Es ist bekannt, dass im Vergleich mit der Pathologie die Sensitivität der MRT tendenziell nicht sehr hoch ausfällt, oft auch geringer als die Spezifität. In der MRT-Diagnostik können aufgrund der beschränkten Auflösung sehr kleine venöse Gefäße, die jedoch mikroskopisch sichtbar sind, schlichtweg nicht immer sicher detektiert werden [53]. Es konnte jedoch auch gezeigt werden, dass die MRT in einigen Fällen eine EMVI detektiert konnte, die der pathologischen Standarddiagnostik zunächst entgangen ist [45,53]. Daher ist die Pathologie in dieser Kategorie nicht als optimaler Goldstandard zu sehen und die erreichbaren Genauigkeiten der MRT im Vergleich mit der Pathologie relativ zu werten, die in der Literatur auch sehr stark variieren. Ein Vorteil der MRT ist, dass eine EMVI in vivo entlang der gesamten Tumorausdehnung evaluiert werden kann.

Unbestritten ist die prognostische und mögliche therapeutische Relevanz einer EMVI. Aus der Histopathologie ist bekannt, dass die EMVI beim kolorektalen Karzinom ein schlechter prognostischer Faktor sowohl für Überlebens- und Rezidivraten als auch für das Auftreten von Fernmetastasen ist [50].

Für die prognostische Bedeutung der Detektion einer EMVI in der MRT hinsichtlich metachroner Metastasen konnte eine Studie an 53 Patienten der Verlaufsbeobachtung zeigen, dass das relative Risiko, eine Fernmetastasierung nach einem Jahr bei positivem EMVI-Befund der MRT (nach dem EMVI-Score von Smith et al. 2007) im Gegensatz zu EMVI-negativen Patienten knapp 4-fach erhöht war (6,7% vs. 24,5%). Die Autoren schlossen aus dieser Studie, dass die Evaluation einer EMVI in der MRT im Primärstaging eine wichtige Hilfestellung in der Therapieentscheidung leisten könne, um Patienten zu identifizieren, die von einem intensivierten Therapieschema wie einer neoadjuvanten Therapie profitieren könnten [13].

Smith et al. (2007) konnten an ihrem Kollektiv von 94 Patienten zeigen, dass die Rezidivrate MRT-befundlich EMVI-positiver Patienten mit einem EMVI-Score von 3-4 im 3-Jahres-Follow-Up mehr als doppelt so hoch war, als bei EMVI-negativen Patienten mit einem EMVI-Score von 0-2 (74% vs. 35%) und dass die Ergebnisse sehr gut mit dem histologischen EMVI-Status korrelierten [47]. Die Mercury II-Studie

konnte zeigen, dass die EMVI im MRT-Staging ein unabhängiger Risikofaktor für einen postoperativ positiven CRM darstellt [6].

Eine weitere Studie, die den EMVI-Score nach Smith et al. (2007) verwendete, konnte eine Korrelation an 86 primär operativ therapierten Patienten zwischen MR-befundlich positiver EMVI und dem histologischen Grading, Lymphknotenstatus sowie der VEGF-Expression in T3-Karzinomen nachweisen. Die Autoren schlugen den positiven EMVI-Score der MRT als Biomarker der Bildgebung für die Aggressivität von T3-Karzinomen vor, der hilfreich sein könnte, im multidisziplinären Therapieansatz die Entscheidung bzgl. einer neoadjuvanten Therapie zu treffen. Auch konnte die Studie zeigen, dass die Ergebnisse zwischen zwei Befundern mit dem EMVI-Score gut reproduzierbar waren und sich eine EMVI, gewertet als LV1-Stadium der Pathologie (lymphatische und/oder venöse Gefäßinvasion Grad I) mit einer Sensitivität von 58,33% sowie einer Spezifität von 92% beurteilen ließ (allerdings entsprechen die Kriterien nicht exakt der Definition einer histopathologischen EMVI Anm. d. Autors) [55].

In der aktuellen S3-Leitlinie „Kolorektales Karzinom“ stellt in der konsensbasierten Therapieempfehlung für T3-Tumore des mittleren Drittels der bildmorphologische Ausschluss einer EMVI eine Bedingung für die Empfehlung einer primär operativen Therapie dar [28]. Eine Befunderhebung, mit der nachweislich nach dem aktuellen Stand der Literatur eine hohe Spezifität im Vergleich mit der Pathologie erreicht werden kann, ist daher besonders wichtig.

Zusammenfassend ist aufgrund der Relevanz der EMVI in der aktuellen Diagnostik und Therapie eine möglichst strukturierte und reproduzierbare Befunderhebung der MRT-Diagnostik essenziell. Der EMVI-Score nach Smith et al. (2007) scheint aufgrund der zitierten Studienergebnisse und Leitlinienempfehlungen daher sehr gut für die Standarddiagnostik anwendbar.

#### **4.6. Schlussfolgerungen aus der Diskussion**

- Bzgl. des *T-Stagings* konnte für die Stadien T1-T4 eine Sensitivität von 42% 63%, 68% und 54% erreicht werden bei einer Spezifität von 96%, 75%, 74% und 96%. Die Sensitivität der T1-, T3- und T4-Stadien sowie die Spezifität der T2-Stadien lagen hierbei deutlich unter dem Niveau vergleichbarer Studien. Es könnte in einer zukünftigen Studie reevaluiert werden, ob mit dem

neuen Scanner, der seit 04/2019 in Betrieb ist, Verbesserungen, insbesondere in der Sensitivität im T-Staging erreicht werden konnten.

- In der Konfidenzintervallanalyse der Reliabilität bzgl. der Übereinstimmung des T-Stagings und CRM-Status konnte keine Abhängigkeit der Ergebnisse von der Tumordicke festgestellt werden. Somit ergab sich in dieser Auswertung kein Potenzial für eine Verbesserung der Standarddiagnostik.
- In der Beurteilung der Invasionstiefe in das perirektale Fettgewebe der T3-Stadien konnte in Bezug auf eine höhergradige Infiltration  $\geq 5$  mm eine Sensitivität von 52%, eine Spezifität 78%, ein PPW 61% sowie ein NPW von 71% erreicht werden, die therapeutisch relevante Spezifität war hierbei als moderat zu werten, die Sensitivität als zu gering. In dieser Kategorie wäre eine prospektive Abweichungsanalyse anhand von kleinen Fallzahlen sinnvoll, da die Abweichungen hinsichtlich der potenziell erreichbaren Genauigkeiten der MRT in dieser Studie zu hoch erschienen: 68% der Fälle (46 der 68 untersuchten Datensätze) wichen um  $\geq 2$  mm von den Messungen der Pathologie ab. Ursächlich könnten u.a. Partialvolumeneffekte aufgrund einer zu stark unterschiedlichen Kippung zwischen den perpendicular auf die Tumorchse eingekippten MRT-Sequenzen und den histologischen Schnitten eine Rolle gespielt haben, eine diesbzgl. interdisziplinäre Klärung mit der Pathologie könnte aufschlussreich sein.
- Die Einteilungsmöglichkeit von T3-Substadien in lediglich zwei Klassen (T3a/b und T3c/d) anstatt wie bisher vier Klassen (T3a, T3b, T3c, T3d), könnte die Befundkonsistenz weiter erhöhen. Nach aktueller Studienlage und Leitlinienempfehlungen ist ohnehin lediglich eine Invasionstiefe  $>5$  mm prognostisch und ggf. therapeutisch relevant, entsprechend der Kategorie T3c/d, weshalb diese Einteilung auch in der Literatur vielfach vorgeschlagen wird.
- Im *Nodalstaging* konnte eine Sensitivität von 85%, eine Spezifität von 45%, ein NPW von 84% sowie ein PPW von 47% erreicht werden. In der vorgelegten Studie tendierte die MRT-Diagnostik zum Overstaging, auch in der Literatur fand sich eine generelle Tendenz zum Overstaging. Ungeachtet der Ergebnisse war der Lymphknotenstatus in diesen wie in vergleichbaren retrospektiven Studien im präoperativen MRT-Staging als unsicher zu werten. Das retrospektive Studiendesign stellt hierbei die Hauptlimitation der Ergebnisse dar: Retrospektiv kann schlichtweg nicht ermittelt werden, welcher

Lymphknoten im Einzelnen zu einem positiven LK-Status des MRT-Stagings geführt hat und ob dies den wahren Gegebenheiten entsprach. Die Gesamtschau der Lymphknoten sämtlicher retrospektiver Studien ist somit die große Black Box, die retrospektiv nicht ausgeleuchtet werden kann und in der sich sowohl falsch-positiv, als auch falsch-negativ befundene Lymphknoten finden, die jedoch dann in einen ausgewiesenen Lymphknotenstatus münden. Dieser Lymphknotenstatus kann sich dann – man muss zu dem Schluss kommen: Mehr oder weniger zufällig - mit dem Lymphknotenstatus der Pathologie decken, der als Goldstandard gilt. Die therapeutische Konsequenz hinsichtlich einer möglichen Übertherapie bei gleichzeitig hoher diagnostischer Unsicherheit ist jedoch so relevant, dass ein positiver Lymphknotenstatus der MRT bei Kontakt oder Nähe von suspekten Lymphknoten zur mesorektalen Faszie im präoperativen Staging zukünftig zumindest nicht mehr als Kriterium für CRM-Positivität herangezogen werden sollte.

- Im *CRM-Status* konnte eine Sensitivität von 78%, eine Spezifität von 84%, ein PPW von 17% sowie ein NPW von 99% erreicht werden. Die Spezifität sowie der hohe NPW, die für die Rechtfertigung rein operativer Verfahren, insbesondere der T3-Stadien relevant sind, fielen, auch im Vergleich mit der Literatur, sehr hoch aus.
- Es sollte reevaluiert werden, ob MR-befundliche T1- und T2-Stadien mit Darmwandnähe zur mesorektalen Faszie weiterhin als Kriterium für CRM-Positivität verwendet werden sollten, wie es bisher gehandhabt wurde. Dies führte in diesem Kollektiv ausschließlich zu (insgesamt 6) falsch-positiven Befunden (17% aller falsch-positiver Befunde im CRM-Status).
- In der Reliabilitätsanalyse der Kappa-Auswertung des CRM-Status fand sich aufgrund eines auffälligen Datenausreißers nach Reevaluation ein diagnostischer Pitfall, der zu systemisch relevanten Fehlbefundungen i.S. eines falsch-positiven CRM-Status führte und auch mit einem Understaging in der T4-Kategorie assoziiert war. Aufgrund dieser Analyse wurde eine Verbesserungsmöglichkeit für die Befundvorlage im RIS gefunden: Künftig sollte eine standardisierte Evaluation der Tumorphöhe zur peritonealen Umschlagfalte erfolgen, am besten in Form einer geänderten Befundvorlage. Somit könnten strukturelle Fehlbefundungen aufgrund der Verwechslungsmöglichkeit der peritonealen Umschlagfalte mit der mesorektalen Faszie, wie sie in dieser

Studie in 8,5% aller falsch-positiven Befunde im CRM-Status vorkamen, bestmöglich vermieden werden.

- Bzgl. der Evaluation einer EMVI könnte für die zukünftige Standarddiagnostik das EMVI-Scoring nach Smith et al. (2007) in die Befundvorlage im RIS übernommen werden, da dies zu einer verbesserten Befundkonsistenz und sichereren Diagnosestellung führen kann, wie in zahlreichen Studien nachgewiesen werden konnte. Eine statistische Auswertung in der Kategorie EMVI im Vergleich mit der Pathologie war aufgrund der Datenlage nicht möglich.

## **5. Zusammenfassung**

Im Klinikum Neuperlach nimmt das Primärstaging des Rektumkarzinoms mittels Magnetresonanztomographie (MRT) eine zentrale Rolle für die Therapieentscheidung im interdisziplinären Ansatz ein. Ziel der Studie war es, einen Querschnitt der Standarddiagnostik in der radiologischen Abteilung an primär operativ therapierten Patienten im Zeitraum 01.01.2008-31.12.2015 als Qualitätssicherungsmaßnahme abzubilden. Hierzu wurde das MRT-Staging mit den postoperativen Befunden der Pathologie als Goldstandard retrospektiv verglichen. Ausgewertet wurden das T-Staging, die Evaluation des Tumorwachstums in Bezug auf den zirkumferenziellen Resektionsrand (CRM), das Lymphknotenstaging, die Infiltrationstiefe in das perirektale Fettgewebe der T3-Stadien sowie die extramurale venöse Tumordinvasion (EMVI). Hierbei wurde mittels statistischer Auswertung und Vergleich mit Studienergebnissen der Literatur überprüft, welche diagnostische Performance erreicht werden konnte und ob es Verbesserungspotenzial für die Standarddiagnostik gab. Im T-Staging und CRM-Status wurde zusätzlich untersucht, ob und inwiefern sich die Tumordicke auf die Reliabilität ausgewirkt hat.

Im *T-Stadienvergleich* erfolgte die Stadieneinteilung entsprechend der maximalen transmuralen Tumorausdehnung analog zur Pathologie, gemessen als intermediäres Tumorsignal in den T2-gewichteten Sequenzen. Für die Stadien T1-T4 wurde im Gesamtkollektiv eine Sensitivität von 42%, 63%, 68% und 54% erreicht, die Spezifität betrug 96%, 75%, 74% und 96%. Die Sensitivität der T1-, T3- und T4-Stadien sowie die Spezifität der T2-Stadien lag unter der Performance ähnlicher Studien der Literatur, ansonsten wurden vergleichbare Ergebnisse erreicht.

In der *Feinauswertung der T3-Stadien* wurde hinsichtlich einer relevanten höhergradigen Invasionstiefe in das perirektale Fettgewebe > 5 mm eine Sensitivität von 52%, eine Spezifität von 78%, ein positiver Vorhersagewert (PPW) von 61% sowie ein negativer Vorhersagewert (NPW) von 71% erreicht. Auffallend waren höhere Messabweichungen zur Pathologie, als hinsichtlich der potenziell erreichbaren Genauigkeiten der MRT zu erwarten gewesen wären. Hierbei könnten unterschiedliche Partialvolumeneffekte bei den Messungen der MRT und Pathologie die entscheidende Rolle gespielt haben, was in einer künftigen Studie mit prospektivem Studiendesign näher untersucht werden könnte. Die vorgeschlagene Einteilung der T3-

Substadien (entsprechend der Infiltrationstiefe  $\leq$  bzw.  $>$  5 mm) in lediglich zwei Klassen kann die Befundkonsistenz weiter verbessern.

Im *CRM-Status* konnte im Gesamtkollektiv eine Sensitivität von 78%, eine Spezifität von 84%, ein PPW von 17% sowie ein NPW von 99% erreicht werden. Als Malignitätskriterium für CRM-Positivität wurde der Abstand einer Tumorformation (inklusive suspekter Lymphknoten)  $\leq$  1 mm zur mesorektalen Faszie definiert. Es wurden sehr gute Ergebnisse in der Spezifität sowie dem NPW erreicht, was für die Rechtfertigung rein operativer Therapieverfahren von speziellen T3-Stadien relevant ist. Auch wurde eine Verbesserungsmöglichkeit für die hausinterne standardisierte Befundvorlage für das MRT-Staging von Rektumkarzinomen gefunden.

Im T-Stadienvergleich und im CRM-Status konnte kein Einfluss einer Metrik nach Tumorrhöhe auf die Reliabilität festgestellt werden.

Im *Lymphknotenstaging* konnte eine Sensitivität von 85%, eine Spezifität von 45%, ein NPW von 84% sowie ein PPW von 47% erreicht werden. In dieser und in vergleichbaren Studien fand sich eine Tendenz zum Overstaging. Die Hauptlimitation der Aussagekraft der Ergebnisse war hierbei im retrospektiven Studiendesign zu sehen. Die Malignitätskriterien eines suspekten Lymphknotens in der MRT wurden anhand einer irregulären Außenkontur und/oder eines gemischten, inhomogenen T2-Signals definiert. Methodisch bedingt ist der präoperative N-Status der MRT, obwohl therapeutisch relevant, weiterhin als unsicher zu betrachten. Suspekte Lymphknoten der MRT sollten aufgrund des Risikos einer Übertherapie deshalb künftig nicht mehr so stark gewichtet werden, als dass sie den CRM-Status beeinflussen können.

Die Evaluation der *EMVI* konnte auf Basis des vorliegenden Datenmaterials nicht statistisch ausgewertet werden. Die EMVI in der MRT wurde anhand einer geschlängelten Ausdehnung von Tumorsignal in einer vaskulären Struktur außerhalb der Muskularis propria definiert. In der EMVI-Evaluation konnte eine Verbesserungsmöglichkeit gefunden werden, die erweiterte Bildkriterien einbezieht und zu einer sichereren Diagnose führen kann.

Zusammenfassend konnten in der vorgelegten Studie wichtige Erkenntnisse und Verbesserungsmöglichkeiten für die Standarddiagnostik des Rektumkarzinoms mittels MRT für die radiologische Abteilung des Klinikum Neuperlach gefunden werden.

## **6. Literaturverzeichnis**

1. Akagi Y, Shirouzu K, Fujita S, Ueno H, Takii Y, Komori K, Ito M, Sugihara K: Predicting oncologic outcomes by stratifying mesorectal extension in patients with pT3 rectal cancer: a Japanese multi-institutional study. *International Journal of Cancer* 131 :1220-1227 (2012)
2. Al-Sukhni E, Milot L, Fruitman M, Beyene J, Victor JC, Schmocker S, Brown G, McLeod R, Kennedy E: Diagnostic Accuracy of MRI for Assessment of T Category, Lymph Node Metastases, and Circumferential Resection Margin Involvement in Patients with Rectal Cancer: A Systematic Review and Meta-analysis. *Annals of Surgical Oncology* 19: 2212-2223 (2012)
3. Andreola S, Leo E, Belli F, Bufalino R, Tomasic G, Lavarino C, Baldini MT, Meroni E: Manual dissection of adenocarcinoma of the lower third of the rectum specimens for detection of lymph node metastases smaller than 5 mm. *Cancer* 77: 607-612 (1996)
4. Asbach P, Taupitz M: Primäres MRT-Staging des Rektumkarzinoms. in Kreis ME, Asbach P (Hrsg). *MRT-basierte Chirurgie des Rektumkarzinoms*. Springer, Berlin, Heidelberg, New York, S. 55 (2020)
5. Balyasnikova S, Brown G: Optimal Imaging Strategies for Rectal Cancer Staging and Ongoing Management. *Current Treatment Options in Oncology*. 17: 32 (2016)
6. Battersby NJ, How P, Moran B, Stelzner S, West N P, Branagan G, Strassburg J, Quirke P, Tekkis P, Pedersen BG, Gudgeon M, Heald B, Brown G, MERCURY II Study Group (2016): Prospective Validation of a Low Rectal Cancer Magnetic Resonance Imaging Staging System and Development of a Local Recurrence Risk Stratification Model: The MERCURY II Study. *Annals of surgery* 263: 751-760 (2016)
7. Branagan G; Chave H; Fuller C; McGee S; Finnis D. Can magnetic resonance imaging predict circumferential margins and TNM stage in rectal cancer? *Diseases of the Colon & Rectum* 47: 1317-1322 (2004)
8. Brandhuber M: Anorektale Anatomie in der MRT. in Kreis ME, Asbach P (Hrsg). *MRT-basierte Chirurgie des Rektumkarzinoms*. Springer, Berlin, Heidelberg, New York, S.14-25 (2020)

9. Brouwer NP, Stijns RC, Valery L, Nagtegaal ID, Beets-Tan RG, Fütterer J, Verhoeven RH, de Wilt JH: Clinical lymph node staging by imaging in colorectal cancer: A flip of the coin? *Journal of Clinical Oncology* 35: 15\_suppl, e15160-e15160 (2017)
10. Brown G: Diagnostic accuracy of preoperative magnetic resonance imaging in predicting curative resection of rectal cancer: Prospective observational study. *British Medical Journal*: 333: 779-782 (2006)
11. Brown G, Radcliffe AG, Newcombe RG, Dallimore NS, Bourne MW, Williams G T: Preoperative assessment of prognostic factors in rectal cancer using high-resolution magnetic resonance imaging. *British Journal of Surgery (BJS)* 90: 355-364 (2003)
12. Brown G, Richards CJ, Bourne MW, Newcombe RG, Radcliffe AG, Dallimore NS, Williams GT: Morphologic predictors of lymph node status in rectal cancer with use of high-spatial-resolution MR imaging with histopathologic comparison. *Radiology*. 227: 371-377 (2003)
13. Bugg WG, Andreou AK, Biswas D, Toms AP, Williams SM: The prognostic significance of MRI-detected extramural venous invasion in rectal carcinoma. *Clinical Radiology* 69: 619-623 (2014)
14. Burton S, Brown G, Daniels I, Norman A, Swift I, Abulafi M, Wotherspoon A, Tait, D: MRI identified prognostic features of tumors in distal sigmoid, recto-sigmoid, and upper rectum: treatment with radiotherapy and chemotherapy. *International Journal of Radiation Oncology – Biology – Physics* 65: 445-451 (2006)
15. Chand M, Siddiqui MR, Swift I, Brown G: Systematic review of prognostic importance of extramural venous invasion in rectal cancer. *World journal of gastroenterology* 22: 1721-1726 (2016)
16. Daniels, I. Extramural depth of tumor invasion at thin-section MR in patients with rectal cancer: results of the MERCURY study. *Radiology* 243:132-139 (2007)
17. Dworak O: Number and size of lymph nodes and node metastases in rectal carcinomas. *Surgical Endoscopy* 3: 96-99 (1989)
18. Dworak O: Morphology of lymph nodes in the resected rectum of patients with rectal carcinoma *Pathology – Research and Practice* 187: 1020-1024 (1991)

19. Ferri M, Laghi A, Mingazzini P, Iafrate F, Meli L, Ricci F, Passariello R, Ziparo V: Pre-operative assessment of extramural invasion and sphincteral involvement in rectal cancer by magnetic resonance imaging with phased-array coil. *Colorectal Disease* 7: 387-393 (2005)
20. Glynne-Jones R, Wyrwicz L, Tiret E, Brown G, Rödel C, Cervantes A, Arnold D: Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Surgical Oncology* 28: 22-40 (2017)
21. Gosens MJ, Klaassen RA, Tan-Go I, Rutten HJ, Martijn H, van den Brule AJ, Nieuwenhuijzen GA, van Krieken JH, Nagtegaal I: Circumferential Margin Involvement Is the Crucial Prognostic Factor after Multimodality Treatment in Patients with Locally Advanced Rectal Carcinoma. *Clinical Cancer Research* 13: 6617-662 (2007)
22. Haanappel A, Kroon HM, Schaap DP, Bedrikovetski S, Dudi-Venkata NN, Lee HX, Thomas ML, Liu J, Van Der Valk, MJ, Rutten HJ, Beets GL, Kusters M, Sammour T: Lateral Lymph Node Metastases in Locally Advanced Low Rectal Cancers May Not Be Treated Effectively With Neoadjuvant (Chemo) Radiotherapy Only. *Frontiers in oncology* 9: 1355 (2019)
23. [https://www.amboss.com/de/wissen/Epidemiologie\\_und\\_Wahrscheinlichkeiten](https://www.amboss.com/de/wissen/Epidemiologie_und_Wahrscheinlichkeiten) (30.11.2019)
24. [https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Darmkrebs/darmkrebs\\_node.html](https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Darmkrebs/darmkrebs_node.html) (25.05.2020)
25. <https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/krebsarten/darmkrebs/therapie.html> (12.08.2021)
26. <https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/krebsarten/weitere-krebsarten/endedarmkrebs-rektumkarzinom/definition-haeufigkeit.html> (12.08.2021)
27. <https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/krebsarten/weitere-krebsarten/endedarmkrebs-rektumkarzinom/tumorstadien.html> (19.11.2019)
28. <https://www.leitlinienprogrammonkologie.de/leitlinien/kolorektales-karzinom/> Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF): S3-Leitlinie Kolorektales Karzinom, Langversion 2.1, 2019, AWMF Registrierungsnummer: 021/007OL (30.10.2019)

29. Jhaveri KS, Hosseini-Nik H, MRI of Rectal Cancer: An Overview and Update on Recent Advances 205: 42-55 (2015)
30. Kim JH, Beets GL, Kim MJ, Kessels AG, Beets-Tan RG: High-resolution MR imaging for nodal staging in rectal cancer: are there any criteria in addition to the size? *European Journal of Radiology* 52: 78–83 (2004)
31. Kim Y, Cha SW, Pyo J, Kim NK, Min BS, Kim MJ, Kim H: Factors Related to Preoperative Assessment of the Circumferential Resection Margin and the Extent of Mesorectal Invasion by Magnetic Resonance Imaging in Rectal Cancer: A Prospective Comparison Study. *World Journal of Surgery* volume 33: 1952-1960 (2009)
32. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174 (1977)
33. Li XT, Sun YS, Tang L, Cao K, Zhang XY: Evaluating local lymph node metastasis with magnetic resonance imaging, endoluminal ultrasound and computed tomography in rectal cancer: a meta-analysis. *Colorectal Disease* 17: 129-135 (2015)
34. Lord A, D'Souza N, Shaw A, Nigel D, Brown G: The Current Status of Nodal Staging in Rectal Cancer. *Current Colorectal Cancer* 15: 143-148 (2019)
35. Matsuoka H, Nakamura A, Sugiyama M, Hachiya J, Atomi Y, Masaki T: MRI diagnosis of mesorectal lymph node metastasis in patients with rectal carcinoma. What is the optimal criterion? *Anticancer research* 24: 97-101 (2004)
36. McHugh ML: Interrater reliability: the kappa statistic. *Biochemia medica* 22: 276-82 (2012)
37. Merkel S, Mansmann U, Siassi M, Papadopoulos T, Hohenberger W, Hermanek P: The prognostic inhomogeneity in pT3 rectal carcinomas. *International Journal of Colorectal Disease* 30: 298-304 (2001)
38. Ma X, Li X, Xu L, Shi D, Tong T, Huang D, Ding Y, Cai S, Peng J: Characteristics and Prognostic Significance of Preoperative Magnetic Resonance Imaging-Assessed Circumferential Margin in Rectal Cancer. *Gastroenterology research and practice*, 2015, 410150 (2015)
39. Park JS, Jang YJ, Choi, GS, Park SY, Kim HJ, Kang H, Cho SH: Accuracy of Preoperative MRI in Predicting Pathology Stage in Rectal Cancers: Node-for-Node Matched Histopathology Validation of MRI Features. *Diseases of the Colon & Rectum* 57: 32-38 (2014)

40. Pham TT, Stait-Gardner T, Lee CS, Barton M, Graham PL, Liney G, Wong K, Price WS: Correlation of ultra-high field MRI with histopathology for evaluation of rectal cancer heterogeneity. *Scientific Reports* 9: 9311 (2019)
41. Pratik T, Guo W, Yang C, Bimal R, Zeng M: Clinical Feasibility Assessment of T3 Sub-Stage in Rectal Cancer Using MRI. *Iranian Journal of Radiology* 15: 16801 (2018)
42. Raychaudhuri B, Cahill D: Pelvic fasciae in urology. *Annals of the Royal College of Surgeons of England* 90: 633-637 (2008)
43. Ruppert R, Junginger T, Ptok H, Strassburg J, Maurer CA, Brosi P, Sauer J, Baral J, Kreis M, Wollschlaeger D, Hermanek P, Merkel S, Chichester UK: Oncological outcome after MRI-based selection for neoadjuvant chemoradiotherapy in the OCUM Rectal Cancer Trial. *British Journal of Surgery* 105: 1519-1529 (2018)
44. Shihab OC, Moran BJ, Heald RJ, Quirke P, Brown G: MRI staging of low rectal cancer. *European Radiology* 19: 643-650 (2009)
45. Shihab OC, Taylor F, Salerno G, Heald RJ, Quirke P, Moran BJ, Brown G: MRI predictive factors for long-term outcomes of low rectal tumours. *Annals of Surgical Oncology* 18: 3278-3284 (2011)
46. Shin R, Jeong SY, Yoo HY, Park KJ, Heo SC, Kang GH, Kim WH, Park JG: Depth of mesorectal extension has prognostic significance in patients with T3 rectal cancer. *Diseases of the Colon & Rectum* 55: 1220-1228 (2012)
47. Smith NJ, Barbachano Y, Norman AR, Swift RI, Abulafi AM, Brown G: Prognostic significance of magnetic resonance imaging-detected extramural vascular invasion in rectal cancer. *British Journal of Surgery* 95: 229-236 (2007)
48. Steel MC, Woods R, Mackay JM, Chen F: Extent of mesorectal invasion is a prognostic indicator in T3 rectal carcinoma. *ANZ Journal of Surgery* 72 :483-487 (2002)
49. Strassburg J, Lewin A, Ludwig K, Kilian L, Linke J, Loy V, Knuth P, Püttcher O, Ruehl U, Stöckmann F, Hackenthal M, Hopfenmüller W, Huppertz A: Optimised surgery (so-called TME surgery) and high-resolution MRI in the planning of treatment of rectal carcinom. *Archives of Surgery* 392: 179-188 (2007)
50. Talbot IC, Ritchie S, Leighton MH, Hughes, AO, Bussey R, Henry J, Morson BC: Spread of rectal cancer within veins: Histologic features and clinical significance. *The American Journal of Surgery* 141: 15-17 (1981)

51. Taylor FG, Quirke P, Heald RJ, Moran BJ, Blomqvist L, Swift IR, Sebag-Montefiore D, Tekkis P, Brown G: Preoperative High-resolution Magnetic Resonance Imaging Can Identify Good Prognosis Stage I, II, and III Rectal Cancer Best Managed by Surgery Alone: A Prospective, Multicenter, European Study. *Annals of Surgery* 253: 711-719 (2011)
52. Taylor FG, Quirke P, Heald RJ, Moran BJ, Blomqvist L, Swift IR, Sebag-Montefiore D, Tekkis P, Brown G: Preoperative magnetic resonance imaging assessment of circumferential resection margin predicts disease-free survival and local recurrence: 5-year follow-up results of the MERCURY study. *Journal of clinical oncology* 32: 34-114 (2014)
53. Tripathi P, Rao SX, Zeng MS: Clinical value of MRI-detected extramural venous invasion in rectal cancer. *Journal of Digestive Diseases* 18: 2-12 (2017)
54. Xie H, Zhou X, Zhuo Z, Che S, Xie L, Fu W: Effectiveness of MRI for the assessment of mesorectal fascia involvement in patients with rectal cancer: A systematic review and meta-analysis. *Digestive Surgery* 31: 123-134 (2014)
55. Yu J, Huang DY, Xu HX, Li Y, & Xu QH: Correlation Between Magnetic Resonance Imaging-Based Evaluation of Extramural Vascular Invasion and Prognostic Parameters of T3 Stage Rectal Cancer. *Journal of Computer Assisted Tomography* 40: 537-542 (2016)
56. Zhang G, Cai YZ, Xu GH: Diagnostic Accuracy of MRI for Assessment of T Category and Circumferential Resection Margin Involvement in Patients With Rectal Cancer: A Meta-Analysis. *Diseases of the Colon & Rectum* 59: 789-799 (2016)

## Danksagung

In besonderem Maße bedanke ich mich bei Frau Prof. Dr. med. Andrea Rieber-Brambs für die Erarbeitung des Dissertationsthemas und die Möglichkeit, diese Arbeit unter Ihrer Betreuung durchführen zu dürfen. Ihre fachliche und organisatorische Unterstützung in sämtlichen Fragestellungen, ihre Geduld, Leitung und Motivation waren für mich über den gesamten Prozess der Arbeit von unschätzbarem Wert.

Mein besonderer Dank gilt auch Fr. Dr. Lange für ihre fachliche Beratung sowie Unterstützung bei der Datenerhebung und Auswahl des Bildmaterials. Ihr Engagement hat wesentlich zum Gesamtergebnis der Arbeit beigetragen.

Ganz herzlich bedanke ich mich auch bei meiner Freundin, Regina Danner, die mich während der gesamten Zeit sehr entlastet, motiviert und liebevoll begleitet hat und bei ihrem Bruder, Matthias Danner, für die fachliche Unterstützung bei der Datenanalyse.

Nicht zuletzt bedanke ich mich bei meiner Familie und meinen Freunden, auf die ich mich mein gesamtes Leben lang immer verlassen konnte und kann, was mir sehr viel bedeutet.

## **Lebenslauf**

Der Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt.