

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

Direktor: Prof. Dr. med. Jens Ricke

**Detektion von Imaging Biomarkern zur Charakterisierung und  
Prognoseabschätzung bei Patienten mit metastasiertem kolorektalen  
Karzinom**



Dissertation

zum Erwerb des Doktorgrades der Medizin

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von

Nils Jäger

aus Lemgo

2022

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

Berichterstatter: Prof. Dr. med. Wolfgang Kunz

Mitberichterstatter: Prof. Dr. med. Dominik Modest  
Prof. Dr. med. Jens Neumann

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

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 06.10.2022

## Inhaltsverzeichnis

Abkürzungsverzeichnis .....	5
Zusammenfassung.....	8
1 Einleitung .....	9
1.1 Das Kolorektale Karzinom (KRK).....	9
1.1.1 Epidemiologie .....	9
1.1.2 Volkswirtschaftliche Bedeutung .....	10
1.1.3 Ätiologie, Klassifikation und Krankheitsverlauf .....	12
1.1.4 Früherkennung und Diagnostik .....	20
1.1.5 Therapie .....	24
1.1.6 Beurteilung des Therapieansprechens und der Überlebenschancen..	35
1.2 Was bedeutet „Radiomics“?.....	40
2 Zielsetzung der Arbeit .....	47
3 Material und Methoden.....	48
3.1 Datensammlung.....	48
3.2 Segmentierung der Gesamttumorlast der Leber .....	48
3.3 Vorstellung der analysierten prognostischen Modelle .....	50
3.4 Datenanalyse .....	52
3.4.1 Statistik.....	52
3.4.2 Machine Learning.....	54
3.4.3 Verwendete Arbeitsprogramme.....	55
4 Ergebnisse.....	56
4.1 Zusammenstellung des Patientenkollektivs.....	56
4.2 Deskriptive Statistik – Basisdaten des untersuchten Patientenkollektivs	58
4.3 Segmentierungen der Gesamttumorlast der Leber .....	59
4.4 Datenanalyse und Überlebensvorhersage .....	60

## Inhaltsverzeichnis

5	Diskussion .....	67
5.1	Stärken und Limitationen der Studie .....	72
5.2	Ausblick.....	74
	Abbildungsverzeichnis.....	77
	Tabellenverzeichnis .....	79
	Literaturverzeichnis .....	80
	Anhang.....	95
	Veröffentlichungen .....	106
	Danksagung.....	107
	Eidesstattliche Erklärung .....	108

## Abkürzungsverzeichnis

1-YS	1-year survival
AIC	Akaike information criterion
APC-Gen	Adenomatöses Polyposis Coli Gen
ARP	Aerts radiomics prior model
AUC	Area under the curve
BIC	Bayesian information criterion
BMBF	Bundesministerium für Bildung und Forschung
BRAF	Rapidly accelerated fibrosarcoma homolog B1
BSC	Best supportive care
CAD	Computer-aided diagnosis and detection
CEA	Carcinoembryonic antigen
CI	Konfidenzintervall
C-index	Concordance index
CPH	Cox proportional hazard
CR	Complete response
CRC	Colorectal cancer
CRLM	Colorectal liver metastases
CT	Computertomografie
CUP	Cancer of unknown primary
CV	Cross validation/Kreuzvalidierung
DALY	Disability-adjusted life year
EGFR	Epidermal growth factor receptor
ESMO	European Society for Medical Oncology
FAP	Familiäre adenomatöse Polyposis
FDR	False discovery rate

## Abkürzungsverzeichnis

FME	Fraunhofer MEVIS
FOBT	Fäkaler okkultter Bluttest
FP	Fluoropyrimidine
gFOBT	Guajak-basierter fäkaler okkultter Bluttest
GMS	Geometric metastatic spread model
HNPCC	Hereditary non-polyposis colorectal cancer
HPP	Hyperplastischer Polyp
HR	Hazard ratio
HU	Houndsfield unit
IEN	Intraepitheliale Neoplasie
iFOBT	Immunologischer fäkaler okkultter Bluttest
KRAS	Kirsten rat sarcoma viral oncogene homolog
KRK	Kolorektales Karzinom
LLD	Liver limited disease
LV	Leucovorin
MAPK	Mitogen-activated protein kinase
mCRC	Metastatic colorectal cancer
mKRK	Metastasiertes kolorektales Karzinom
MMR	DNA-Mismatch-Reparatursystem
MSI	Mikrosatelliten-Instabilität
MSS	Mikrosatelliten-Stabilität
mt	Mutant
NED	No evidence of disease
OMD	Oligometastatic disease
ORS	Overall response rate
OS	Overall survival
PD	Progressive disease

## Abkürzungsverzeichnis

PET	Positronenemissionstomografie
PFS	Progression free survival
PR	Partial response
PTS	Primary tumor sidedness
QIB	Quantitative imaging biomarker
RECIST	Response evaluation criteria in solid tumors
ROC	Receiver operating characteristic
ROI	Region of interest
RQS	Radiomic quality score
RT	radiation therapy
SAT	Segmentation Annotation Tool
SA/V	Surface-area-to-volume ratio
SD	Stable disease
SIRT	Selective internal radiation therapy
SM	Syn-/metachronous
SPLIT	Zeitlicher 2/3-Split
SSA	Sessiles serratiertes Adenom
TACE	Transarterial chemoembolisation
TBS	Tumor burden score
Tis	Carcinoma in situ (nach TNM-Klassifikation)
TNM	Tumor, Node (Lymphknoten), Metastase
TSA	Traditionelles sessiles Adenom
UICC	Union Internationale Contre le Cancer
WLTB	Whole liver tumor burden
wt	Wild-type
VEGF	Vascular endothelial growth factor

## **Zusammenfassung**

Die vorliegende Arbeit zielt darauf ab, auf Basis von Computertomographie (CT)-Bildern frühere und präzisere Therapieentscheidungen bei Patienten mit metastasiertem kolorektalen Karzinom zu ermöglichen. Durch eine detaillierte, automatisierte und quantitative medizinische Bildanalyse könnte der Tumor beziehungsweise dessen Metastasen bereits bei Erstdiagnose prätherapeutisch charakterisiert und hinsichtlich der Prognoseabschätzung für den einzelnen Patienten analysiert werden. Eine quantitative Bildanalyse mit Techniken des maschinellen Lernens ermöglicht dies unter Umständen mit einer höheren Trefferquote als dies mit der derzeit eingesetzten manuellen Erfassung der Tumorgröße oder Verteilung von Metastasen möglich ist.

Konkretes Ziel dieser Arbeit war es deshalb, die Vorhersage des 1-Jahres-Überlebens von Patienten mit metastasiertem kolorektalen Karzinom (mKRRK) zu verbessern und die Vorhersagekraft verschiedener quantitativer Imaging Biomarker (QIB) über eine geometrische und radiomische Analyse des Gesamttumorvolumens der Leber im Vergleich zum etablierten Tumor Burden Score (TBS) abzuschätzen.

Durch die anvisierte prätherapeutische Risikostratifizierung und Prognoseabschätzung mittels Verfahren des maschinellen Lernens und der sich prospektiv daraus abzuleitenden Weiterentwicklung einer individualisierten Therapieplanung entsteht hier in der Zukunft ein direkter Mehrwert für Patienten mit kolorektalen Karzinomen. Eine gezieltere Überlebensvorhersage könnte bedeuten, dass Patienten mit metastasiertem kolorektalen Karzinom schneller der richtigen Therapie zugeführt werden oder ihnen die Nebenwirkungen einer möglicherweise unnötigen intensivierten Chemotherapie erspart werden könnten. Der konkrete Mehrwert für den Patienten spiegelt sich in einer Verbesserung der verbliebenen Lebensqualität und potenziell auch in einem Überlebensvorteil aufgrund der angestrebten präziseren Therapieentscheidungen wider. Diese Rationale ist konkreter Gegenstand der vorliegenden Arbeit und wird im weiteren Verlauf analysiert und diskutiert. Die vorliegende Arbeit setzt somit unmittelbar an einem weitverbreiteten klinischen Problem (der bisher unzureichenden Möglichkeit einer bildbasierten, prätherapeutischen Risikostratifizierung und Prognoseabschätzung) an und versucht dieses auf innovative Art und Weise zu lösen.



# 1 Einleitung

## 1.1 Das Kolorektale Karzinom (KRK)

### 1.1.1 Epidemiologie

Jede achte Krebserkrankung betrifft den Darm (1).

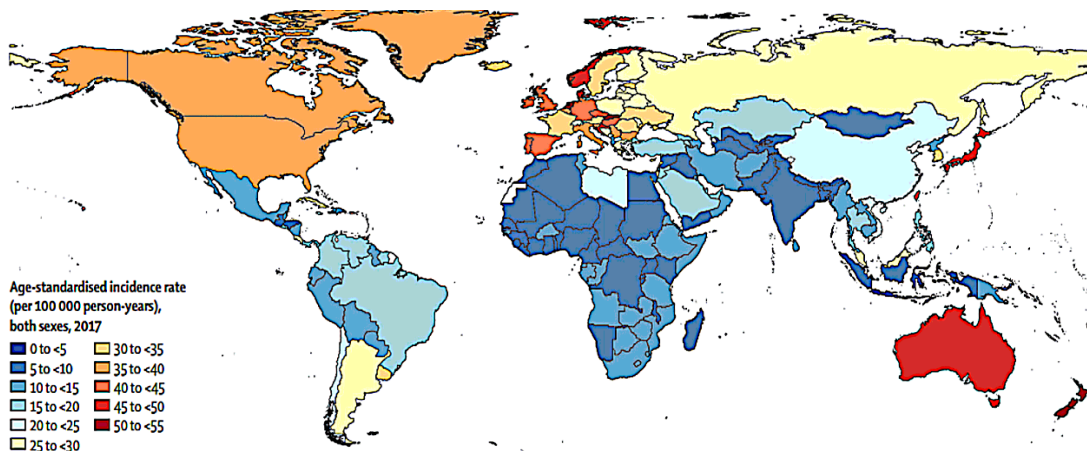
Mit rund 1,8 Millionen jährlichen Neuerkrankungen und fast 900.000 Todesfällen weltweit (Stand 2017) steht das kolorektale Karzinom global an dritter Stelle der häufigsten Krebserkrankungen und ist die zweithäufigste krebsbedingte Todesursache (2).

Unter der gemeinsamen Betrachtung beider Geschlechter ist es in Deutschland mit rund 60.600 jährlichen Neuerkrankungen und 24.200 Todesfällen (Stand 2018) nach dem Brustkrebs das häufigste Malignom, kutane Malignome außer Acht gelassen. Es hat aber im Gegensatz zu diesem ein deutlich schlechteres relatives 5-Jahres-Überleben von nur ca. 62% und ist so nach dem Lungenkrebs die häufigste krebsbedingte Todesursache, obwohl die altersstandardisierten Sterberaten in Deutschland für das KRK innerhalb der letzten 10 Jahre schon um mehr als 20% zurückgingen (1).

Das KRK ist eine Erkrankung des älteren Menschen mit einem mittleren Erkrankungsalter von 75 Jahren bei Frauen und 72 Jahren bei Männern. Mehr als die Hälfte aller Darmkrebspatienten ist zum Zeitpunkt der Erstdiagnose also älter als 70 Jahre und nur ca. 10% der Diagnosen werden vor dem 55. Lebensjahr gestellt (1).

In den Industrienationen Westeuropas ist die Inzidenz des KRK (Männer/Frauen pro 100.000: ca. 39/25) fast doppelt so hoch wie im weltweiten Schnitt (Männer/Frauen pro 100.000: ca. 21/14) und mehr als doppelt so hoch wie in den Entwicklungs- oder Schwellenländern Afrikas (Männer/Frauen pro 100.000: ca. 14/10) (3) (vgl. Abbildung 1).

**Abbildung 1:** Weltweite altersstandardisierte Inzidenzraten beim kolorektalen Karzinom pro 100.000 Einwohner pro Jahr, übernommen aus (2)



Im Zusammenhang mit der zunehmenden Adaptation der westlichen Lebensweise in diesen Ländern und der stetig steigenden weltweiten Lebenserwartung des Menschen (4, 5) wird erwartet, dass die globale Inzidenz und Mortalität des kolorektalen Karzinoms bis zum Jahre 2030 im Vergleich zu 2012 um 60% auf insgesamt 2.2 Millionen neuer Fälle und 1.1 Millionen Tode pro Jahr ansteigen wird (6).

### 1.1.2 Volkswirtschaftliche Bedeutung

Das KRK war im Jahre 2016 für ca. 19,14 Millionen DALY weltweit verantwortlich (7). DALY – *disability adjusted live year* – beschreibt die Summe aus der Anzahl verlorener Lebensjahre durch vorzeitigen Tod, von der durchschnittlichen Lebenserwartung des Menschen ausgehend, und der Anzahl der Jahre, die unter Behinderung mit der letztlich zum Tode führenden Krankheit gelebt wurde. Es ist damit ein Ausdruck für die verlorenen krankheitsfreien Jahre durch eine Krankheit und ein Marker für die gesellschaftliche Bedeutung dieser, also der Krankheitsbürde für den Menschen (8).

Die Gesamtheit der krebsbedingten Krankheitsbürde für die Menschheit ließ sich im Jahre 2016 in 244,57 Millionen DALY pro Jahr ausdrücken. Das KRK hat daran einen Anteil von rund 8%, das macht es nach dem Lungen- und Leberkrebs zur Krebsform mit der insgesamt drittgrößten Krankheitsbürde für den Menschen (7).

Zu den verlorenen Lebensjahren summieren sich die direkten und indirekten Kosten für das Gesundheitssystem und den Staat durch Darmkrebs. Die indirekten Kosten lassen sich in verlorenen Einkünften der an Darmkrebs erkrankten Bevölkerung und den daraus resultierenden geringeren Steuereinnahmen messen. Erstere beliefen sich in den USA im Jahre 2015 auf 9,4 Milliarden US-Dollar und machten damit 10% der durch Krebserkrankungen verlorenen Einkünfte aus (9). Zu den direkten Kosten für das Gesundheitssystem gehören die Behandlungskosten für das KKK. Diese beliefen sich, alle Krankheitsstadien eingeschlossen, in den USA im Jahre 2000 auf 7,49 Milliarden US-Dollar und werden bis zum Jahre 2020 auf voraussichtlich 13,03 Milliarden US-Dollar ansteigen, sich damit also innerhalb von 20 Jahren nahezu verdoppeln. (10). Grund dafür ist auch die in den letzten Jahrzehnten marktreif gewordene kostspielige *targeted therapy* mit monoklonalen Antikörpern, deren Rolle in Zukunft stark zunehmen wird (11-13).

Eine Behandlung des metastasierten Kolonkarzinoms über 6 Monate mit dem auch heute noch in Deutschland zur Erstlinientherapie gehörenden Chemotherapie-Schema FOLFIRI (14) kostete im Jahre 2016 in den USA ca. 13.400 US-Dollar (15). Durch die Zugabe des monoklonalen Antikörpers und EGFR-Inhibitors Cetuximab steigen diese Kosten um mehr als das 5-Fache auf ca. 84.100 US-Dollar an (15). Im Zuge der zunehmenden Rolle der kostspieligen Krebsimmuntherapie in Zukunft sowie durch neu erschlossene Forschungsschwerpunkte, z.B. an Checkpoint-Inhibitoren und CAR-T-Zellen (16, 17), werden auch die Behandlungskosten für Darmkrebs weiter ansteigen und Fragen nach Kosteneffizienz und Wirtschaftlichkeit künftig einen großen Stellenwert haben. Eine große Hürde zur Verbesserung der Kosteneffizienz dieser Immuntherapien wird auch die Bestimmung neuer prädiktiver Biomarker sein, um die richtigen Patienten für diese teuren Therapien selektieren zu können, da die derzeit auf dem Markt befindlichen monoklonalen Antikörper gegen Darmkrebs gerade in Schwellen- und Entwicklungsländern (noch) nicht kosteneffizient sind (18).

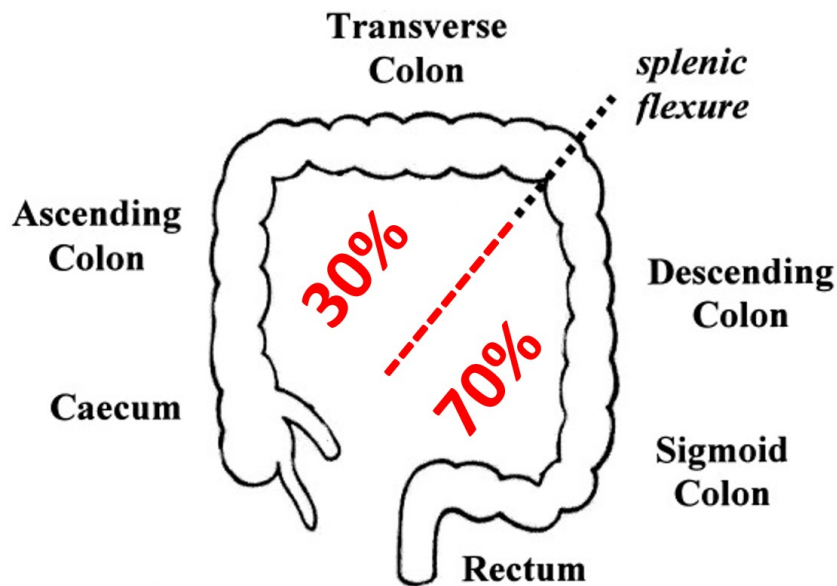
Die Behandlung von Darmkrebspatienten bindet auch im deutschen Gesundheitssystem viele Ressourcen. So gibt es deutschlandweit derzeit (Stand 2019) insgesamt 293 von der deutschen Krebsgesellschaft zertifizierte Darmzentren. Das sind mehr, als für jede andere Krebsentität existieren; die zweitmeisten Krebszentren gibt es für den Brustkrebs mit deutschlandweit 280 Brustzentren (19).

Das KRK ist also eine Krankheit von globaler Tragweite, dessen weltweite Inzidenz, Mortalität und Relevanz für die Gesundheitssysteme in der Zukunft stark zunehmen wird.

### **1.1.3 Ätiologie, Klassifikation und Krankheitsverlauf**

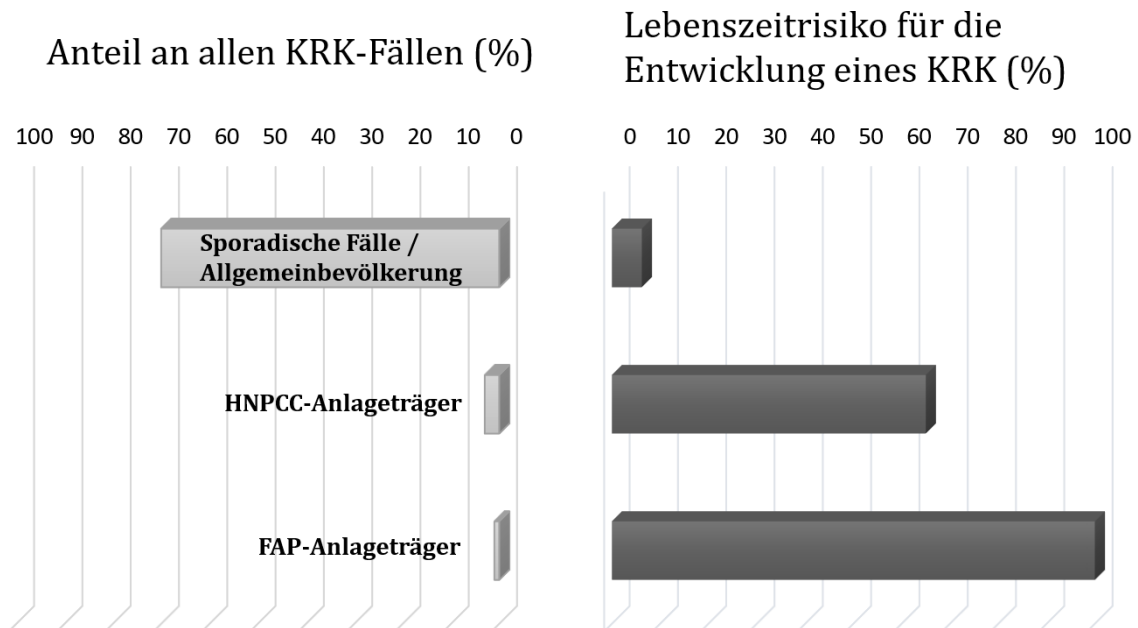
Ungefähr 30% der KRK-Fälle treten familiär gehäuft auf, aber nur bei ca. 5% sind die genetischen Ursachen und klinischen Auswirkungen so genau beschrieben, dass ihnen definierte erbliche Krankheiten zugeschrieben werden können (20). Die häufigste Ursache für erblichen Darmkrebs, die etwa 3% aller KRK-Fälle ausmacht, ist das hereditäre, nicht-polypöse Kolonkarzinom-Syndrom (HNPCC), auch Lynch-Syndrom genannt. Es beruht auf einer autosomal-dominant vererbten Mutation in einem von vier Genen des DNA-Mismatch-Reparatursystems (MMR). Bei einem Lebenszeitrisiko von 50-80% (21) (sonst 5,5% bei Frauen und 6,9% bei Männern (1)) und einem medianen Erkrankungsalter von bereits 45 Jahren (sonst 75 bei Frauen und 72 bei Männern (1)) führt es zu Darmtumoren, die zu über 50% im rechtsseitigen Kolon liegen (22) (sonst finden sich 70% der Tumore aboral der linken Kolonflexur und nur 30% rechtsseitig (14, 23)) und ist mit einer Mikrosatelliteninstabilität (MSI) als Zeichen gestörter DNA-Reparatur sowie einem erhöhten Risiko für extraintestinale Neoplasien assoziiert (20, 22, 24) (vgl. Abbildung 2).

**Abbildung 2:** Schema des Dickdarms mit Häufigkeit der Tumorlokalisationen in den Darm-Abschnitten, modifiziert nach (25)



Die zweithäufigste erbliche Darmkrebsform entsteht auf dem Boden einer familiären adenomatösen Polyposis (FAP). Sie ist für ca. 1% der KKK-Fälle verantwortlich, wird ebenso autosomal-dominant vererbt und bewirkt durch eine Mutation im APC-Tumorsuppressorgen eine multiple Ausbreitung von kolorektalen Adenomen bereits in jungen Jahren. Das Lebenszeitrisiko für die Entwicklung eines KKK ist quasi 100%, das mediane Erkrankungsalter liegt bei 39 Jahren und die derzeitige einzige kurative Therapieoption besteht aus einer prophylaktischen Proktokolektomie (20, 22).

**Abbildung 3:** Prävalenz und Lebenszeitrisiko der Entwicklung eines KKK bei verschiedenen Risikogruppen

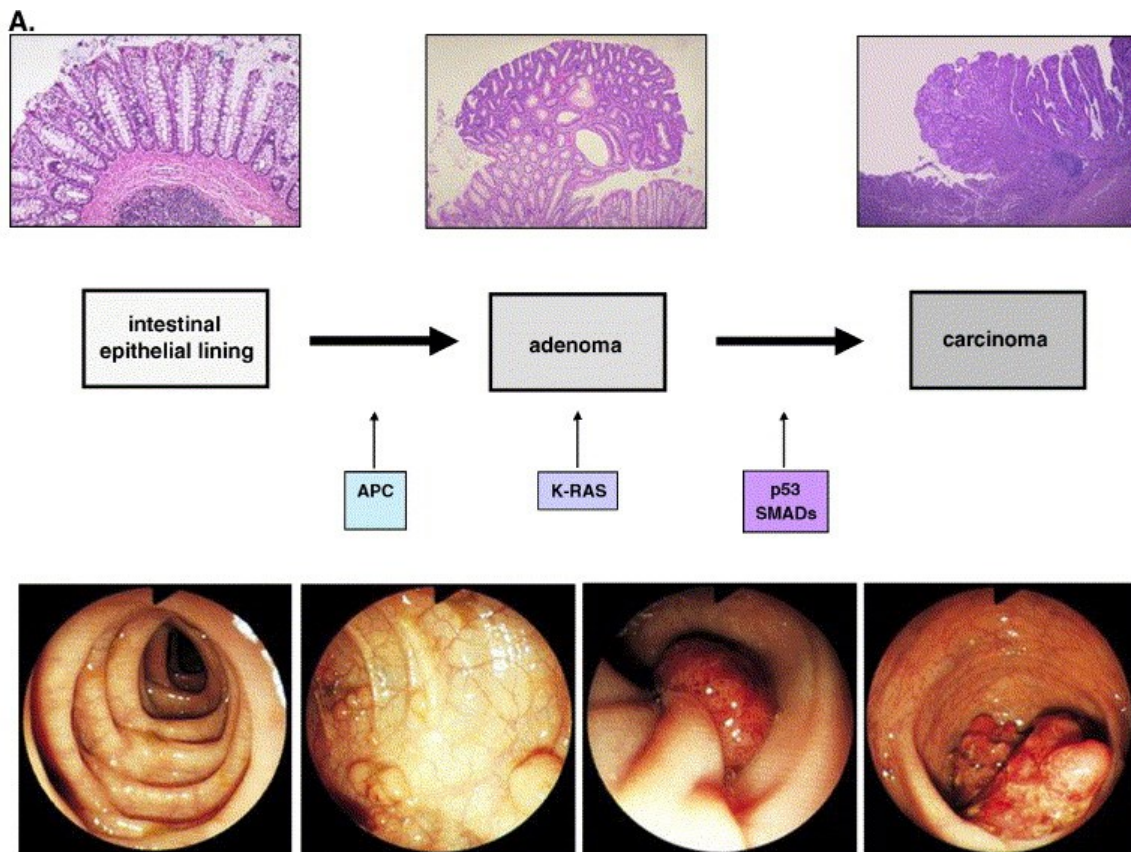


Da die Mehrheit der KKK aber sporadisch und in hohem Alter auftreten, ist die Krankheit von einem hohen Einfluss erworbener Risikofaktoren gekennzeichnet. Dazu gehören abseits von familiärer Vorbelastung und erblichen Krankheitsbildern (26) die chronisch entzündlichen Darmerkrankungen wie Colitis ulcerosa oder Morbus Crohn. Ebenso konnten Alkoholkonsum, Rauchen, eine westliche ballaststoffarme und fettreiche Ernährungsweise mit großen Mengen an rotem verarbeiteten Fleisch und wenig Obst und Gemüse sowie ein von Adipositas und Bewegungsmangel gekennzeichnete Lebensstil als Risikofaktoren identifiziert werden. Über die Hälfte aller KKK-Fälle lassen sich mit potenziell vermeidbaren Risikofaktoren in Verbindung bringen (27). Generell lassen sich zudem einige geschlechtsspezifische Unterschiede für das KKK ausmachen. So treten beispielsweise Darmtumore insgesamt häufiger und früher bei Männern auf als bei Frauen, sie liegen bei Frauen häufiger im rechtsseitigen Kolon als bei Männern und Frauen weisen bei intensiverer Chemotherapie ein erhöhtes Auftreten von schweren Nebenwirkungen auf (28, 29). Frauen scheinen aber dennoch ein insgesamt besseres Gesamtüberleben zu haben als Männer (30).

Auch der Nachweis histologisch gesicherter Adenome bei einer Koloskopie erhöht das Darmkrebsrisiko, denn Adenome bilden in fast allen Fällen die Vorstufe eines KKK (14, 31). Dieses entwickelt sich dabei zu 70-80% stufenweise über einen

längeren Zeitraum von ca. 10 Jahren aus einem adenomatösen Polypen heraus über die sog. Adenom-Karzinom-Sequenz, basierend auf einem Modell von Fearon und Vogelstein, durch mehrere aufeinander folgende Mutationen in definierten Genen, beginnend mit dem APC-Gen (s. Abbildung 4) (22, 32).

**Abbildung 4:** Die klassische Adenom-Karzinom-Sequenz nach (32), übernommen aus (33)

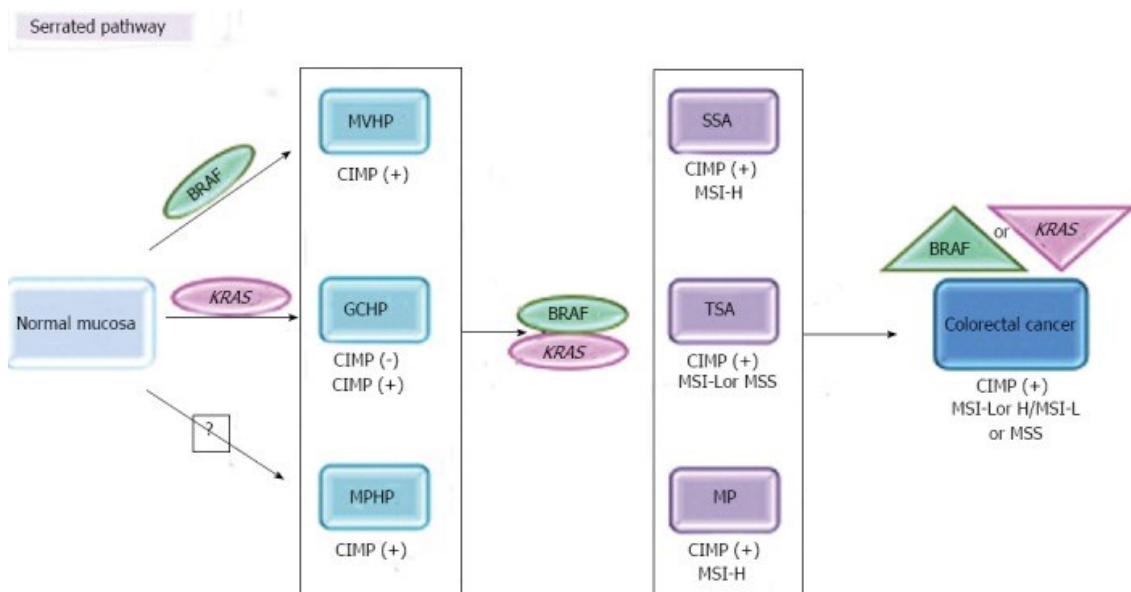


Entsprechend des Stadiums auf dem Weg zur malignen Entartung entlang der Adenom-Karzinom-Sequenz werden Dysplasien, akkurater als sogenannte intraepitheliale Neoplasien (IEN) bezeichnet, histologisch in Low-Grade-IEN und High-Grade-IEN eingeteilt (34).

„Klassische“ neoplastische Polypen werden des Weiteren je nach Histologie in tubulöse, tubulovillöse und villöse Adenome unterteilt, wobei das tubulöse Adenom die häufigste und das villöse Adenom die seltenste Form darstellt, letzteres aber die höchste Entartungsrate besitzt. Diese „klassischen“ Adenome machen ca. 88-95% aller Adenome aus (34-36).

Weitere molekulare Entstehungswege des KRK wurden in den letzten Jahren bekannt. Einerseits existiert neben der typischen Adenom-Karzinom-Sequenz die sog. „serratierte Route“, die zur Entstehung des sog. sessilen serratierten Adenoms (SSA) als Vorstufe zum KRK führt (37, 38) (vgl. Abbildung 5). Initiierend ist hier eine Mutation in einer von zwei Proteinkinasasen, entweder dem BRAF- oder dem KRAS-Onkogen, die jeweils wichtige Rollen in der sog. MAP-Kinase-Kaskade spielen, welche Auswirkungen auf Apoptose, Zellwachstum und -differenzierung hat (37, 39). Die Sequenz ist in etwa 20-30% der KRK-Fälle ursächlich (22), führt eventuell schneller zu einer malignen Entartung als der klassische Weg (14, 35) und kann deshalb mit eine Ursache für die sog. Intervallkarzinome sein, als welche KRK bezeichnet werden, die trotz adäquatem Koloskopie-Screening mit einer Häufigkeit von ca. 2 pro 1000 Patientenjahren auftreten (40).

**Abbildung 5:** Der serratierte Karzinogeneseweg, übernommen aus (37)



Der serratierte Karzinogeneseweg als Beispiel für weitere molekulare Entstehungswege des KRK mit wichtigen Schlüsselmutationen. MVHP: Microvesicular hyperplastic polyp; GCHP: Globet cell hyperplastic polyp; MPHP: Mucin-poor hyperplastic polyp; SSA: Sessile serrated adenoma; TSA: Traditional serrated adenoma; MP: Mixed polyp; MSI-H: High microsatellite instability; CIMP: CpG island methylator phenotype.

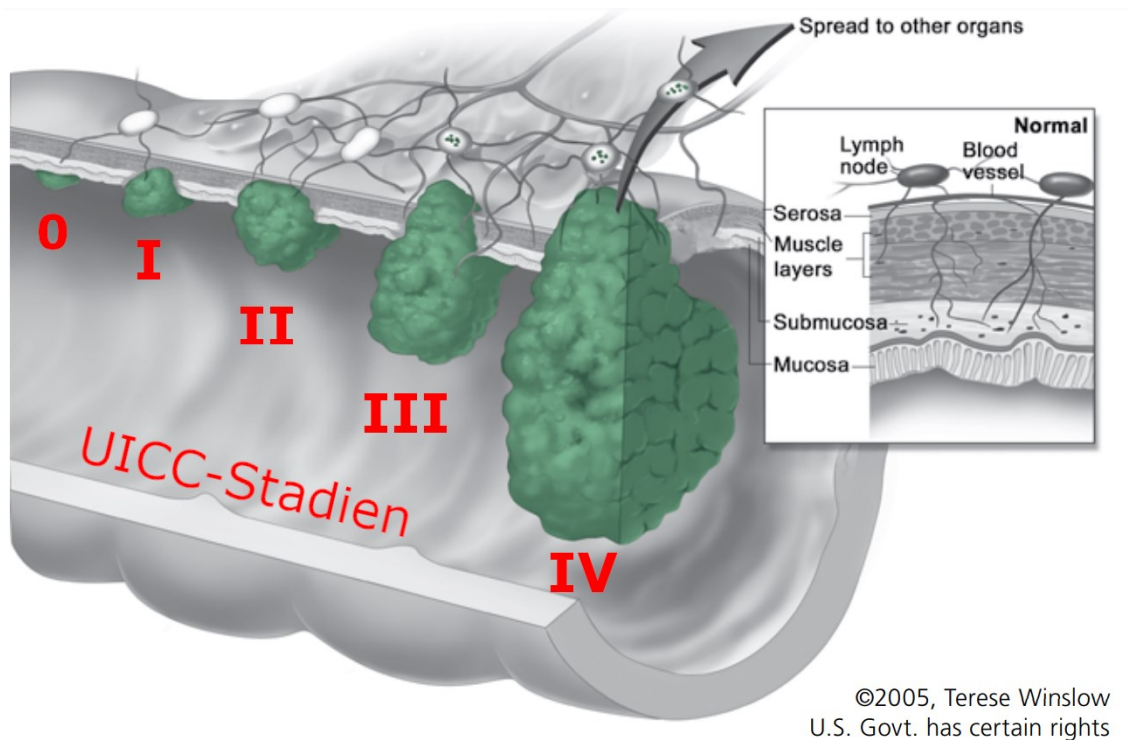
Die entstehenden SSA als mögliche alternative Vorstufe zum Darmkrebs unterscheiden sich sowohl in der Morphologie, Histologie und in der Lokalisation von den klassischen Adenomen. So ragen sie nicht polypös ins Darmlumen hervor, was ihr Auffinden in der Endoskopie erschwert, haben ein serratiertes,



„sägezahnartiges“ Kryptenepithel und kommen verstärkt im rechtsseitigen Kolon vor. Sie machen ca. 3-8% aller Adenome aus (35). Andererseits existiert ein gemischter Karzinogeneseweg, der die molekularen Veränderungen beider Entstehungswege in sich vereint und als resultierendes Adenom das villöse Adenom oder das sog. traditionelle sessile Adenom (TSA) zur Folge haben kann (35). Das TSA macht ca. 1-3% aller Adenome aus und entspricht in seiner Histologie dem SSA, ragt aber wie die klassischen Adenome polypös als IEN ins Darmlumen vor und kommt überwiegend im linksseitigen Kolon vor (35).

Die Tumorausbreitung des KKR wird in von der *Union Internationale Contre le Cancer* definierte Stadien im Rahmen einer TNM-Klassifikation mit den Untergruppen der Ausbreitung des Primärtumors, des regionalen Lymphknotenbefalls und dem Vorhandensein von Fernmetastasen eingeteilt. Beim UICC-Stadium 0 liegt ein noch nicht tiefere Schleimhautschichten infiltrierendes Carcinoma in situ (Tis) vor. Infiltriert der Primärtumor bereits die Submukosa (T1) bis hin zur Muscularis propria der Darmwand (T2), wächst aber noch nicht über diese hinaus, liegt das UICC-Stadium I vor. Bei Infiltration darüber hinaus bis hin zu einer Invasion in perikolisches Fettgewebe (T3), ins Peritoneum oder Nachbarorgane (T4), aber noch ohne Lymphknoteninvasion (N0) oder Fernmetastasierung (M0), erreicht er das UICC-Stadium II. Beim Stadium III liegen dann bereits regionale Lymphknotenmetastasen vor (N1-2) und sobald der Tumor Fernmetastasen bildet (M1), befindet er sich im Stadium IV (22). Diese Fernmetastasen sind es auch, die – sofern nicht resektabel – im Verlauf prognosebestimmend sind (14).

**Abbildung 6:** Bildhafte Darstellung der KRK-Stadien nach der UICC-Klassifikation, modifiziert nach (41)



Diese UICC-TNM-Klassifikation ist immer noch der Goldstandard im Tumorstaging, d.h. der Einteilung der Ausbreitung, und der wichtigste prognostische Faktor bei der Erstdiagnose eines KRK sowie grundlegender Wegweiser für eine stadienadaptierte Therapie (42).

So sinken die 5-Jahres-Überlebensraten beim KRK in Korrelation zu UICC-Stadium I bis IV drastisch. Bei Erstdiagnose eines Tumors im Stadium I liegt diese noch bei bis zu 95%, in Stadium II bei bis zu 85%, 55-65% im Stadium III und lediglich 5% im Stadium IV (22). Die in der folgenden Tabelle dargestellte vereinfachte Form der UICC-Klassifikation entspricht im Wesentlichen der „veraltete[n] aber immer noch gebräuchliche[n]“ (43) sogenannten Dukes-Klassifikation, die den Grad der Tumordinfiltration ebenso in vier Stadien von A bis D einteilt.

**Tabelle 1:** Vereinfachte UICC-Klassifikation des kolorektalen Karzinoms und Dukes-Stadien nach (14) und (43)

UICC-Stadium	Dukes-Stadium	Primärtumor	Lymphknotenstatus	Fernmetastasen
0	/	Tis	N0	M0
I	A	T1-T2	N0	M0
II	B	T3-T4	N0	M0
III	C	jedes T	N1-N2	M0
IV	D	jedes T	jedes N	M1

Da das KRK sich im Verlauf aber erst relativ spät symptomatisch manifestiert, wird es nicht immer sofort erkannt. Charakteristische Frühsymptome fehlen, deshalb kann es oft erst bei Blutbeimengungen im Stuhl und einer möglicherweise daraus entstehenden chronischen Blutungsanämie, durch plötzliche Änderungen im Stuhlverhalten bis hin zur Ileussyptomatik sowie einer B-Symptomatik, bestehend aus Fieber, Nachtschweiß und Gewichtsverlust, auffallen (22). In drei von vier Fällen wird ein KRK erst nach dem Auftreten von Symptomen diagnostiziert und in mehr als der Hälfte (51%) aller Fälle befindet sich das KRK bei Erstdiagnose bereits im Stadium III oder IV (44). Jeder Vierte hat bei der Erstdiagnose eines KRK bereits Fernmetastasen (14). Insgesamt ist der häufigste Weg eine hämatogene Metastasierung, welche mit 50-70% am öftesten die Leber und zu 10-30% die Lunge betrifft (45). Insbesondere der Leber kommt dabei nicht nur im Falle des KRK eine herausragende Rolle zu. Sie ist einer der häufigsten Metastasierungsorte maligner Erkrankungen generell. So weisen bis zur Hälfte aller malignen Erkrankungen eine Leberbeteiligung auf und Lebermetastasen sind um ein Vielfaches häufiger als primäre Lebertumore (46). Gastrointestinale Tumore wie das KRK bilden dabei den größten Anteil der Lebermetastasen verursachenden Primärtumore, und Lebermetastasen finden sich auch am häufigsten als Erstmanifestation einer malignen Erkrankung bei unklarem Primarius (= *Cancer of Unknown Primary*, CUP-Syndrom) (47). Im Falle des KRK spielt hierbei auch die Lokalisation des

Primärtumors eine Rolle. So ist die Häufigkeit von Lungenmetastasen beim Rektumkarzinom höher als beim Kolonkarzinom und nimmt von oral nach aboral im Verlauf vom rechtsseitigen Kolonkarzinom bis zum Rektumkarzinom um mehr als das Doppelte zu (48, 49). Dies hängt damit zusammen, dass die venöse Drainage eines tiefen Rektumkarzinoms nicht wie die des restlichen Darmtraktes als Pfortadertyp primär über die Leber erfolgt, sondern über die tiefen und mittleren Rektalvenen als Cava-Typ direkt in die Lunge (50-53). Des Weiteren ist eine lymphogene Metastasierung des KRK in regionale oder ferne Lymphknoten oder eine direkte Metastasierung ins Peritoneum möglich, welches 5-15% aller Metastasierungsorte ausmacht (45). Selten metastasiert ein KRK in andere Orte wie das ZNS oder die Knochen (45).

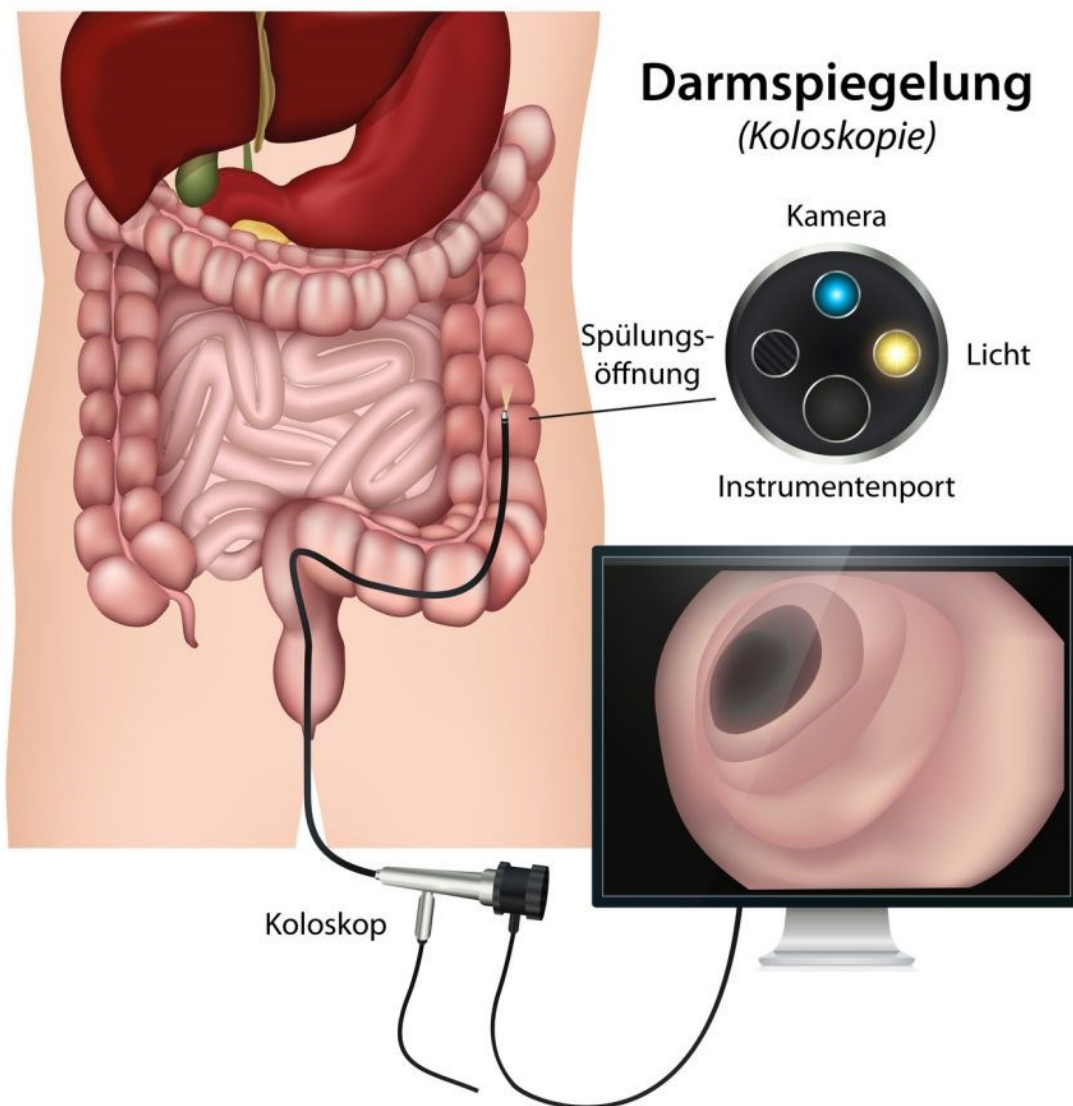
#### **1.1.4 Früherkennung und Diagnostik**

Da das KRK ein Karzinom des älteren Menschen ist, sollten zur Risikovermeidung in der Normalbevölkerung neben der Vermeidung der o.g. Risikofaktoren regelmäßige Früherkennungsuntersuchungen ab dem 50. Lebensjahr erfolgen (14).

Idealerweise bestehen diese zwischen dem 50. und 54. Lebensjahr aus der jährlichen Kombination einer digital rektalen Untersuchung mit einem immunchemischen Test auf okkultes Blut im Stuhl, dem iFOBT, der mittlerweile den früher verwendeten Gujakttest (herkömmlicher gFOBT, auch Hämoccult®) abgelöst hat (31, 54). Dieser Testart kommt deshalb eine besondere Bedeutung zu, da der FOBT als unabhängiger Prädiktor für eine frühe Diagnose eines sich entweder noch im Stadium I oder II befindlichen KRK identifiziert werden konnte und Anämien in 60% und Blut im Stuhl in 45% der Fälle einer KRK-Erstdiagnose als vorausgegangene Symptome beschrieben wurden, was diese damit zu den beiden häufigsten Begleitsymptomen des KRK macht (44).

Ab dem 55. Lebensjahr ist der Goldstandard zur Früherkennung aufgrund größter Sensitivität und Spezifität dann die komplette Koloskopie (14). Der Stellenwert dieser Früherkennungsmethode hat in der jüngsten Vergangenheit noch weiter zugenommen, sodass insbesondere Männern mittlerweile bereits ab dem 50. Lebensjahr die Koloskopie empfohlen wird (55).

**Abbildung 7:** Schema zur Koloskopie, übernommen aus (56)



© bilderzweig / Fotolia.

Bei gesundem Befund soll eine Koloskopie dabei nach 10 Jahren wiederholt werden. Werden allerdings dabei Polypen entdeckt, können diese abgetragen werden und das empfohlene Intervall für die nächste Koloskopie verkürzt sich, sofern dabei  $>2$  oder  $>1\text{cm}$  messende Adenome entdeckt werden, auf 3 Jahre. Können diese nur schrittweise in „piecemeal“-Technik oder inkomplett abgetragen werden, oder werden  $>5$  Adenome oder gar frühe Karzinome entdeckt, verkürzt sich das Intervall für die nächste Kontrollkoloskopie weiter (14).

Als Alternative für eine komplette Koloskopie werden ein zweijährliches FOBT-Screening, die Sigmoidoskopie, die Kapselendoskopie, genetische Stuhltests sowie die CT-/MR-Koloskopie genannt (14, 31). Der Vorteil der Koloskopie besteht jedoch

darin, dass zum Beispiel im Gegensatz zur Sigmoidoskopie das gesamte Kolon eingesehen werden kann und im Vergleich zur CT-/MR-Koloskopie gefundene Adenome direkt abgetragen werden können.

Da ein FOBT zwar nicht invasiv, aber weniger spezifisch als eine Koloskopie ist und beispielsweise ebenso auf ein Hämorrhoidalleiden hindeuten kann, schließt sich auch einem positiven Befund im FOBT trotzdem immer eine Koloskopie als nächster Schritt in der Diagnostik an (14).

Werden Polypen bei einer Koloskopie vollständig abgetragen, verringert sich das relative Darmkrebsrisiko für diese Patienten um 90% (14).

Während Abdomen-Sonographie und Röntgen-Thorax die primären Untersuchungen in der weiteren Basisdiagnostik nach Diagnose eines KRK darstellen, kommt bei unklarem Befund oder Verdacht auf Fernmetastasen der Computertomographie (CT) eine zentrale Rolle in der weiteren Diagnostik, dem präoperativen Tumorstaging und der Verlaufsbeurteilung während der Therapie zu (14).

Obwohl die Magnetresonanztomographie (MRT) zwar sensitiver für das Erfassen einzelner kleiner Leberherde sowie der Gesamttumorlast der Leber ist, ist die CT insbesondere für die Beurteilung der Lagebeziehungen der Lebermetastasen zum umliegenden Gewebe und damit auch am besten für die Beurteilung der Resektabilität von Lebermetastasen geeignet und wird mit der MRT in der klinischen Praxis gleichwertig zur Resektionsplanung verwendet (14, 57, 58). Dies hat einen relevanten klinischen Einfluss, insbesondere da sich die Frage eines kurativen Behandlungsansatzes lediglich bei einer potenziell chirurgisch resektablen Krebserkrankung stellt (14). Obwohl in der S3-Leitlinie die Sonographie noch immer als primäre Methode der Wahl aufgeführt ist, um Lebermetastasen auszuschließen, unterliegt die CT geringeren Sensibilitätsschwankungen und einer höheren diagnostischen Genauigkeit zur Metastasendetektion (59). Sie ist beim Kolonkarzinom am besten für die präoperative lokale Ausbreitungsdiagnostik und das bildbasierte Tumorstaging geeignet (14, 57-60). Lediglich beim Staging des Rektumkarzinoms ist die MRT der CT hinsichtlich der lokalen Tumorausbreitung bekanntermaßen überlegen (61). Die CT stellt auch in der Früherkennung eine ergänzende Alternative zur kompletten Koloskopie bei eingeschränkter Durchführbarkeit, z.B. aufgrund von Adhäsionen oder eines stenosierenden Tumors,

dar (14). Hier erweist sie sich gegenüber der Koloskopie aufgrund der gleichzeitigen Detektierbarkeit anderer abdomineller Erkrankungen wie beispielsweise dem Aortenaneurysma als effizient und eventuell sogar kostengünstiger (62). Die CT ist zudem mit dem geringsten Zeitaufwand für die Untersuchung und in der Regel auch mit geringeren Kosten verbunden als die MRT (63-66).

Wie die MRT kann auch eine PET/CT-Untersuchung eine bildgebende Alternative zur Diagnostik und Verlaufskontrolle beim KRK darstellen. Ein Nachteil dieser Untersuchungsmethode ist jedoch, dass beispielsweise bei Patienten, die unter systemischer, zytotoxischer oder zytostatischer Chemotherapie stehen, durch die medikamenteninduzierte Hemmung des Tumormetabolismus auch die für die PET-Komponente entscheidende Aufnahme des radioaktiv markierten Glukose-Tracers in die Tumorzellen gehemmt wird. Dies kann unter Umständen eine reduzierte Sensibilität zur Folge haben. Aus all den oben genannten Gründen ist die CT deshalb die derzeit am besten geeignete Methode mit dem geringsten zeitlichen Untersuchungsaufwand, der breitesten Verfügbarkeit zum Staging und zur Verlaufsbeurteilung in der bildbasierten Therapieplanung des metastasierten KRK (14, 67).

Standardprotokoll für das CT-basierte, prätherapeutische Staging des KRK ist dabei ein Multidetektor-Spiral-CT mit Kontrastmittelgabe in der portalvenösen Phase, da sich damit insbesondere der Fernmetastasenstatus gut erfassen lässt (53, 68). Ein Multidetektor-Spiral-CT beschreibt dabei eine moderne Bauweise des CT-Geräts, welches durch mehrere Detektoren in der Lage ist, mehrere Schichten gleichzeitig und mit kontinuierlichem Vorschub des Patienten aufzunehmen. Die Untersuchung dauert dadurch in der Regel nur wenige Sekunden (69). Für die CT-Untersuchung mit Darstellung der Strukturen in portalvenöser Kontrastmittelpase erfolgt die Aufnahme idealerweise 45-50 Sekunden nach Beginn der Applikation eines jodhaltigen Kontrastmittels (wie z.B. 100-120 ml Iohexol mit einer Geschwindigkeit von 2-3 ml pro Sekunde) (53). Absolute Kontraindikationen gegen die CT bestehen keine, lediglich bei einer Schwangerschaft sollte der Einsatz der CT aufgrund der Strahlenbelastung für das Ungeborene vermieden werden. Jodallergie, Kontrastmittelallergie, Hyperthyreose sowie Nierenerkrankungen sind relative Kontraindikationen gegen den Einsatz von Kontrastmittel in der CT (70).

Unter die präoperative Tumordiagnostik fällt neben der Bildgebung auch die molekularbiologische Diagnostik. Standard ist hierbei die Bestimmung des

karzinoembrionalen Antigens (CEA) im Blut als Verlaufsparemeter, der im Falle von Lebermetastasen auch prognostischen Wert hat. Auch gibt es Hinweise auf prädiktive Informationen, die ein erhöhtes CEA enthält, da solche Patienten signifikant besser auf EGFR-Antikörper anzusprechen scheinen (71). Die Marker Ca19-9 und Ca 125 stehen als alternative Tumormarker zur Debatte (14).

Vor dem Start einer medikamentösen Tumorthapie sind ebenfalls einige molekularbiologische Testungen erforderlich, um die Wirksamkeit bestimmter Chemotherapeutika oder Antikörper abschätzen zu können und prognostische Information zu gewinnen. Von klinischer Bedeutung für die Erstlinientherapie sind dabei die Tests auf RAS- und BRAF-Mutationen. So wirkt beispielsweise der EGFR-Antikörper Cetuximab nur auf Tumore, die einen RAS-Wildtyp aufweisen (14, 72-75). Ebenso wie nachgewiesene RAS-Mutationen sind die meisten BRAF-Mutationen von negativ prädiktiver Bedeutung (14, 76-79).

### **1.1.5 Therapie**

Die Therapie des KRK richtet sich nach vielen Faktoren. Dem UICC-Stadium kommt eine entscheidende Bedeutung bei der Therapiewahl zu.

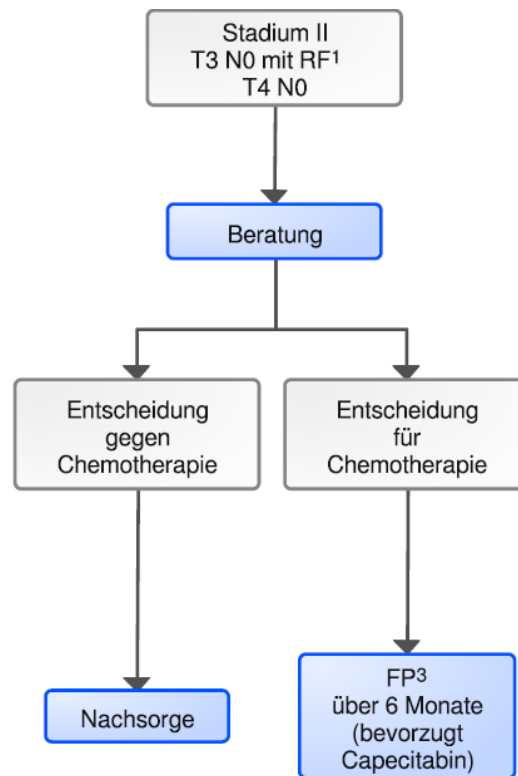
Im UICC-Stadium I-III des Kolonkarzinoms besteht ein kurativer Therapieansatz, in welchem die radikale chirurgische Resektion des Primärtumors den höchsten Stellenwert hat. Standard sind dabei eine Ausräumung von mindestens 12 regionären Lymphknoten im Abflussgebiet des Tumors, ein Sicherheitsabstand von mindestens 10 cm Darm vom Tumorrund ausgehend sowie eine En-bloc-Resektion tumoradhärenter Organe. Im Stadium I ist keine adjuvante Chemotherapie indiziert, da sie die Prognose nicht verbessert (31)

Im Stadium II ist die Datenlage bezüglich des Nutzens einer adjuvanten Chemotherapie heterogener. Aufgrund der derzeit in der S3-Leitlinie angeführten Studien wird sie nur für Patienten mit ausgewählten Risikofaktoren für ein erhöhtes Rezidivrisiko empfohlen und ist immer eine Risiko-Nutzen-Abwägung (14, 80, 81). Diese sind ein T4-Stadium des Primarius, eine Lymph- oder Blutgefäßinfiltration, weniger als 12 perioperativ untersuchte Lymphknoten, eine Tumorperforation oder intraoperativer Tumoreinriss sowie Notfalloperationen. Ebenso profitieren



Patienten mit einer MSS eher von einer adjuvanten Chemotherapie im Stadium II als Patienten mit einer MSI (31).

**Abbildung 8:** Therapieschema bei lokal fortgeschrittenem KRK, übernommen aus (31)



<sup>1</sup>RF - Risikofaktoren

<sup>3</sup>FP – 5-Fluoruracil (*flouopyrimidine*)

Eindeutiger ist die Datenlage im Stadium III. Auch hier ist der Therapieanspruch kurativ, es wird jedoch eindeutig eine perioperative, adjuvante Chemotherapie empfohlen (31).

Das aktuell wirksamste Schema hier besteht aus dem ältesten gegen Darmkrebs wirksamen Chemotherapeutikum 5-Fluoruracil (5-FU) und Oxaliplatin. Das 5-FU wird entweder in Kombination mit der wirksamkeitssteigernden Folinsäure (Leucovorin (LV)) als Langzeitinfusion über 24-48 Stunden oder in Form der Prodrug Capecitabin oral gegeben. Die beiden Darreichungsformen können dabei in ihrer Wirksamkeit als gleichwertig erachtet werden (82, 83). Dieses Schema nennt sich je nach verwendetem Fluoropyrimidin dann FOLFOX oder CAPOX und gehört auch in der späteren palliativen Situation zur Erstlinientherapie. Andere Chemotherapeutika wie monoklonale Antikörper oder der Topoisomerase-I-

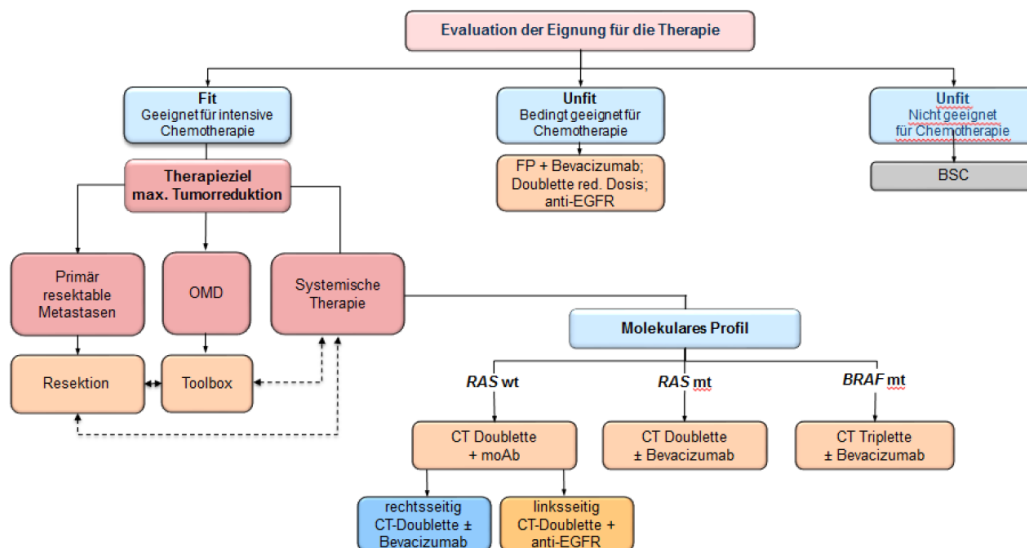
Inhibitor Irinotecan, wie sie in der palliativen Situation zum Einsatz kommen, finden im Stadium III des Kolonkarzinoms jedoch noch keine Anwendung (14, 31).

Die häufigsten Nebenwirkungen dieser medikamentösen Tumorthherapie sind in Bezug auf das 5-FU Diarrhoe und Stomatitis, wohingegen Oxaliplatin vor allem Nausea, Mukositis und eine Polyneuropathie verursachen kann.

Die auf das 5-FU zurückzuführenden Nebenwirkungen können sich bei genetischen Störungen des 5-FU-Abbaus verstärken, ansonsten ist die Wechselwirkung der oralen Prodrug Capecitabin mit Protonenpumpenhemmern nennenswert, da diese in Kombination zu einer Wirkungsverminderung des Capecitabins führen können (31).

Ab einer Entfernung des aboralen Tumorrandes von weniger als 16 cm zur Anokutanlinie zählt ein maligner Darmtumor als Rektumkarzinom (14). Unterschiede in der Therapie des Rektumkarzinoms im Vergleich zu der des Kolonkarzinoms bestehen hierbei vor allem in den Stadien II und III, da im Falle des Rektumkarzinoms zusätzlich eine neoadjuvante Radiochemotherapie vor der operativen Resektion angestrebt wird. Die adjuvante Therapie richtet sich jedoch, sofern keine Bestrahlung erfolgt, nach dem Vorgehen beim Kolonkarzinom (14).

**Abbildung 9:** Therapiestruktur im Stadium IV, übernommen aus (14)

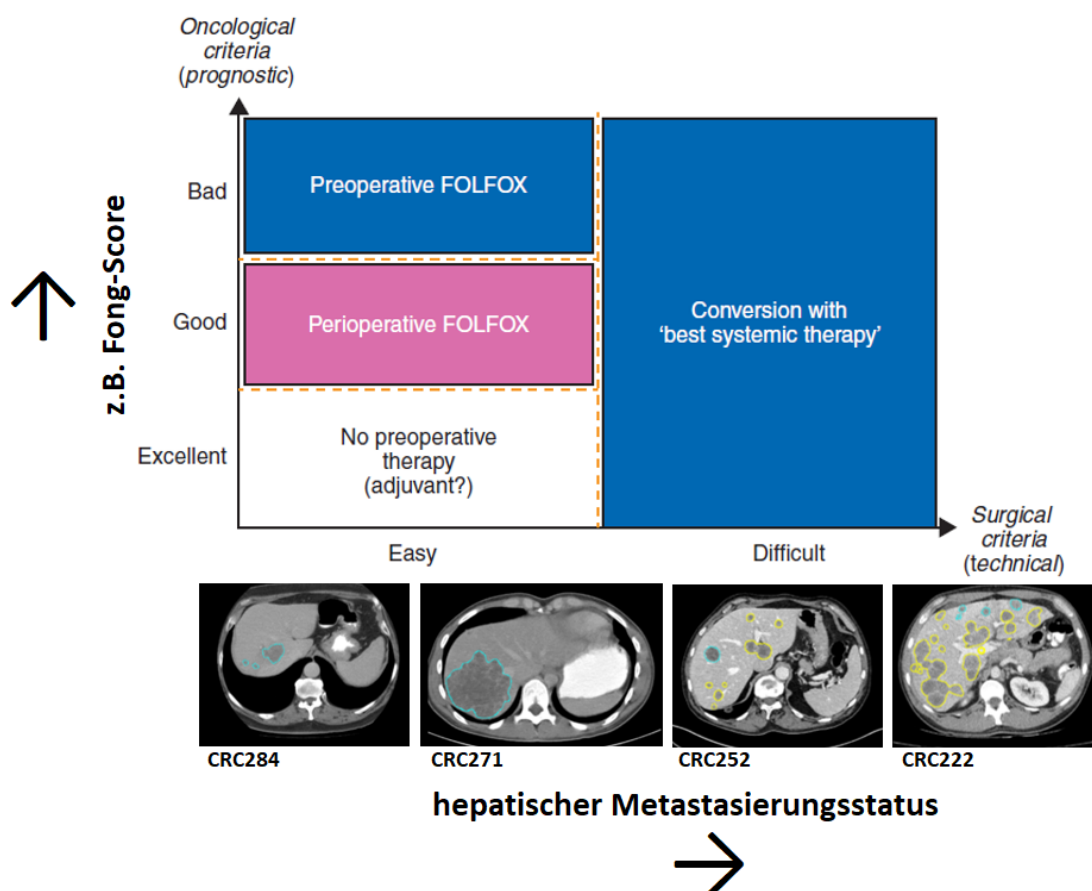


Im UICC-Stadium IV haben die Patienten bereits Fernmetastasen entwickelt, am häufigsten in der Leber. Hier besteht zwar bei mittlerweile bis zu 25% der Patienten

mit synchron hepatisch metastasiertem KRK auch die Chance auf Heilung, die Ausgangslage ist jedoch oft eine palliative (31, 84, 85) (vgl. Abbildung 9).

Die Krankheitsausbreitung, gemessen an Anzahl, Größe, Lokalisation und damit der Resektabilität der (Leber-)Metastasen, ist dabei neben onkologischen Kriterien von entscheidender Bedeutung für die zu verfolgenden Therapieziele und die Prognose des Patienten (vgl. Abbildung 10).

**Abbildung 10:** Auswirkungen chirurgischer und onkologischer Kriterien auf die Therapie des „fitten“ KRK-Patienten, modifiziert nach (86)



Ergänzt wurden, beispielhaft als chirurgische Kriterien, CT-Schnittbilder verschiedener hepatischer Metastasierungsstadien mit unterschiedlicher Resektabilität (anhand von CRC284, CRC271, CRC252 und CRC 222, vierer Patienten dieser wissenschaftlichen Arbeit) sowie der Fong-Score als eine exemplarische Zusammenstellung relevanter onkologischer Kriterien.

Sofern möglich, ist zunächst eine chirurgische Resektion der Lebermetastasen das Mittel der ersten Wahl (vgl. Abbildung 9 und 10). Ob eine solche Sinn macht, entscheidet sich sowohl aufgrund von technischen und onkologischen Kriterien, die jedoch bisher nicht international einheitlich prospektiv evaluiert sind. So empfiehlt die *European Society for Medical Oncology* (ESMO) aber hinsichtlich der technischen

Umsetzbarkeit, dass sich bei einer R0-Resektion mindestens 30% des ursprünglichen Lebergewebes erhalten lassen sollten, bzw. das Lebergewicht/Körpergewicht-Verhältnis bei mindestens 0,5 liegen sollte. Das entspräche einem Gewicht von 350 Gramm bei einem 70 kg schweren Menschen (86).

**Abbildung 11:** Der Fong-Score, übernommen aus (87)

<b>Table 5. CLINICAL RISK SCORE FOR TUMOR RECURRENCE</b>						
<b>Score</b>	<b>Survival (%)</b>					<b>Median (mo)</b>
	<b>1-yr</b>	<b>2-yr</b>	<b>3-yr</b>	<b>4-yr</b>	<b>5-yr</b>	
0	93	79	72	60	60	74
1	91	76	66	54	44	51
2	89	73	60	51	40	47
3	86	67	42	25	20	33
4	70	45	38	29	25	20
5	71	45	27	14	14	22

Each risk factor is one point: node-positive primary, disease-free interval <12 months, >1 tumor, Size >5 cm, CEA >200 ng/ml.

Die onkologischen Kriterien, die zur Beurteilung der Sinnhaftigkeit einer primären Metastasenresektion herangezogen werden, sollen Aufschluss über das Rezidivrisiko geben, da Patienten mit hohem Risiko und vielen prognostisch negativen Faktoren eher vom sofortigen Beginn einer systemischen Therapie profitieren. Hierzu gehören die Anzahl der Metastasen, eine extrahepatische Metastasierung oder weitere Kriterien von prognostisch negativer Bedeutung, wie sie exemplarisch im Fong-Score zusammengefasst sind (vgl. Abbildung 11) (87). Umgekehrt kann man erwägen, bei fehlenden Risikofaktoren und primär resektabler Erkrankung die perioperative medikamentöse Tumorthherapie

auszulassen, da es Hinweise dafür gibt, dass diese Patienten nach der Resektion eher vom alleinigen chirurgischen Eingriff profitieren (88).

Ist eine R0-Resektion nicht möglich oder sind schwierige Operationsbedingungen aufgrund der Krankheitsausbreitung zu erwarten, sollte mit einer systemischen Tumortherapie als Konversionstherapie begonnen werden, um potenziell resektable Lebermetastasen nach einer Reevaluation nach einem Zeitraum von 12-16 Wochen, in welchem das Maximum einer möglichen Reduktion der Tumorgröße infolge der systemischen Therapie erreicht wird, sekundär doch noch reseziieren zu können (86). Durch die Resektion von Lebermetastasen lässt sich die 5-Jahres-Überlebenschance auf bis zu 40% steigern (22).

Bei Patienten mit moderatem und hohem Rezidivrisiko und resektablen Lebermetastasen wird aufgrund der Ergebnisse der EORTC 40983 Intergroup Studie eine medikamentöse Chemotherapie nach dem bereits im Stadium III in der medikamentösen Tumortherapie zum Einsatz kommenden FOLFOX/CAPOX-Schema jeweils über 3 Monate vor und nach der Operation empfohlen (31, 84).

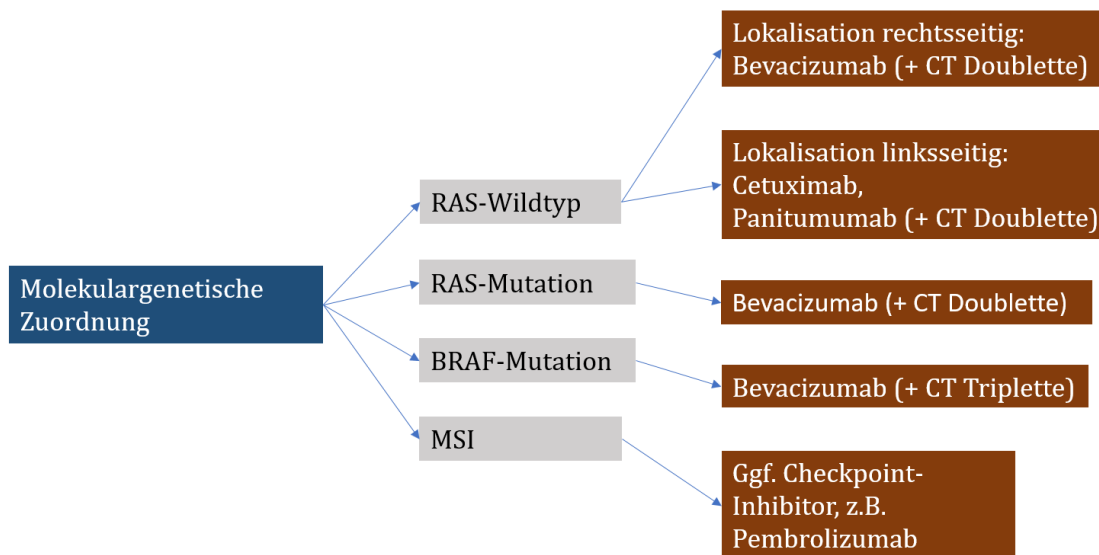
In der palliativen Situation wird zudem der Einsatz von Irinotecan anstatt Oxaliplatin zusammen mit 5-FU/LV im sog. FOLFIRI-Schema in der Erstlinientherapie als gleichwertig erachtet. Bei Patienten mit gutem Allgemeinzustand kann auch ein Triplet aus allen drei Chemotherapeutika, das sog. FOLFOXIRI-Schema, erwogen werden. Dieses ist potenziell am wirksamsten, jedoch auch mit den meisten Nebenwirkungen behaftet (86). Da insbesondere rechtsseitig gelegene KRK eine tendenziell schlechtere Prognose haben als linksseitige, wird dieses Triplet gerade in dieser Patientengruppe empfohlen (86).

In den letzten Jahrzehnten hat zudem das molekulargenetische Profil des KRK im Rahmen der *targeted therapy* therapeutische Bedeutung erlangt. Mittlerweile existieren in der Erstlinientherapie zwei Arten von Antikörpern, die in Kombination mit den gängigen Chemotherapie-Schemata die Überlebenszeit und das progressionsfreie Überleben der Patienten weiter verlängern können (86).

So existieren zum einen die EGFR-Antikörper Cetuximab und Panitumumab, die gegen den *epithelial growth factor receptor* gerichtet sind. Zum anderen gibt es den VEGF-Antikörper Bevacizumab, der als Angiogenesehemmer gegen den *vascular endothelial growth factor* wirkt.

Die EGFR-Antikörper haben zwar bessere Ansprechraten als die VEGF-Antikörper, wie *Heinemann et al.* in der FIRE-3 Studie nachweisen konnten (vgl. Abbildung 13). Sie wirken jedoch nur bei Tumoren, deren RAS- oder BRAF-Onkogene nicht mutiert sind (89). Da aber ca. 40% aller KRK RAS-Mutationen und 11,5% BRAF-Mutationen aufweisen, muss hier auf Bevacizumab als Alternative ausgewichen werden (86, 90) (vgl. Abbildung 12).

**Abbildung 12:** Systemtherapie nach molekularer Einordnung des mKRK



## Einleitung

Auch hinsichtlich der Wahl des Fluoropyrimidins gibt es Einschränkungen bei den EGFR-Antikörpern. Da die gastrointestinalen Nebenwirkungen bei Capecitabin stärker als bei 5-FU sind und sich zu den GI-Nebenwirkungen unter Cetuximab summieren, und da Cetuximab bei intermittierend gegebenem 5-FU keinen zusätzlichen Nutzen zu bringen scheint, kann die Kombination von Irinotecan und/oder Oxaliplatin mit Capecitabin in Kombination mit EGFR-Antikörpern nicht empfohlen werden (82, 86, 91, 92).

Ebenso spielt die Lokalisation für die Wahl des Antikörpers in der intensiven Chemotherapie eine Rolle. So scheint Bevacizumab bei rechtsseitig gelegenen Kolonkarzinomen besser zu wirken als EGFR-Antikörper (93). Aufgrund ihrer tendenziell schlechteren Prognose gibt es aus der TRIBE Studie Hinweise dafür, dass gerade in diesen Fällen eine möglichst intensive Chemotherapie aus dem FOLFOXIRI-Triplett in Kombination mit Bevacizumab für die Patienten den größten Nutzen haben könnte (94) (vgl. Abbildung 13).

**Abbildung 13:** Auszug ausgewählter Studien zum KRK, übernommen aus (95)

Trial	Design	N° of patients BRAFm	Line of treatment	Therapy	mOS	mPFS	ORR
Keynote-164	Phase II	23/124 (18.5%)	≥2 lines	Pembro	-	-	34.7%
Checkmate-142	Phase II	29/119 (24.3%)	≥2* lines	Nivo + Ipi	-	-	55%
Corcoran <i>et al</i>	Phase I	142/142 (100%)	≥2 lines	Dabra + Pani Dabra + Pani + Trame Trame + Pani	13.2 months 8.2 months -	3.5 months 4.2 months 2.6 months	10% 21% 0%
BEACON	Phase III	665/665 (100%)	≥2 lines	Enco + Bini + Cetu Enco + Cetu CT + Cetu	9.3 months HR: 0.38, <i>p</i> < 0.001 9.3 months HR: 0.40, <i>p</i> < 0.001 5.9 months	4.3 months 4.2 months 1.5 months	26.8% 19.5% 1.8%
SWOG S1406	Phase II	106/106 (100%)	≥2 lines	Irino + Cetu + Vemu Irino + Cetu	- -	4.2 months 2 months HR: 0.50, <i>p</i> = 0.001	17% 4%
Checkmate-142	Phase II	17/45 (37.7%)	1 line	Nivo + Ipi	-	-	71%
Keynote-177	Phase III	57/307 (18.5%)	1 line	Pembro CT	NR 23.5 months	16.5 months 8.2 months HR: 0.60, <i>p</i> = 0.0002	43.8% 33.1%
Loupakis <i>et al</i>	Phase II	25/25 (100%)	1 line	FOLFOXIRI + Beva	24.1 months	11.8 months	72%
Masi <i>et al</i>	Phase II	10/57 (17.5%)	1 line	FOLFOXIRI + Beva	23.8 months HR = 0.76, NS	12.8 months HR = 0.89, NS	90%
TRIBE	Phase III	28/508 (5.5%)	1 line	FOLFOXIRI + Beva FOLFIRI + Beva	19 months 10.7 months HR = 0.54, NS	7.5 months 5.5 months HR = 0.57, NS	56% 42%
TRIBE-2	Phase III	66/679 (9.7%)	1 line	FOLFOXIRI + Beva FOLFOX + Beva	27.4 months 22.5 months HR = 0.82, <i>p</i> = 0.032	6.2 months 5.6 months HR = 0.87, NS	- -

Abbreviations: mCRC, Metastatic colorectal cancer; N°, Number; BRAFm, BRAF mutated; Beva, Bevacizumab; Pembro, Pembrolizumab; Nivo, Nivolumab; Ipi, ipilimumab; Dabra, Dabrafenib; Vemu, Vemurafenib; Pani, Panitumumab; Trame, Trametinib; Enco, Encorafenib; Cetu, Cetuximab; Irino, Irinotecan; CT, Chemotherapy; Bini, Binimetinib; M, Months; NR, Non reach; NS, Non statistically significant

Zusammengefasst wird also bei linksseitigen RAS-Wildtyp-Tumoren ein EGFR-Antikörper hergenommen, bei rechtsseitigen und bei RAS- oder BRAF-mutierten Tumoren ein VEGF-Antikörper (14).

Die Ursachen für den Abbruch einer intensiven Chemotherapie sind entweder ein Krankheitsprogress oder intolerable Toxizität der Therapie selbst. Da diese insbesondere bei Oxaliplatin-haltigen Schemata unter kontinuierlicher Gabe schneller eintreten kann als ein Krankheitsprogress, hat sich gezeigt, dass eine intermittierende Gabe von Oxaliplatin keine Nachteile für die Patienten hinsichtlich der Überlebenszeit bringen muss (96). Es wird also, entsprechend den ESMO-Empfehlungen, die intensive Therapie nach dem FOLFOX/CAPOX-Schema plus Antikörper lediglich über 6-8 je zweiwöchige Zyklen als sogenannte Induktionstherapie gegeben, bevor sich eine Erhaltungstherapie bis zum nächsten Krankheitsprogress anschließt, die im Idealfall lediglich aus einem Fluoropyrimidin plus Antikörper, jedoch ohne Oxaliplatin, besteht. Liegt im Verlauf ein erneuter Krankheitsprogress vor, so kann eine Re-Induktion einer intensiven Chemotherapie über denselben Zeitraum oder gegebenenfalls der Wechsel auf ein anderes Chemotherapieschema, die Zweitlinie, erfolgen (31, 86, 96).

Da die Studienlage für Oxaliplatin-haltige Schemata derzeit besser ist als für Schemata mit Irinotecan, wird dort momentan nach wie vor die intensive Chemotherapie so lange empfohlen, wie sie für den Patienten verträglich ist (86).

Als dritte Entscheidungsebene, neben der Resektabilität und der Molekulargenetik des Tumors, gibt die ESMO zudem Empfehlungen zur Einteilung der Patienten in „fit“ oder „unfit“ anhand des klinischen Allgemeinzustandes heraus (vgl. Abbildung 14). Dies hat insbesondere in Bezug darauf, welche KRK-Patienten mit nicht initial resektablen Lebermetastasen für eine intensive adjuvante Chemotherapie infrage kommen, einen Mehrwert (86).



**Abbildung 14:** Revidierte ESMO-Patientengruppen mit mKRK und nicht initial resektablen Metastasen, übernommen aus (86)

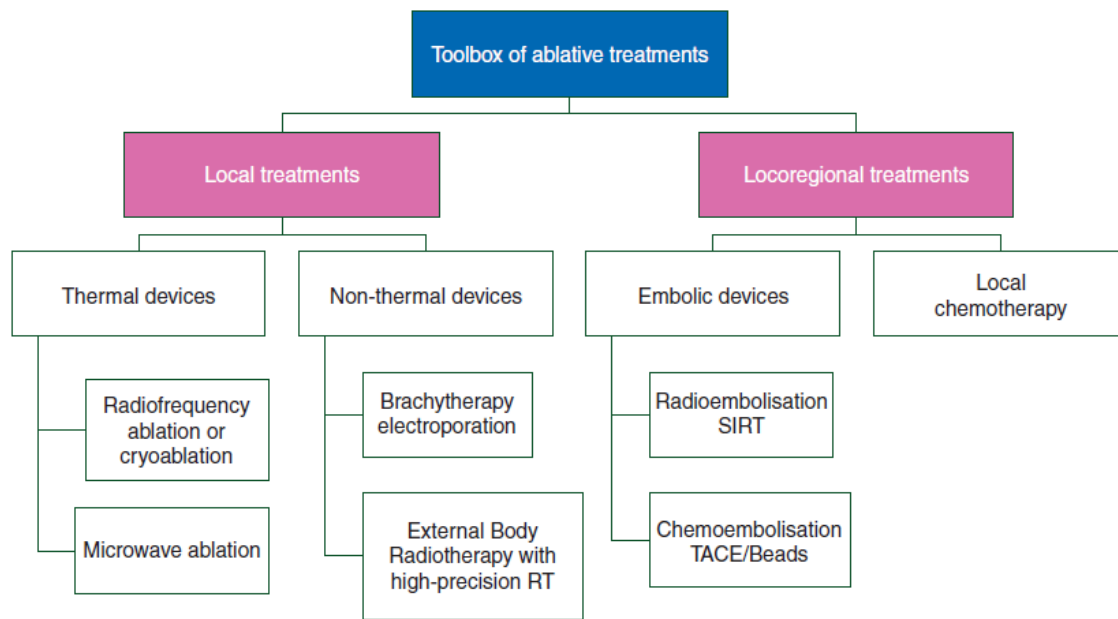
Table 6. Revised ESMO groups for treatment stratification of patients according to whether patients are 'fit' or 'unfit'			
Patient's classification	'Fit' patients		'Unfit' patients
	Group 1	Group 2	
Clinical presentation	A) Conversion and achievement of NED B) Impending clinical threat, impending organ dysfunction and severe (disease-related) symptoms Treatment biomarker driven: <i>RAS</i> wt, <i>RAS</i> mt, <i>BRAF</i> mt patient subgroups	Asymptomatic patients No impending clinical threat Resection not an option Treatment biomarker driven: <i>RAS</i> wt, <i>RAS</i> mt, <i>BRAF</i> mt patient subgroups	Best supportive care
Treatment goal	A) Cytoreduction, followed by R0 resection; NED achieved by LAT B) Improvement of symptoms and hence avoidance of rapid evolution and prolonged survival	Disease control and hence prolonged survival	Palliative

LAT, local and ablative therapy; mt, mutant; NED, no evidence of disease; wt, wild-type.

Bei Patienten, deren Metastasen nicht resektabel sind und die nicht fit genug für eine intensive Chemotherapie sind, steht eine weniger intensive Therapie bestehend aus einem Dosis-reduzierten Regime oder einem alleinigen Fluoropyrimidin plus Antikörper neben der symptomatischen Therapie im Sinne eines *best supportive care* (BSC) im Vordergrund (86).

Insbesondere im Falle einer fortgeschrittenen Metastasierung gibt es in der palliativen Situation neben der systemischen Therapie zudem noch die Optionen lokalablativer Verfahren mit dem Ziel, die Überlebenszeit der Patienten weiter zu verlängern oder akute Komplikationen infolge der Metastasierung zu verhindern. (86). Hier sind die Radiofrequenzablation, die Lasertherapie und die stereotaktische Bestrahlung nennenswert, viele groß angelegte Studien existieren bezüglich der Ergebnisse jedoch noch nicht (31). Die verschiedenen derzeit gängigen Möglichkeiten einer lokalablativen Metastasenbehandlung stellt die ESMO in einer „Toolbox“ zusammen (vgl. Abbildung 15) (86).

**Abbildung 15:** ESMO-Toolbox ablativer Verfahren, übernommen aus (86)



ESMO-Toolbox ablativer Verfahren. SIRT: selective internal radiation therapy, RT: radiation therapy, TACE: transarterial chemoembolisation

Zusammengefasst unterliegt die optimale Therapie im Stadium IV also einer Entscheidungsfindung auf drei Ebenen:

1. Sind die Metastasen in Bezug auf die anhand von technischen sowie onkologischen Kriterien evaluierte Krankheitsausbreitung resektabel?
2. Welche *targeted therapy* kann auf Basis des molekulargenetischen Profils des Tumors eingesetzt werden?
3. Sind die Patienten anhand des klinischen Allgemeinzustandes geeignet für eine intensive medikamentöse Tumorthherapie?

Insgesamt beträgt das mediane Gesamtüberleben von Patienten mit mKRK mittlerweile rund 30 Monate – mehr als das Doppelte an Überlebenszeit wie noch vor 20 Jahren (86).

### 1.1.6 Beurteilung des Therapieansprechens und der Überlebenschancen

Nachdem im Jahre 1981 die WHO erstmals einheitliche Kriterien zur Beurteilung des Therapieansprechens bei soliden malignen Tumorerkrankungen veröffentlicht hat, ist der derzeitige Goldstandard die Beurteilung aufgrund der RECIST-Kriterien, die 2000 publiziert und 2009 in der Version 1.1 überarbeitet wurden (97). Hierbei handelt es sich um die manuelle Messung unidimensionaler Tumordurchmesser von insgesamt bis zu fünf repräsentativen und wiederholt messbaren *target lesions*, wovon je bis zu zwei im selben befallenen Organ liegen dürfen. Als messbar zählt eine Läsion, wenn sie im längsten Durchmesser mindestens 10 mm misst. Auch bis zu zwei Lymphknoten können als eigenständiges „Organ“ in die Messung als *target lesions* einbezogen werden (98). Da nicht pathologisch vergrößerte Lymphknoten aber normalerweise bis zu 10 mm im Durchmesser groß sind, hat man sich darauf geeinigt, hier nur definitiv pathologisch vergrößerte Lymphknoten in die Messung mit einzubeziehen, deren kleinster Durchmesser mindestens 15 mm beträgt (97).

Am besten anzuwenden sind die RECIST-Kriterien dabei in Schnittbildgebungen. Am ehesten wird die CT als ubiquitär verfügbares Diagnostikum mit großer Reproduzierbarkeit empfohlen. Die MRT ist ebenso möglich und in Einzelfällen, wie z.B. bei einfach beurteilbaren Lungenmetastasen, auch ein klassisches Röntgen, wobei hier die Mindestgröße der Läsionen für eine ausreichende Beurteilbarkeit auf 20 mm festgelegt wurde. Der Ultraschall wird aufgrund seiner Untersucherabhängigkeit nicht empfohlen (97).

Bei den *target lesions* sollte für eine optimale Beurteilung des Ansprechens der systemischen Tumorthherapie keine chirurgische Exzision oder Vorbehandlung durch lokoregionäre Techniken erfolgen, ansonsten erfolgt eine Klassifizierung als *non evaluable* (98).

Damit die Patienten anhand der RECIST-Kriterien evaluiert werden können, sollten sie zum Zeitpunkt der Baseline-Messung, der Staging-CT vor Therapieeinleitung, mindestens eine messbare Läsion aufweisen. Die Baseline-Messung sollte so nah wie möglich am Therapiebeginn stattfinden und nie mehr als 4 Wochen davor (97).

Die Durchmesser der bis zu fünf *target lesions* werden nun zu einem *baseline sum diameter* aufsummiert. Alle anderen Läsionen werden als *non-target lesions*

ebenfalls markiert, welche im Verlauf lediglich als "vorhanden", "nicht vorhanden" oder als "eindeutige Progression" deklariert, jedoch nicht wie die *target lesions* neu vermessen werden (97).

Normalerweise alle 6-8 Wochen, möglichst zusammenfallend mit dem Ende eines Chemotherapie-Zyklus, findet nun eine (Re-)Evaluierung des Therapieansprechens mithilfe einer erneuten CT-Messung, dem *follow-up*, statt. Es wird erneut die Durchmessersumme der in der *baseline* vermessenen *target lesions* errechnet (97).

Hieraus wird das Therapieansprechen der Patienten nach den RECIST-Kriterien nun in vier Kategorien eingeteilt.

**Tabelle 2:** Die RECIST-Kriterien, Version 1.1, nach (97)

Kategorien des Therapieansprechens	Abkürzung	Veränderung im Summendurchmesser der vermessenen <i>target lesions</i>
Komplettes Ansprechen ( <i>complete remission</i> )	CR	Verschwinden aller <i>target-</i> und <i>non-target lesions</i> , sowie Reduktion aller Lymphknoten auf einen minimalen Durchmesser von max. 10 mm.
Partielles Ansprechen ( <i>partial remission</i> )	PR	Mindestens 30% Reduktion, mit Referenz zum <i>baseline sum diameter</i>
Progrediente Erkrankung ( <i>progressive disease</i> )	PD	Mindestens 20% Anstieg, mit Referenz zum schmalsten im Verlauf gemessenen Summendurchmesser, mindestens jedoch ein Größenzuwachs von 5 mm. Ebenso zählt das Auftauchen jeglicher neuer Läsionen als PD
Stabile Erkrankung ( <i>stable disease</i> )	SD	Länge zwischen -30% und +20% im Vergleich zum schmalsten im Verlauf gemessenen Summendurchmesser

Diese sind:

1. Bilden sich alle sichtbaren Läsionen im Verlauf vollständig und alle Lymphknoten auf eine normale Größe von im schmalsten Durchmesser maximal 10 mm zurück, spricht man von einer *complete response* (CR).
2. Eine Reduktion des *baseline sum diameters* im Verlauf um mindestens 30% im Vergleich zum *baseline sum diameter* bezeichnet man als *partial response* (PR).
3. Bei einem *target lesion sum diameter*, der um mindestens 20% größer als die schmalste im Verlauf gemessene Durchmessersumme ist, spricht man von einer *progressive disease* (PD). Diese schmalste im Verlauf gemessene Summe muss dabei nicht der in der *baseline* gemessenen Summe entsprechen. Mindestens muss zur Kategorisierung einer PD jedoch ein absoluter

Größenzuwachs von mindestens 5 mm erfolgen. Ebenso zählt das Auftauchen jeglicher neuer Läsionen im Verlauf als PD.

4. Liegt der Größenunterschied des *target lesion sum diameters* im Bereich zwischen -30% und +20% im Vergleich zur schmalsten im Verlauf gemessenen Durchmessersumme, spricht man von einer *stable disease* (SD).

Die bisher gültigen RECIST-Kriterien in der Version 1.1 berücksichtigen mit der unidimensionalen Messung des Tumordurchmessers derzeit lediglich anatomische Gesichtspunkte in der Evaluierung des Therapieansprechens.

Eine Größenzunahme des unidimensionalen Tumordurchmessers um 20%, welcher nach den RECIST-Kriterien noch als SD klassifiziert werden würde, entspräche einer Flächenänderung von 44% und einer Volumenänderung von 73% in der Tumormasse. In Anbetracht dessen wird insbesondere die teilweise Therapieansprache (PR) und die stabile Erkrankung (SD) im Rahmen der RECIST-Kriterien nur undifferenziert betrachtet.

Es gibt deshalb zunehmend Ansätze, die bestehenden RECIST-Kriterien für bestimmte Krankheiten und weitere Messmethoden zu modifizieren. So wurden die EORTC- und PERCIST-Kriterien auf Basis der RECIST-Kriterien für die PET-Diagnostik entwickelt. Diese unterscheiden, ebenso wie die RECIST-Kriterien, vier Kategorien, die jedoch anstelle von anatomischen Gesichtspunkten nun den veränderten Metabolismus des Tumorgewebes unter Therapie zur Beurteilung des Therapieansprechens solider Tumore in der PET heranziehen (99, 100).

Ebenso wurden für das Ovarialkarzinom und das Prostatakarzinom Leitlinien veröffentlicht, die den Wert des Tumormarkers CA-125 und des prostataspezifischen Antigens PSA in die Beurteilung des Therapieansprechens mit einbeziehen. (101-105)

Auch gibt es von anderen Forschungsgruppen Bestrebungen, die RECIST-Kriterien an bestimmte Krebsformen anzupassen beziehungsweise krankheitsspezifisch zu modifizieren. Die RECIST-Forschungsgruppe nennt in ihrer Publikation von 2016 bereits diverse Beispiele hierfür (101).

Auch gibt es mit der *early tumour shrinkage* und der *depth of response* neue Ansätze, die die „Responder“-Gruppe im Therapieansprechen insbesondere unter einer

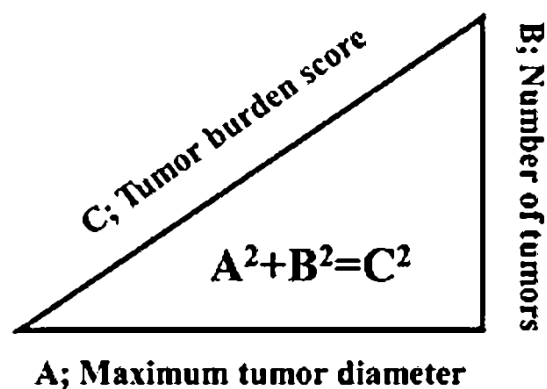
*targeted therapy* weiter unterteilen können und prädiktive Information enthalten, was die Notwendigkeit einer schnellen Vorhersage des Therapieansprechens verdeutlicht (106).

Darüber hinaus existieren auch hinsichtlich kolorektaler Lebermetastasen Bestrebungen, in der Beurteilung des Therapieansprechens auf Bevacizumab morphologische Kriterien unabhängig von der Größenänderung des Tumors unter Therapie heranzuziehen, die besser mit dem Überleben der Patienten korrelieren als die RECIST-Kriterien (107, 108).

Zudem ließe sich mit einer automatisierten Messung der Tumordurchmesser oder einer Tumorummetrie anstelle einer rein manuellen Vermessung eine signifikante Zeitersparnis für den befundenden Radiologen bei gleicher oder sogar geringerer Messvariabilität erbringen (109-111).

Mit dem *tumor burden score* (TBS) existiert aber auch seit 2017 ein neuer, rein geometrischer Ansatz, mit welchem das Überleben von Patienten mit metastasiertem kolorektalen Karzinom mit größerer Genauigkeit als mit bisher etablierten morphologischen Kriterien wie dem Fong-Score vorhergesagt werden konnte (vgl. Abbildung 16) (112).

**Abbildung 16:** Der „Tumor Burden Score“, übernommen aus (112)



Während die RECIST-Kriterien noch keine intratumorale Gewebheterogenität berücksichtigen, deren Veränderung mittlerweile als stärkster Auslöser für die Entwicklung einer Resistenz gegen die medikamentöse Tumorthapie gesehen wird (113), finden sich neuerdings für einige Tumorentitäten wie das Mamma-, Prostata- und Lungenkarzinom Datenbanken und Befundungssysteme, mit deren

Hilfe versucht wird, diese Heterogenität strukturiert im radiologischen Befund zu beschreiben (114-117).

Im Anblick neuer Therapieformen herrscht also ein Umdenken und das Bestreben, die bisher vorherrschende rein anatomische Messung des Tumordurchmessers zur Beurteilung des Therapieansprechens, um neue Aspekte zu erweitern oder zu ersetzen. Es werden zunehmend multidimensionale Messungen, metabolische Komponenten, Tumormarker und die differenzierte Morphologie des Tumors selbst herangezogen, um neue Kriterien und neue Biomarker zu finden, mit denen sich ein Therapieansprechen noch schneller und besser mit einem progressionsfreien Überleben und Gesamtüberleben des Patienten korrelierend erfassen ließe. Die Patienten könnten einen schnelleren Therapiewechsel bei schnellerem Erkennen einer fehlenden Ansprache erhalten und gegebenenfalls nicht nur psychologisch, sondern auch von einem besseren progressionsfreien Überleben und Gesamtüberleben profitieren.

## 1.2 Was bedeutet „Radiomics“?

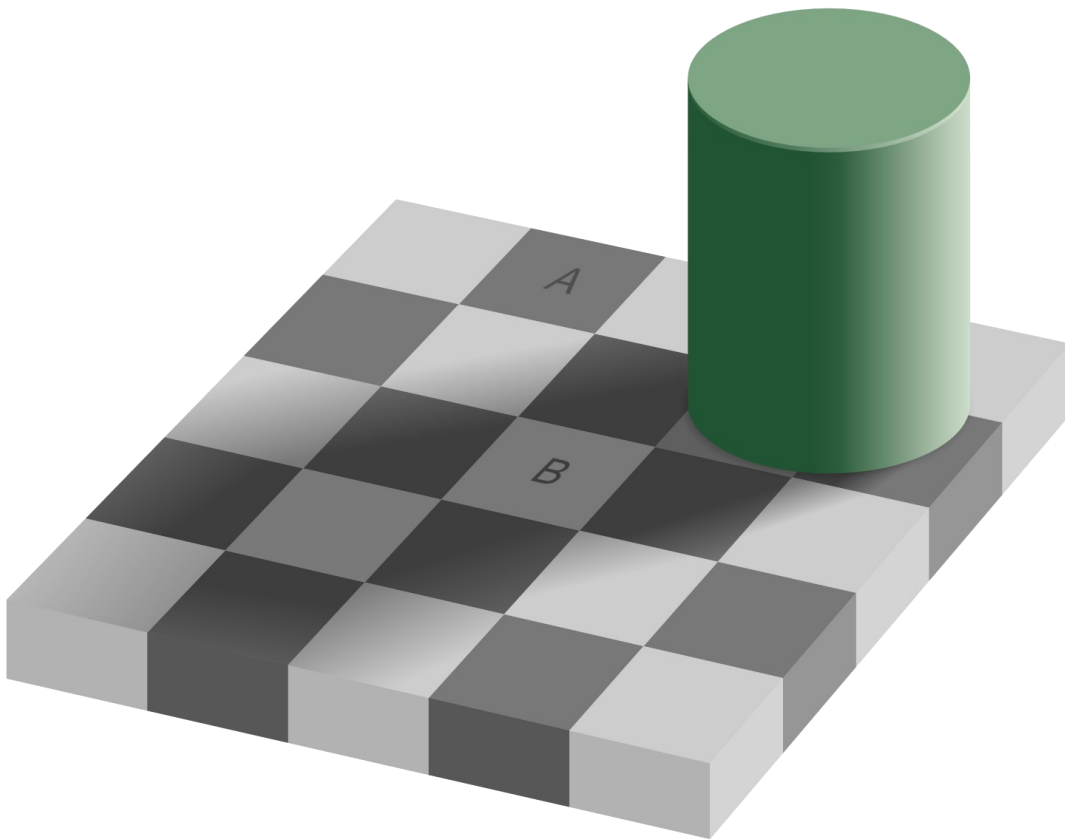
Mit der Überschrift ihres Papers „*Images are more than pictures, they are data*“ beschreiben *Gillies et al.* (2015) die Grundidee von Radiomics, einem Begriff, der sich im letzten Jahrzehnt in der quantitativen Analyse radiologischer Bilddaten etabliert hat: Durch die Umwandlung von medizinischen Bildern in quantitative Datensätze, die Informationen über die zugrundeliegende Pathophysiologie beinhalten, können diese strukturiert analysiert werden und als Entscheidungshilfe dienen (115).

Der Begriff „Radiomics“ beschreibt die quantitative Bildanalyse radiologischer Schnittbildgebungen (CT, MRT, PET/CT) auf der Theorie, dass diese Bilder nicht nur visueller Natur sind, sondern Daten enthalten, die die zugrundeliegende Pathophysiologie bis hin zur Genetik der beobachteten Struktur widerspiegeln, die dem bloßen Auge des Radiologen aber verborgen bleiben (115).

In seinem Vortrag auf dem *European Congress of Radiology* (ECR) 2019 in Wien gibt *Philippe Lambin*, Autor diverser Studien zum Thema „Radiomics“, diesbezüglich ein Beispiel für die Fehlbarkeit des menschlichen Auges. Auf dieser optischen Täuschung, die Abbildung 17 entnommen werden kann, erscheinen zwei graue Felder, hier mit „A“ und „B“ gekennzeichnet, mit einer beispielhaften Röntgendichte von je exakt 100 Hounsfield-Einheiten (HU) dem Betrachter nur durch die ihnen angrenzenden verschieden grauen Felder unterschiedlich hell, obwohl sich die Grautöne der beiden Felder von Bedeutung nicht unterscheiden (118). Da das menschliche Auge zudem weitaus weniger Grautöne unterscheiden kann, als theoretisch in der digitalen Anzeige von CT- oder MRT-Datensätzen möglich wären, ist es nur logisch, dass sich hinter den radiologischen Bildgebungen Informationen verbergen müssen, die sich nicht mit dem menschlichen Auge erfassen lassen.



**Abbildung 17:** Die Fehlbarkeit des Radiologen nach (118), übernommen aus (119)



Radiomics erweitert dabei die Idee hinter bereits vorbestehenden CAD-Systemen (*computer-aided diagnosis and detection systems*), für welche insbesondere beim Brustkrebs einige Studien existieren, die bisher jedoch nur dichotom zwischen gut- und bösartigen Tumoren unterscheiden können (115, 120, 121).

Durch statistische und maschinengestützte Analyse radiomischer Bildmarker („Radiomics Features“) und damit der Bestimmung zusätzlicher Information aus den Bilddaten sowie gegebenenfalls aus der Kombination mit klinischen Patientendaten ließen sich allerdings nicht nur diagnostische (z.B. „Ist die Läsion benigne oder maligne?“), sondern auch prädiktive (z.B. „Wirkt die Chemotherapie?“) oder prognostische (z.B. „Wie lange lebt der Patient noch?“) Modelle errechnen, die klinische Endpunkte vorhersagen und den Patienten am Ende davon profitieren lassen können. In den letzten Jahren wurden hauptsächlich Studien zum unterstützenden Potential von Radiomics in der Onkologie veröffentlicht, prinzipiell ist jedoch eine Vielzahl von Anwendungsmöglichkeiten denkbar.

Da beispielsweise in der histopathologischen Untersuchung Tumorproben mit einer

Häufigkeit von wahrscheinlich bis zu 23% im Nachhinein nicht korrekt klassifiziert werden und eine pathologische Zweitmeinung durchaus Konsequenzen für die anstehenden Therapieentscheidungen haben kann (122-125), bietet sich eine korrelierende radiomische Analyse in Zukunft auch als Ergänzung und Bestätigung der Resultate, bspw. einer Biopsie, an bzw. könnte sogar eine Alternative zu dieser darstellen und dazu beitragen, unnötige Biopsien zu reduzieren oder nur die Tumore zu biopsieren, die am wahrscheinlichsten die gesuchte Information enthalten (115).

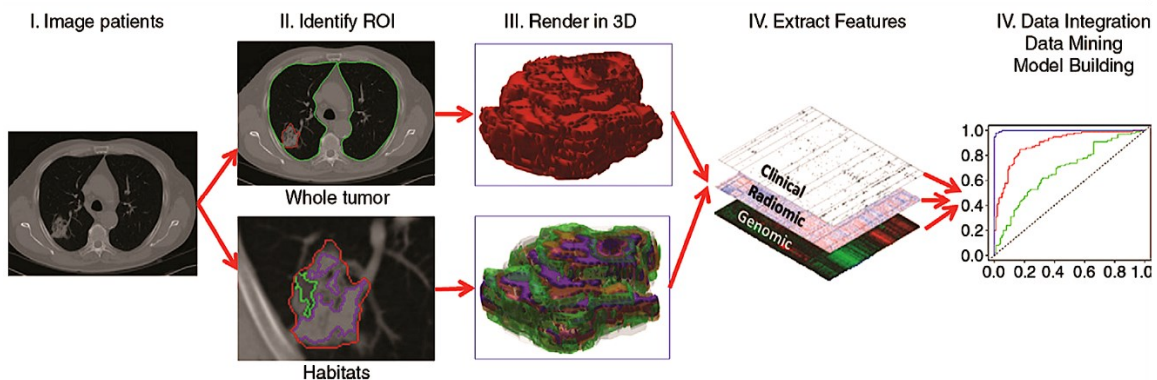
Ebenso schafft die quantitative Bilddatenanalyse eine Schnittstelle zur Genetik. So konnten Korrelationen von Radiomics Features mit dem Genexpressionsmuster untersuchter Tumorentitäten festgestellt werden (126-131). Diese These, dass aus den radiologischen Bilddaten auf das Genom des Tumors Rückschlüsse gezogen werden können, wird unter dem Begriff der „Radiogenomics“ zusammengefasst (115). Dies stellt jedoch ein eigenständiges Themengebiet dar und geht über den Rahmen dieser Arbeit hinaus.

*Gillies et al.* beschreiben in ihrer Arbeit den Arbeitsprozess der radiomischen Bilddatenanalyse in sechs Schritten. Diese können der untenstehenden Abbildung 18 entnommen werden und lauten (115):

1. Bilddatenakquise aus geeigneten Schnittbildgebungen wie der CT, MRT oder PET/CT.
2. Definieren der Regionen von Interesse (*regions of interest*, ROI) für die weitere Auswertung, z.B. Tumormetastase oder Gesamtorgane
3. Segmentieren der ROI, d.h. Definieren der Grenzen und Nachzeichnen der Konturen der auszuwertenden Bildabschnitte, ggf. mit semiautomatischer Computerunterstützung
4. Extraktion von zunächst rein deskriptiven radiomischen Bildmarkern, sog. „Radiomics Features“
5. Zusammenführen der gesammelten Features zu einer auswertbaren Datenbank
6. „Data Mining“ in Bezug auf den Findungsprozess hin zu einem diagnostischen, prädiktiven oder prognostischen Modell mithilfe von Klassifikatoren wie der „klassischen“ logistischen Regression oder „Machine Learning“-Ansätzen wie

dem Random Forest, welcher sich insbesondere bei größeren Datenmengen als Ansatz maschinellen Lernens eignet (132, 133) und gegebenenfalls in Kombination mit anderen klinischen Markern als Modell zur Unterstützung bei klinischen Entscheidungen (*decision support model*).

**Abbildung 18:** Radiomischer „Workflow“, übernommen aus (115)



Als kritischster Punkt wird dabei die Tumorsegmentierung angesehen, da bei manuellen Messungen eine deutliche Intra- und Interobserver-Variabilität besteht (134, 135) und bei vollständig automatisierter Tumorsegmentierung das nötige menschliche Fachwissen bei abweichenden Gegebenheiten fehlt. Eine optimale Lösung könnte deshalb eine Computer-unterstützte Segmentierung mit manueller Nachkorrektur durch den Radiologen darstellen, um die Untersucherabhängigkeit zu reduzieren und gleichzeitig die automatische Vorzeichnung bei Abweichungen von der Norm, die nicht vom Algorithmus erfasst wurden, korrigieren zu können (115, 136). Ein Vorteil der Tumorsegmentierung ist zudem, dass der Tumor nicht nur unidimensional wie nach den RECIST-Kriterien, sondern durch die ganzheitliche Markierung als ROI auch dreidimensional betrachtet und das Volumen in einer Tumorummetrie exakt vermessen werden kann.

In der Extraktion der radiomischen Features aus den vorhandenen ROI unterscheidet man nun unterschiedliche Klassen von Bildmarkern. Zum einen gibt es die semantischen Features, die im Allgemeinen dem Wortschatz des Radiologen entspringen und Größe, Form, Lokalisation, Nekrose usw. beinhalten. Die andere Gruppe von extrahierbaren Features sind die agnostischen Features. Diese basieren nicht auf den klinischen Bezeichnungen des Radiologen zur Befundbeschreibung,

sondern stellen rein mathematisch statistische Beschreibungen der ROI auf verschiedenen statistischen Ebenen dar. So werden die Features 1. statistischer Ordnung wie Median, Mittelwert, Minimum, Maximum, Dichte, Streuung etc., welche lediglich uniforme Voxelbeschreibungen bezeichnen, von Textur-Features 2.- und höherer statistischer Ordnung unterschieden, die die Beziehung benachbarter Voxel bis hin zu gleichartigen oder unterschiedlichen Voxelmustern über das gesamte Bild mit verschiedensten statistischen Methoden und Filtern beschreiben können. Diese quantitative Texturanalyse wurde erstmals von *Haralick et al.* 1973 eingeführt (115, 137) und findet in Radiomics eine Anwendung als Ansatz zur Beschreibung einer Tumorerheterogenität, die dem bloßen Auge des Radiologen verborgen bleibt. Diese sog. Haralick-Features haben teilweise auch heute noch Relevanz; so konnte in einer Studie mithilfe von Haralick-Features gesundes Prostatagewebe von Prostatakarzinomen unterschieden und zwischen niedrigem und hohem Gleason-Score weiter differenziert werden (138). Die heutige Zahl der teilweise auf dieser Bildanalyse basierenden Möglichkeiten zur mathematischen und statistischen Beschreibung von komplexen radiologischen Bilddaten ist jedoch noch sehr viel größer als damals; so gibt es mittlerweile weit über 1000 Radiomics Features (139), mit denen die ROI beschrieben werden kann und welche allesamt potenziell diagnostische, prädiktive oder prognostische Bildmarker darstellen könnten.

Das Themenfeld unter dem Namen „Radiomics“ wurde zuerst von *Lambin et al.* im Jahre 2012 publiziert (140) und 2017 in einem Review vom selben Autor um den Versuch erweitert, einen Qualitäts-Score mit dem Ziel der Vereinheitlichung und Vergleichbarkeit der publizierten Resultate in diesem noch jungen Forschungsfeld einzuführen, für welches derzeit noch keine einheitlichen Leitlinien existieren (141). Der eingeführte Qualitätsscore soll deshalb anhand von insgesamt maximal 36 zu vergebenden Punkten für 16 Kriterien, die die Qualität des Arbeitsablaufes über die verschiedenen Arbeitsschritte hin zu einem Radiomics-Modell beschreiben, die Qualität der publizierten Studie einheitlich vergleichend bewerten können, da die bisherigen Qualitätsstandards der publizierten Studien noch uneinheitlich und dringend verbesserungswürdig seien (141, 142). Die bisher publizierten Studienergebnisse sind jedoch vielversprechend und weisen auf ein breites zukünftiges Anwendungsspektrum von Radiomics mit hohem prognostischen Wert für viele Tumorentitäten hin (143-161).

Erwähnenswert ist in diesem Zusammenhang insbesondere eine wegweisende

Studie von *Aerts et al.* aus dem Jahre 2014, in der ein prognostisches Modell über drei multizentrische, völlig unabhängige Kohorten für zwei unterschiedliche Tumorentitäten - das Lungenkarzinom sowie Kopf- und Nackentumore - funktionierte und das Überleben der Patienten vorhersagen konnte (126).

Neuere Studien beurteilen ihre Radiomics Features außerdem zunehmend als sog. „Delta-Radiomics“, d.h. sie beurteilen auch ihre Veränderung anhand von Follow-up-CT-Untersuchungen im Therapieverlauf, basierend auf der Hypothese, dass zeitliche Veränderungen in den Feature-Werten therapiebedingt sind (162), und erzielen damit vielversprechende Ergebnisse in der Vorhersage des Therapieansprechens auf die Behandlung verschiedener Tumorentitäten (163-166).

Auch das wirtschaftliche Potential dieses noch jungen Fachbereiches ist nicht zu unterschätzen, so bieten beispielsweise die Firmen TEXRAD aus Großbritannien ([www.texrad.com](http://www.texrad.com)) und Oncoradiomics aus den Niederlanden ([www.oncoradiomics.com](http://www.oncoradiomics.com)) bereits marktreife Lösungen zur klinischen Entscheidungshilfe für die Detektion von Lungenkarzinomen an.

Ein großes Problem von Radiomics generell ist jedoch (bisher) die Reproduzierbarkeit der Ergebnisse (115). So konnten in zwei Analysen zahlreicher themenverwandter Studien nur 11-25% aller Ergebnisse reproduziert werden und eine weitere Analyse aus dem Jahre 2009 befand mindestens 50% der untersuchten Studien für qualitativ unzureichend, um von klinischem Nutzen zu sein. (115, 167-170). *Lambin* spricht diesbezüglich in seinem Review 2017 von einer „Phantom Study“ als möglichem Qualitätsmarker, in welcher zunächst anhand von „Phantomen“, Demo-Objekten definierter Dichte, vor der eigentlichen Untersuchung am Patienten diejenigen Radiomics Features vorselektiert werden, die keine signifikanten Unterschiede zwischen den verschiedenen verwendeten Untersuchungsgeräten, z.B. CT-Scannern, aufweisen. Ziel ist dabei herauszufinden, welche Radiomics Features auch gegenüber unterschiedlichen Scannern valide Ergebnisse liefern können (141, 171).

Zusammenfassend lässt sich sagen, dass Radiomics ein aufstrebendes junges Forschungsfeld mit klinischem Potential darstellt. Es bleibt jedoch für die Zukunft ein großer Forschungsbedarf, um nach den vielversprechenden Publikationen der letzten Jahre weitere valide und reproduzierbare Ergebnisse erzielen zu können, aus

welchen Systeme zur Entscheidungsunterstützung entwickelt werden können, von welchen der Patient am Ende klinisch profitieren kann. Auch bedarf es trotz ersten Versuchen der Qualitätsprüfung einer stärkeren Vereinheitlichung der Studienprotokolle, da bisher für die quantitative Bilddatenauswertung noch keine einheitlichen Standards und Leitlinien existieren. Dennoch wird das klinische Potential der quantitativen Bildanalyse aus der wachsenden Anzahl an Publikationen in den letzten Jahren deutlich, in welchen bereits gezeigt werden konnte, dass radiologische Bilddaten diagnostische, prädiktive und prognostische Informationen von der Unterscheidung des Differenzierungsgrades maligner Tumore über die Beurteilung des Therapieansprechens bis hin zur Korrelation mit dem Gesamtüberleben der Patienten enthalten, die dem bloßen Auge des Radiologen ohne maschinelle Unterstützung verborgen bleiben.

## 2 Zielsetzung der Arbeit

Da die bisher etablierte Evaluierung des Therapieansprechens mittels RECIST nur eine limitierte Anzahl Läsionen pro Organ miteinschließt, besteht die Möglichkeit, dass klinisch relevante Information über Therapieansprechen und Überleben des Patienten im Behandlungsverlauf verloren geht. Der mögliche Zusatznutzen einer radiomischen Analyse der gesamten Tumormasse eines Patienten in der Vorhersage von dessen Überleben und Therapieansprechen beim metastasierten kolorektalen Karzinom ist bislang noch nicht bekannt. Zudem stellt der Tumor Burden Score (TBS) bisher den komplexesten Ansatz dar, über die Quantifizierung der geometrischen Ausbreitung der hepatischen Tumorlast das Überleben von mKRK-Patienten vorhersagen zu wollen. Insbesondere komplexere geometrische Prädiktionsmodelle mit Klassifikatoren des maschinellen Lernens wurden hierzu bislang nicht publiziert. Deshalb soll die vorliegende Arbeit die Aussagekraft verschiedener etablierter klinischer Vorhersagemodelle und Biomarker mit neuartigen geometrischen und radiomischen Analysen der Gesamttumorlast vergleichen.

**Die Forschungshypothese lautet also:** Mit geometrischer und radiomischer Analyse der Tumorlast der gesamten Leber (*whole liver tumor burden*, WLTB) aus prätherapeutischen CT-Bildern von Patienten mit metastasiertem kolorektalen Karzinom (mKRK) lässt sich das 1-Jahres-Überleben besser vorhersagen als mit bisher etablierten Methoden wie dem Tumor Burden Score (TBS).

Eine mögliche prätherapeutische Vorhersage der Prognose könnte bedeuten, dass Patienten schneller der richtigen Therapie zugeführt werden könnten und so potenziell einen Überlebensvorteil davon haben. So könnte eine Zweitlinien- einer Erstlinientherapie vorgezogen werden, wenn man, ohne den Tumor zur Evaluation erst Wochen nach Therapiebeginn manuell vermessen zu müssen, bereits früher oder sogar noch anhand des Baseline-CT-Scans, Aussagen über die Effektivität dieser treffen könnte. Ebenso könnte man Patienten bei einer validen negativen Überlebensvorhersage die Nebenwirkungen einer intensivierten Chemotherapie bei geringen Heilungsaussichten ersparen und ein Maximum an Lebensqualität erhalten. Die frühe Vorhersage klinischer Endpunkte anhand von Bilddaten würde somit eine bessere Risikostratifizierung im Sinne des Patienten ermöglichen und die personalisierte Medizin unterstützen.

### **3 Material und Methoden**

Alle Patientendaten der vorliegenden Arbeit wurden im Rahmen des BMBF-geförderten PANTHER-Projektes (Akronym für: Patientenorientierte onkologische Therapieunterstützung) gesammelt, an welchem das Institut für klinische Radiologie und die Medizinische Klinik und Poliklinik III (Hämatologie und Onkologie) des Klinikums der Universität München zusammen mit den externen Projektpartnern Fraunhofer MEVIS (FME) und Siemens Healthineers beteiligt waren.

Die retrospektive Auswertung der Patientendaten dieser monozentrischen Studie wurde von der lokalen Ethikkommission unter der Projektnummer 502-16 bewilligt.

#### **3.1 Datensammlung**

Eingeschlossen werden konnten zunächst 485 Patienten mit metastasiertem kolorektalen Karzinom, die zwischen Januar 2007 bis Oktober 2017 im Klinikum der Universität München behandelt wurden (s. Abbildung 19). Von diesen Patienten wurden die entsprechenden CT-Datensätze gesammelt, welche vorwiegend mit den hausinternen Multidetektor-Spiral-CT-Scannern verschiedener Hersteller bei einer Standard-Aufnahmeeinstellung von 120 kV Röhrenspannung aufgenommen wurden. Die Bildaufnahme erfolgte in der portalvenösen Kontrastmittelphase nach gewichtsadaptierter intravenöser Kontrastmittelgabe und die Bildrekonstruktion erfolgte im Weichteilfenster. Die Schichtdicke variierte zwischen 0,75 und 5 mm.

#### **3.2 Segmentierung der Gesamttumorlast der Leber**

Die CT-Daten der eingeschlossenen Patienten wurden unabhängig von den zugehörigen klinischen Daten pseudonymisiert und randomisiert auf das Vorhandensein von Lebermetastasen überprüft.

Diese wurden als DICOM-Dateien aus dem radiologischen Archiv mithilfe des



radiologischen Befundungsprogramms *Syngo.via* (172) anonymisiert exportiert und in eine intern von FME auf Basis der nichtkommerziellen Software MeVisLab (173, 174) entwickelte Online-Plattform, das Segmentation Annotation Tool (SAT), hochgeladen. Im Anschluss wurden die Patientendaten (vom Verfasser dieser wissenschaftlichen Arbeit) in besagtem Programm gesichtet und zunächst bis zu 6 Lebermetastasen pro Leber mit einem Mindestdurchmesser von 10 mm (analog zu den Target-Läsions-Kriterien von RECIST) sowie die Gesamtleber halbautomatisch eingekreist, bzw. *segmentiert*, um Zielbereiche für die später anstehende radiomische Auswertung, sogenannte *regions of interest* (ROI), zu definieren. Dieser Punkt der Datenaufbereitung wird aufgrund der hohen Intra- und Interobserver-Variabilität als die kritischste Stelle im Radiomics-Workflow angesehen, da sie die Grundlage für die anstehende Entwicklung von radiomischen Prädiktionsmodellen darstellt (134, 135).

Um die Vorteile menschlicher und maschineller Segmentierungsmöglichkeiten zu kombinieren, wurde im Anschluss auf Basis der generierten ROI von FME ein auf dem Prinzip des Deep Learning basierender Algorithmus, ein sogenanntes *convolutional neural network* (175), trainiert. Mit diesem ließ sich nun auch die restliche Tumorlast der Leber vollautomatisch als ROI markieren, sodass am Ende eine semi-automatische Segmentierung der kompletten Gesamttumorlast der Leber (WLTB) stand.

Dieser Ansatz wurde gewählt, da in den letzten Jahren deutlich wurde, dass ganzheitliche Ansätze wie die Bestimmung der Gesamttumorlast der Leber geeigneter für die Überlebensvorhersage mit radiomischen Modellen und weniger untersucherabhängig sind als einzelne segmentierte Läsionen im Sinne der RECIST-Kriterien (136, 176).

Manuelle Korrekturen wie das Einzeichnen übersehener Tumore oder das Nachzeichnen von Konturen bei fehlerhafter automatischer Vorsegmentierung waren dabei allerdings zu jedem Zeitpunkt möglich.

Um die Stabilität der später extrahierten Radiomics Features zu erhöhen, wurden im Sinne einer internen Kalibrierung nach *Mühlberg et al.* (177) zudem ROI in der den Körper im CT-Scan umgebenden Luft, in Fettgewebe sowie der Trachea semi-automatisch segmentiert.

### 3.3 Vorstellung der analysierten prognostischen Modelle

Die Extraktion der Radiomics Features aus der segmentierten Gesamttumorlast erfolgte über eine intern entwickelte Software aus dem Pool der PyRadiomics-Feature-Bibliothek (139), welche aus weit über 1000 verschiedenen potenziellen Bildmarkern besteht und mit dem Ziel entwickelt wurde, als Referenz für reproduzierbare Ergebnisse in radiomischen Analysen zu dienen.

Die extrahierten Bildmarker der Gesamttumorlast wurden im Anschluss in 5 verschiedene prognostische Modelle gruppiert:

- i. Imaging priors model:* Dieses Modell vereint bereits in der Onkologie etablierte quantitative Bildmarker wie den TBS, die Lage des Primarius (*primary tumor sidedness (PTS)* “rechtsseitig oder linksseitig”), die Abschwächung des Röntgenstrahls in der CT durch das Tumorgewebe (Attenuierung, *tumor attenuation*) (178), aber auch das Volumen der Gesamttumorlast in der Leber (WLTB). Diese quantitativen Bildmarker wurden zudem auch einzeln auf ihren prognostischen Wert untersucht.
- ii. Clinical priors model:* Analog zum *Imaging priors model* vereint dieses Modell etablierte klinische Parameter, bestehend aus Laborwerten und histopathologischen Befunden. Es enthält die Parameter PTS (wie das *Imaging priors model*), Alter, Geschlecht, Tumorgrading, den Status über syn- oder metachrone Metastasierung, Histologie, den Tumormarker CEA, das UICC-Stadium, Information, ob die Metastasierung auf die Leber beschränkt ist (*liver limited disease (LLD)*) und das TNM-Staging.
- iii. Aerts radiomics prior (ARP) model:* Aufgrund seiner nachgewiesenen Vorhersagekraft für einzelne Zielläsionen in vielen onkologischen Studien wurde zudem das etablierte *Aerts radiomics model* (126, 179-181) auf seine Performance, wenn auf die WLTB angewandt, untersucht. Es besteht aus 4 radiomischen Features, die die Tumorheterogenität und die Kompaktheit der Lebermetastasen beschreiben, nämlich die *nonuniformity*, die *wavelet nonuniformity*, die *energy* und die *compactness*.
- iv. Geometric metastatic spread (GMS) model:* Aus der Hypothese, dass auch die räumliche geometrische Ausdehnung der Lebermetastasen

prognostische Bedeutung hat, erwuchs das GMS-Modell. Es besteht aus der maximalen Distanz der Lebermetastasen entlang der x-, y- und z-Achse des CT-Scanners, *metastatic spread* ( $MS_x$ ,  $MS_y$ ,  $MS_z$ ) genannt, und – in Analogie zum TBS – aus der pythagoreischen Addition dieser. Zudem wurden in das GMS-Modell zur besseren Quantifizierung der Ausbreitung 2 zusätzliche Radiomics Features implementiert, nämlich die *surface-area-to-volume ratio* ( $SA/V$ ) und die *compactness* der räumlichen Metastasenausdehnung.

- v. *Naive model*: Zu guter Letzt wurde ein Modell basierend auf allen extrahierbaren Radiomics Features inklusive der vorher genannten Modelle gebaut (alle PyRadiomics Features (ARP-Features sind enthalten) + GMS + *Imaging priors* + *Clinical priors*). Aufgrund der Vielzahl an verwendeten Features ist das Modell jedoch nach einer *maximum relevance minimum redundancy* (*mRMR*) *feature selection* (182) nur für eine Auswertung mittels Machine Learning geeignet.

Als ein weiteres Modell wurde zudem für eine Subgruppenanalyse innerhalb des *Aerts radiomics prior model* noch das *Aerts heterogeneity model* etabliert. Dieses Modell beschreibt beispielhaft anhand des ARP-Modells die alleinige Aussagekraft der Tumorheterogenität und schließt lediglich die Textur-Features aus dem ARP-Modell ein. Dadurch soll es zwar besonders unempfindlich gegenüber inakkuraten Tumorsegmentierungen sein, aber auch stark anfällig für variable CT-Scannerprotokolle, Bildrauschen und Artefakte (183-185). Um diese Anfälligkeit auszugleichen, wurde das Modell anschließend mit der internen Kalibrierungstechnik nach *Mühlberg et al.* anhand der zuvor ebenfalls segmentierten ROI der Luft, des Fettgewebes und der Trachea als Kontrollregionen auf nicht tumorspezifische Texturveränderungen kalibriert, welche bspw. den verschiedenen CT-Scannern in einem heterogenen, retrospektiv gesammelten Patientenkollektiv geschuldet sein können. Die Vorhersagekraft des Modells wurde nun vor und nach der Kalibrierung miteinander verglichen, um die Möglichkeiten einer Reduzierung der Anfälligkeit innerhalb des ARP-Modells zu untersuchen (177).

### 3.4 Datenanalyse

Ein zusammenfassendes Diagramm der durchgeführten Experimente kann der untenstehenden Abbildung A1 (s. Anhang) entnommen werden.

In der Suche um die Auswahl des besten Klassifikators für die gesuchte Vorhersage des 1-Jahres-Überlebens wurde für die 5 oben genannten Modelle, soweit möglich, jeweils ein statistischer Ansatz und ein Machine-Learning-Ansatz verfolgt. Weitere Erläuterungen finden sich dazu in den folgenden Kapiteln (3.4.1 und 3.4.2) sowie in der zugehörigen Studie von *Mühlberg et al.*, in der Teile dieser wissenschaftlichen Arbeit veröffentlicht sind (186).

#### 3.4.1 Statistik

Das 1-Jahres-Überleben und die Gesamtüberlebenszeit wurden anhand des Datums der Baseline-CT-Messung bis zum Todesdatum gemessen, sofern bekannt. In der deskriptiven Statistik wurden Mittelwerte und Standardabweichung für kontinuierliche Variablen und Häufigkeiten für kategoriale Variablen zur Zusammenfassung der verwendeten Daten und Features verwendet.

Für die univariate Statistik wurden die p-Werte mittels Student's t-Test ermittelt, sofern nach Test auf normalverteilte Daten mittels Shapiro-Wilk-Test und Test auf Varianzhomogenität mittels Levene-Test anwendbar. Alternativ wurde zur Bestimmung der p-Werte der Wilcoxon-Rangsummentest für kontinuierliche Variablen verwendet, der exakte Test nach Fisher für kategoriale Variablen mit 2 Faktoren und der Chi-Quadrat Test bei >2 Faktoren. Ein zweiseitiger p-Wert <0,05 wurde als signifikant bewertet.

Für die sich anschließende multivariate Statistik und Modellkonstruktion wurden die Features zunächst einer multiplen Testkorrektur nach der Benjamini-Hochberg-Methode zur Berechnung der Falscherkennungsrate (*false discovery rate*, FDR) unterzogen, der jeweilige C-index (s.u.) zur Beurteilung der Anpassungsgüte der Features an die Stichprobendaten bestimmt und diese dann mittels Log-likelihood-Test auf Signifikanz überprüft. Im Anschluss wurde aus den univariat signifikanten Features eine weitere Selektion der besten Features entsprechend der Güte des

bayesianischen Informationskriteriums (BIC) vorgenommen. Dieses ist im Gegensatz zum Akaike-Informationskriterium (AIC) von der Stichprobengröße abhängig und gilt bei größeren Stichproben und Modellen aus vielen Parametern als konservativerer Test durch weniger Fehler 1. Art (= fälschliche Zurückweisung der Nullhypothese) (187). Da dies zu einer spezifischeren Modellauswahl führen kann, wird das BIC hier verwendet.

Mittels logistischer Regression wurden dann multivariate Modelle konstruiert und deren Diskriminierungsfähigkeit mittels der *Area Under the Curve* (AUC) der *Receiver Operating Characteristic* (ROC, eine Kurve in einem 1-Spezifität/Sensitivität-Koordinatensystem) und den *Odds Ratios* (OR) bestimmt. Um Werteverzerrungen in den *Odds Ratios* der teilweise sehr unterschiedliche Wertemengen umfassenden Features zu verringern, wurden diese zuvor für eine bessere Vergleichbarkeit der verwendeten Features auf einen Mittelwert von 0 und eine Standardabweichung von 1 normalisiert. Als Konfidenzintervall (CI) wurde ein Wert von 95% gewählt.

Zudem wurden die Modelle neben der Diskriminationsfähigkeit des 1-Jahres-Überlebens auf ihre Prädiktionsfähigkeit von Überlebenszeiten untersucht. Zur Konstruktion prädiktiver Modelle zur Überlebenszeitenvorhersage wurden univariate und multivariate Cox Proportional Hazard (CPH) Modelle benutzt. Zur vergleichenden Überlebenszeitanalyse wurde dann zunächst auf der multivariaten Modellregression ein spezifischer medianer Schwellenwert (sog. *threshold*) bestimmt, welcher das Patientenkollektiv in zwei gleich große Hälften teilt. Entlang dieses Schwellenwertes wurde der Datensatz nun anhand der modellspezifischen Regressionswerte in eine je gleich große *high-risk*-Gruppe mit zugeordneten Modellwerten größer als der o.g. Schwellenwert und eine *low-risk*-Gruppe mit niedrigeren Werten aufgeteilt (=dichotomisiert). Die Überlebenszeiten der *high-risk*- und der *low-risk*-Gruppe müssten nun bei guter Vorhersagekraft des Modells signifikant voneinander abweichen, da im Idealfall der modellspezifische mediane Schwellenwert auch dem Median des Gesamtüberlebens entspricht, d.h. das Modell allen Patienten mit einer kürzeren Überlebenszeit auch einen niedrigeren Modell-Score zuweist und umgekehrt. Sie wurden deshalb im Anschluss vergleichend mittels Kaplan-Meier-Kurven visualisiert, mittels Log-rank-Test auf signifikante Abweichungen überprüft und der *Concordance Index* (C-index, eine Variante der AUC, geeignet für Überlebensregressionen) als Maß für die prädiktive Genauigkeit des

Modells ermittelt. Letzterer beschreibt analog zur diskriminativen AUC das Verhältnis aus erstens der Fläche zwischen der tatsächlichen Abweichung der entlang der nach dem Median des Modell-Scores dichotomisierten Kurven und zweitens der Fläche zwischen der optimalen Abweichung der Überlebenskurven bei Aufteilung der beiden Risikogruppen allein anhand des Medians der (bekannten) Überlebenszeiten, anstatt des Medians der modellspezifischen Wertemenge. Weist das Modell Patienten mit kürzerer Überlebenszeit einen falsch-hohen Score zu und umgekehrt, sinkt die Vorhersagegenauigkeit des Modells; es verringert sich der Abstand zwischen beiden Kaplan-Meier-Kurven und somit auch der C-index.

Aufgrund der Vielzahl verwendeter Features und der daraus folgenden multiplen Testkorrektur ist Modell  $v$ , das *naive model*, nicht für die univariate und multivariate statistische Analyse geeignet.

Für ein tieferes Verständnis der Zusammenhänge zwischen den klinischen Features und Imaging Features wurde eine univariate *Spearman correlation heatmap* mit Dendrogramm erstellt. Dieses kann der untenstehenden Abbildung A2 (s. Anhang) entnommen werden.

Zur weiteren Veranschaulichung wurde zudem versucht, die Stärke des Einflusses verschiedener CT-Scanner auf ausgewählte Radiomics Features über eine ANCOVA-Analyse zu schätzen. Für die Ergebnisse sei auf Tabelle A6 (s. Anhang) verwiesen.

### 3.4.2 Machine Learning

Für den Machine-Learning-Ansatz wurden 2 Methoden benutzt, um für die Vorhersage des 1-Jahres-Überlebens Trainings- und Testdaten zur internen Validierung zu gewinnen:

1. Eine 10x10-fache Kreuzvalidierung (*cross-validation* (CV)), bei welcher das Gesamtkollektiv zufällig in 10 gleich große Kohorten unterteilt wurde, die dann gegenseitig als Trainings- und Testkohorte zueinander in Beziehung gesetzt wurden. Dieser Prozess wurde anschließend 10 Male wiederholt, um die Validität des Ergebnisses auf den Testdaten weiter zu erhöhen.
2. Ein zeitlicher 2/3-Split, wonach die Patienten anhand des Datums ihres Baseline-CT-Scans in Trainings- und Testkohorte unterteilt wurden. Da in der

vorliegenden Arbeit keine externe Validierungskohorte verfügbar war, wurde diese zweite, tendenziell auf Testdaten schlechtere Ergebnisse liefernde interne Validierungsmethode zur Kontrolle mitaufgeführt, da sie einer externen Validierung durch die zeitliche Teilung in 2 unabhängige Patientenkollektive strukturell am nächsten kommt. Ein weiterer Vorteil dieser Methode besteht darin, dass sie auch auf einen möglichen zeitlichen Bias, den sogenannten Batch-Effekt, der durch die unterschiedlichen Zeitpunkte der CT-Aufnahmen im Patientenkollektiv entsteht, hinweisen kann, der durch die kreuzweise Validierung in o.g. Methode ausgehebelt wird.

Als Klassifikator zur Risikostratifizierung wurde ein Random Forest (132, 133) auf der Trainingskohorte anhand der Daten von jedem der 5 oben genannten Modelle trainiert. Die Vorhersagekraft jedes Modells wurde dann anhand der Testkohorte mit prädiktiver AUC, dem C-index und der entsprechenden Signifikanz angegeben.

Analog zum statistischen Ansatz wurden auch hier die Kohorten entlang des Medians der modellspezifischen Wertemenge jeweils zweigeteilt und die Überlebenskurven nach Training auf den Daten der Trainingskohorte aus den Daten der Testkohorte errechnet. Erneut wurde das CPH-Modell für die Berechnung der C-Indexe verwendet, Kaplan-Meier-Kurven berechnet und der Log-rank-Test zur Überprüfung signifikanter Unterschiede zwischen den Risikogruppen angewandt.

### **3.4.3 Verwendete Arbeitsprogramme**

Die hier verwendete Darstellung der Tabellen zum untersuchten Patientenkollektiv erfolgte mit dem Programm Excel<sup>®</sup> von Microsoft.

Die Datenanalyse erfolgte mit dem Programm R (Version 3.3.2, [www.R-project.org](http://www.R-project.org)) und Lifelines (188) für die Entwicklung der Prädiktionsmodelle.

## 4 Ergebnisse

### 4.1 Zusammenstellung des Patientenkollektivs

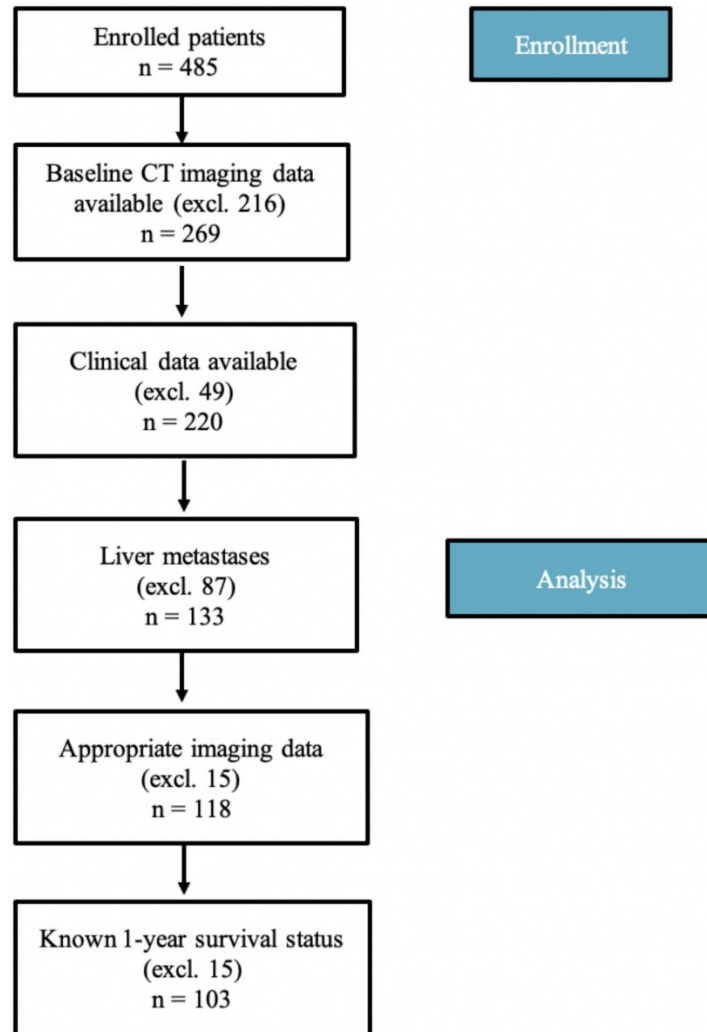
Von den 485 ursprünglich eingeschlossenen Patienten war bei 269 ein Baseline-CT-Scan verfügbar und von 220 Patienten waren ausreichend klinische Daten aus den Klinikdatenbanken extrahierbar. 133 Patienten dieser Kohorte wiesen Lebermetastasen auf. 15 Patienten mussten im Verlauf aufgrund von unzureichender Bildqualität, nativen CT-Scans ohne Kontrastmittel oder Bewegungsartefakten ausgeschlossen werden sowie 15 weitere Patienten aufgrund fehlender Information zum Überlebensstatus oder Follow-up-Information. Übrig blieb eine finale Studienkohorte, bestehend aus 103 Patienten, von welchen 82 zumindest 1 Jahr überlebten.

Die Ausschlusskriterien des zusammengestellten Patientenkollektivs können untenstehender Abbildung 19 entnommen werden.

Die CT-Daten des finalen Patientenkollektivs wurden insgesamt mit über 20 verschiedenen CT-Scannern von 5 verschiedenen Herstellern aufgenommen. Eine detaillierte Auflistung kann der folgenden Tabelle 3 entnommen werden.



**Abbildung 19:** Ausschlusskriterien des zusammengestellten Patientenkollektivs, übernommen aus (186)



Flussdiagramm der Patientenauslese inklusive der zugehörigen Ausschlusskriterien, die zum finalen Patientenkollektiv von final 103 auswertbaren mKRK-Patienten geführt haben.

**Tabelle 3:** Für die Bilddatenakquise des finalen Patientenkollektivs verwendete CT-Scanner

Marke	CT-Modell
Siemens	Biograph 64 SOMATOM Definition AS SOMATOM Definition AS+ SOMATOM Definition Edge SOMATOM Definition Flash SOMATOM Emotion 6 SOMATOM Emotion 16 SOMATOM Emotion Duo SOMATOM Force SOMATOM Sensation 16 SOMATOM Sensation 64
Philips	Brilliance 16P Brilliance 64 iCT 256 Ingenuity CT MX 16-slice
Canon Medical	Alexion Aquilion
GE Healthcare	Brightspeed S Discovery 690 Discovery CT750 HD Optima CT660
Toshiba	Asteion

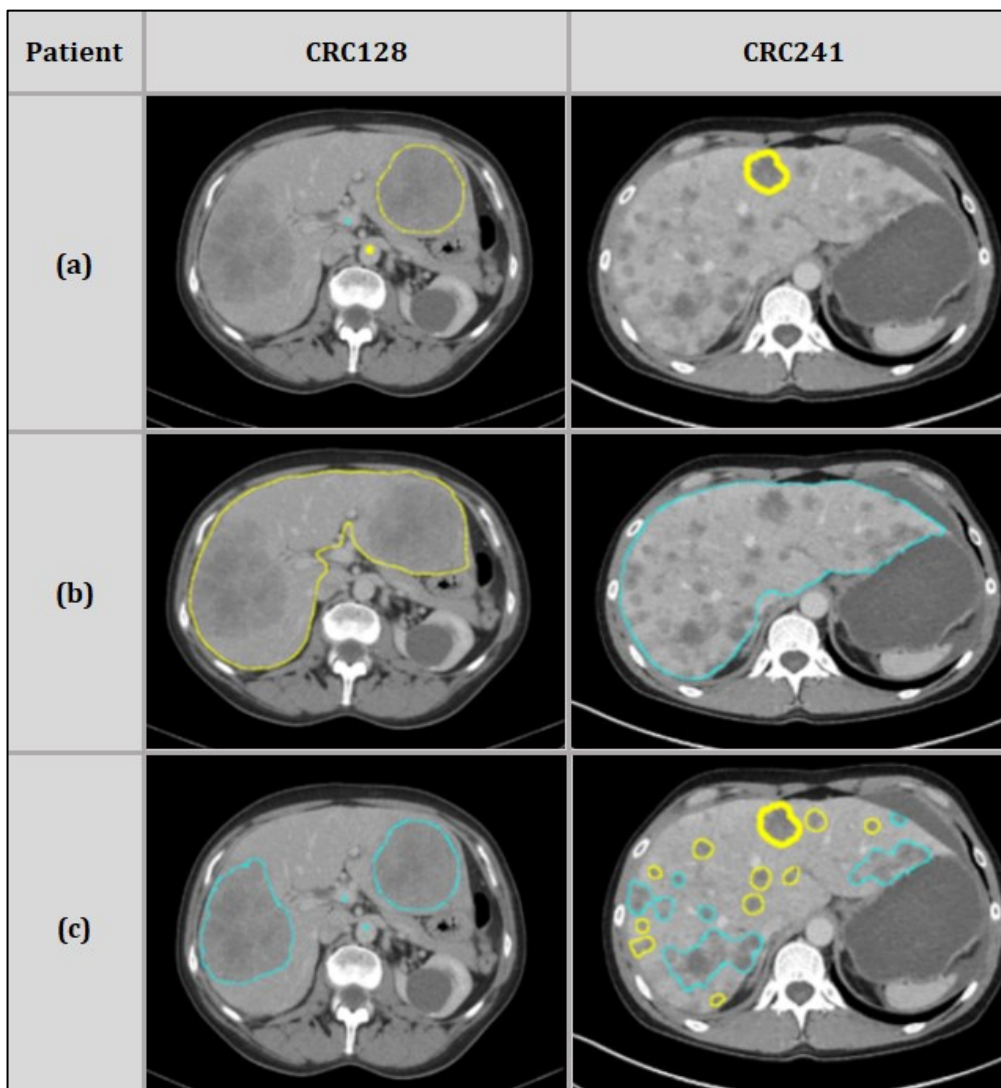
#### 4.2 Deskriptive Statistik – Basisdaten des untersuchten Patientenkollektivs

Von allen 103 zur weiteren Analyse eingeschlossenen Patienten lagen Daten zum Alter bei Erstdiagnose vor. Im Mittel waren die Patienten 61 Jahre alt. Der jüngste Patient war zum Zeitpunkt der Erstdiagnose 33 Jahre alt, der älteste 81 Jahre. Es waren 55 der 103 Patienten und damit 53,4% männlich, die anderen weiblich. Von den Patienten überlebten 82 und damit 79,6% das erste Jahr nach Diagnosestellung, 21 Patienten überlebten es nicht. Die durchschnittliche Überlebenszeit betrug 21,5 Monate. Die übrigen ausführlichen Basisdaten des untersuchten Kollektivs können der untenstehenden Tabelle A1 (s. Anhang) entnommen werden.

### 4.3 Segmentierungen der Gesamttumorlast der Leber

Die Ergebnisse der durchgeführten Tumorlast-Segmentierungen von der Einzelläsionssegmentierung über die Segmentierung der Gesamtleber bis hin zur semiautomatischen Segmentierung der Gesamttumorlast der Leber (WLTB) seien hier beispielhaft anhand folgender Abbildung 20 für zwei Patienten mit unterschiedlicher Ausprägung der hepatischen Metastasierung aufgezeigt.

**Abbildung 20:** Ergebnisse der durchgeführten Tumorsegmentierungen



Der Weg der Bilddatenannotation von der (a) Einzelläsions- über die (b) Gesamtlebersegmentierung bis hin zur semiautomatischen (c) Gesamtlebertumorlastsegmentierung, beispielhaft gezeigt anhand zweier Patienten (CRC128 und CRC241) mit hepatischer Metastasierung unterschiedlicher Ausprägung.

#### 4.4 Datenanalyse und Überlebensvorhersage

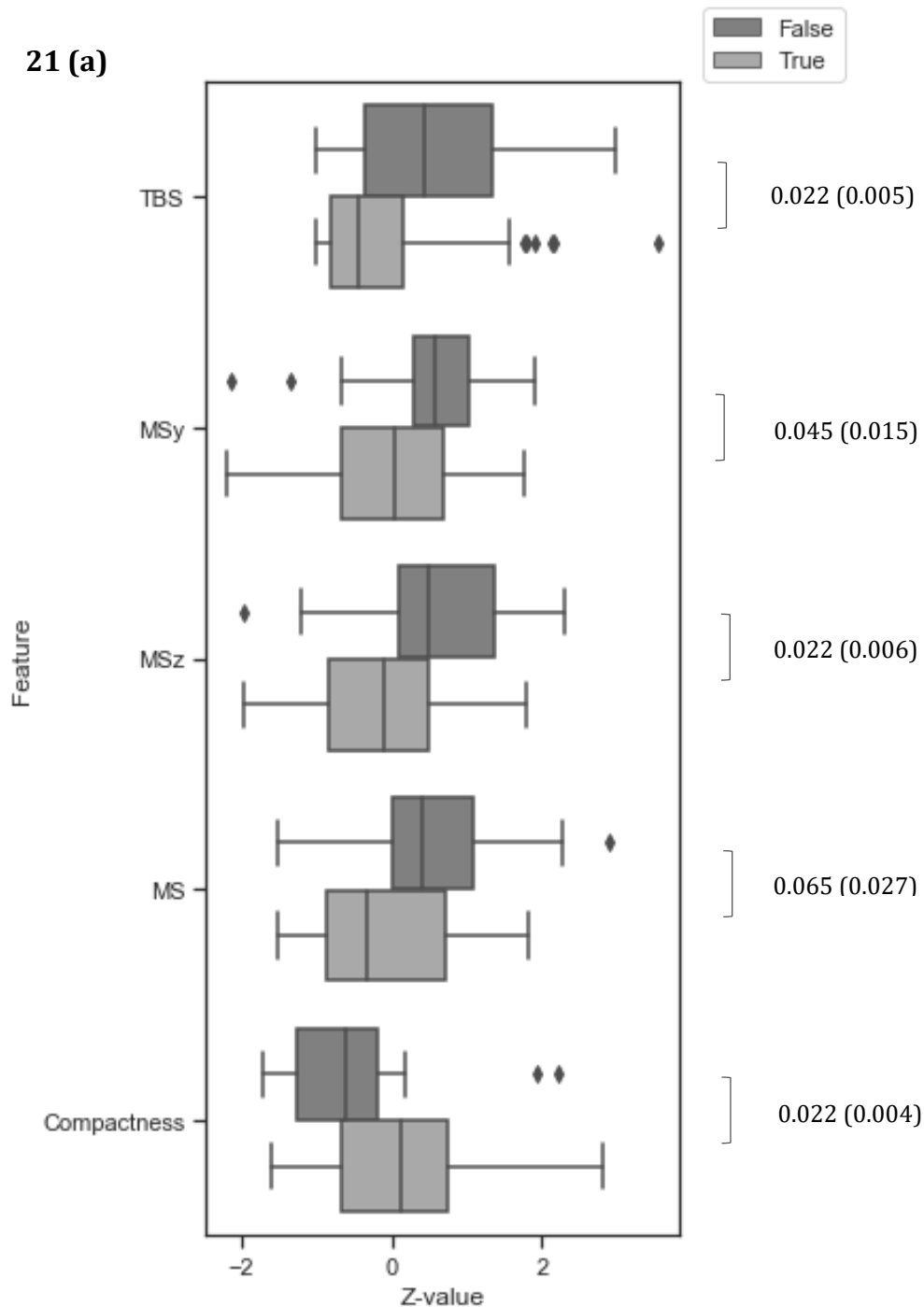
Die Ergebnisse der univariaten statistischen Analyse können Abbildung 21 (a) entnommen werden. Für eine komplette Auflistung aller durchgeführten Tests wird auf untenstehende Tabelle A2 (s. Anhang) verwiesen.

Die in den Modellen *i-v* in der Unterscheidung zwischen 1-Jahres- und Kein-Jahres-Überleben univariat signifikanten Features waren neben dem TBS die GMS-Features MS, MSy, MSz und die Compactness der Tumorausbreitung, die auch eine signifikante Anpassungsgüte und einen C-index zwischen 0,62 und 0,65 zeigten.

Aus der zum tieferen Verständnis der univariaten Zusammenhänge unter den verschiedenen klinischen Biomarkern und Imaging Biomarkern erstellten Heatmap mit Dendrogramm (Abbildung A2, s. Anhang) ging hervor, dass die größten Zusammenhänge jeweils zwischen CEA und WLTV-Volumen und zwischen MSy und dem M-Staging bestehen.

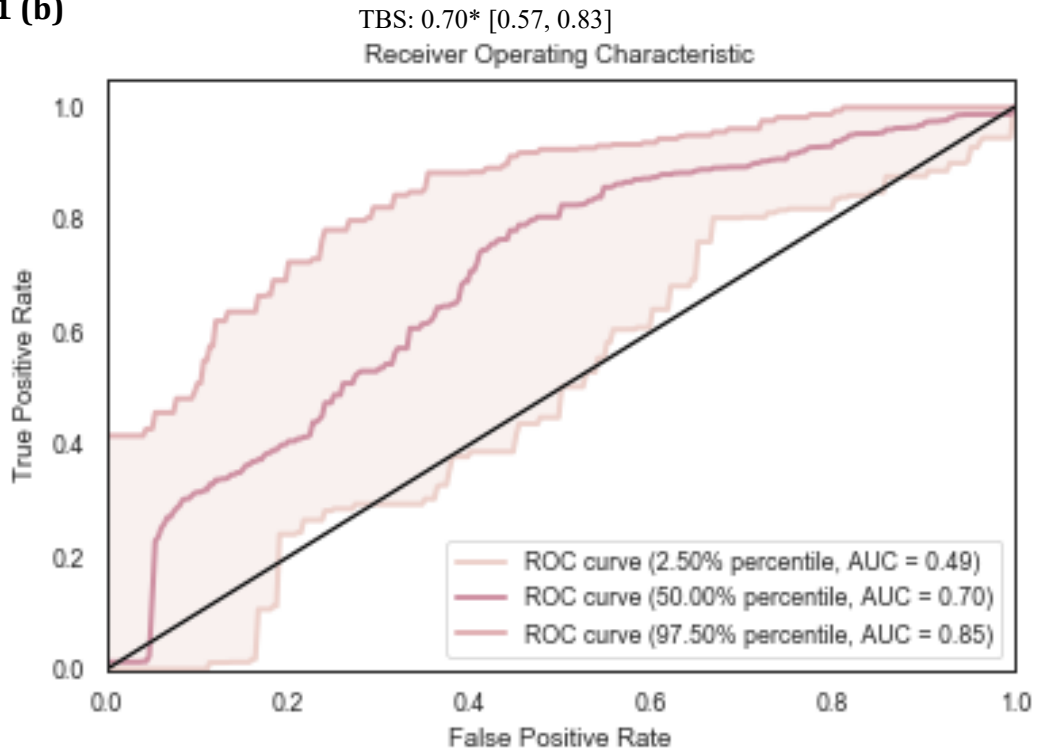
Die vollständigen Ergebnisse der sich anschließenden multivariaten Analyse können untenstehender Tabelle A3 (s. Anhang) entnommen werden. Nach dieser und nach einer Modellselektion mittels BIC blieb von den 5 in der univariaten Analyse signifikanten Features lediglich der TBS als das beste multivariate statistische Modell übrig. Dieses Ein-Feature-Modell wies mit einer AUC von 0,70 [0,57 – 0,83] eine gute Vorhersagekraft in der Vorhersage des 1-Jahres-Überlebensstatus auf (vgl. Abbildung 21 (b)) und zeigte auch in der Überlebenszeitanalyse (vgl. Abbildung 21 (c)) in der Unterscheidung zwischen Hoch- und Niedrigrisikogruppe mit einem signifikanten C-Index von 0,65 eine gute Performance.

**Abbildung 21:** Boxplots zur univariaten statistischen Analyse, ROC der 1-Jahres-Überlebensvorhersage und Kaplan-Meier-Kurve zur Überlebenszeitanalyse im multivariaten statistischen Ansatz, modifiziert nach (186)



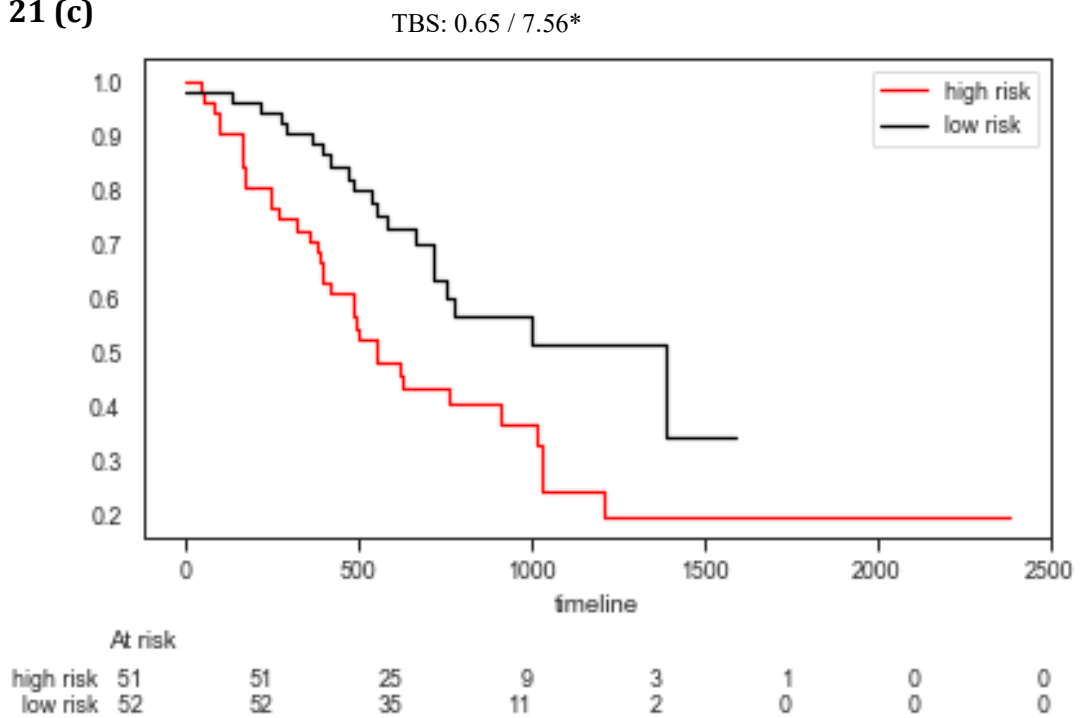
Statistische Analyse zur Abschätzung des 1-Jahres-Überlebens (1-YS) und der Gesamtüberlebenszeit. (a): In der univariaten statistischen Analyse im Log-likelihood-Test signifikante Features, dargestellt als Boxplots zur Diskriminierung des 1-Jahres-Überlebens mit zugehörigen p-Werten nach (vor) multipler Testkorrektur.

21 (b)



(b): Hier zur Veranschaulichung beispielhaft dargestellte ROC für das nach Modellselektion beste multivariate Modell, nur noch bestehend aus dem TBS. Die Ergebnisse in (b) sind dargestellt als die AUC [95%-CI] der ROC. \*: signifikant ( $p < 0.05$ ) besser als informiertes Raten.

21 (c)



(c): Kaplan-Meier-Kurve des multivariaten Modells für *high-* und *low-risk*-Gruppen. Die Ergebnisse in (c) sind dargestellt als C-index des Modells / als Ergebnis des log-rank Tests ( $p < 0.05$ : \*,  $p < 0.01$ : \*\*). TBS: Tumor burden score, MS: Metastatic spread, MSx: Metastatic spread entlang der CT-Scanner-x-Achse (etc.).

Die Ergebnisse der Machine-Learning-Analyse für die Vorhersage des 1-Jahres-Überlebens und die Überlebenszeitanalyse anhand ungesehener Testdaten finden sich in Abbildung 22 sowie in Tabelle A4, Tabelle A5 und Tabelle A6 (s. Anhang).

Zunächst wurden die QIB des *imaging priors model (i)* individuell auf ihre Vorhersagekraft analysiert. Auch hier erreichte der TBS eine gute Vorhersagekraft mit einer AUC von 0,68 [0,54 – 0,79] nach 10x10-facher Kreuzvalidierung (10x10 CV). Die AUCs nach Validierung durch zeitlichen 2/3-Split unterschieden sich in der Machine-Learning-Analyse nicht signifikant von den AUCs nach 10x10 CV und sind aufgrund fehlenden Mehrwertes hier zur besseren Übersicht und Lesbarkeit nicht mit aufgeführt (vgl. Abbildung 22 (a)). Die übrigen QIB WLTB-Volumen, die WLTB-Attenuierung und die PTS zeigten mit AUCs zwischen 0,50 und 0,57 weniger gute Ergebnisse. Das *imaging priors model (i)*, welches all diese QIB beinhaltet, wies mit einer AUC von 0,67 [0,54 - 0,79] eine ähnliche Vorhersagekraft auf wie der TBS allein.

Das *clinical priors model (ii)* zeigte eine AUC von 0,56 [0,43 – 0,67] und erreichte damit keine signifikante Vorhersagbarkeit des 1-Jahres-Überlebens, während jedoch eine zusätzlich analysierte Kombination aus sowohl *imaging priors (i)* und *clinical priors model (ii)* mit einer AUC von 0,66 [0,54 – 0,77] ähnlich gute Ergebnisse im Vergleich mit dem TBS erzielte (nicht tabellarisch aufgeführt).

Sowohl das *GMS model (iv)* als auch das *Aerts radiomics prior model (iii)* waren den anderen Modellen und dem TBS mit einer AUC von 0,73 [0,60 – 0,84] und von 0,76 [0,65 – 0,86] überlegen. Damit hat das ARP-Modell von allen präsentierten Modellen am besten abgeschnitten.

Das zur Ergänzung beispielhaft mitaufgeführte *Aerts heterogeneity model* schnitt insgesamt schlechter ab als das vollständige ARP-Modell. Nach interner Kalibrierung mit den aus den mitsegmentierten, oben aufgeführten ROI bestimmten Kalibrierungsfeatures nach (177) gelang es jedoch, die AUC des Modells nach zeitlichem 2/3-Split von anfänglich 0,67 [0,55 - 0,77] auf 0,73 [0,59 – 0,84] zu erhöhen (vgl. Tabelle A5 (s. Anhang)). Die Ergebnisse nach 10x10CV sind für dieses Modell aufgrund des hohen Rechenaufwandes bei der Komplexität der Deep-Learning-basierten Kalibrierung nicht mit aufgeführt.

Das *naive model (v)* kam in der Prädiktion des 1-Jahres-Überlebens nach 10x10 CV auf eine AUC von 0,65 [0,55 – 0,78] (vgl. Tabelle A4 (s. Anhang)) und stellt mit seinem schlechteren Abschneiden im Vergleich zu beispielsweise dem ARP-Modell die

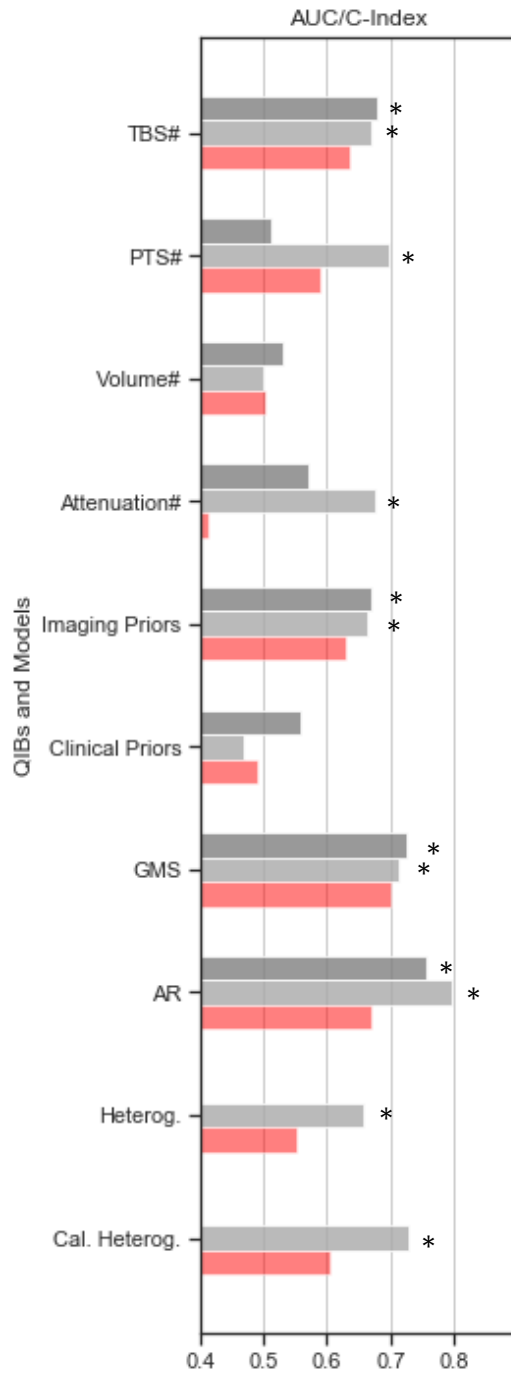
Wichtigkeit von bereits aus Studien gewonnenem, vorgegebenen Wissen in der Konstruktion von prädiktiven Modellen heraus. Eine bestätigende AUC bei temporalem 2/3-Split ist bei fehlendem Mehrwert hier deshalb nicht mehr mit aufgeführt.

Die für die Überlebenszeitanalyse erforderlichen Kaplan-Meier-Kurven zur Visualisierung der Vorhersagekraft der mit einem Random Forest als Machine-Learning-Klassifikator getesteten Modelle auf ungesehenen Testdaten sind in Abbildung 22 (b) zusammengestellt. Der C-index war auch hier für das *GMS model* mit 0,66 und für das *ARP model* mit 0,70 am höchsten und lag über dem TBS, welcher einen C-index von 0,64 aufwies.



**Abbildung 22:** Grafiken zur 1-Jahres-Überlebensvorhersage und zur Überlebenszeitanalyse im Machine-Learning-Ansatz, modifiziert nach (186)

**22 (a)**

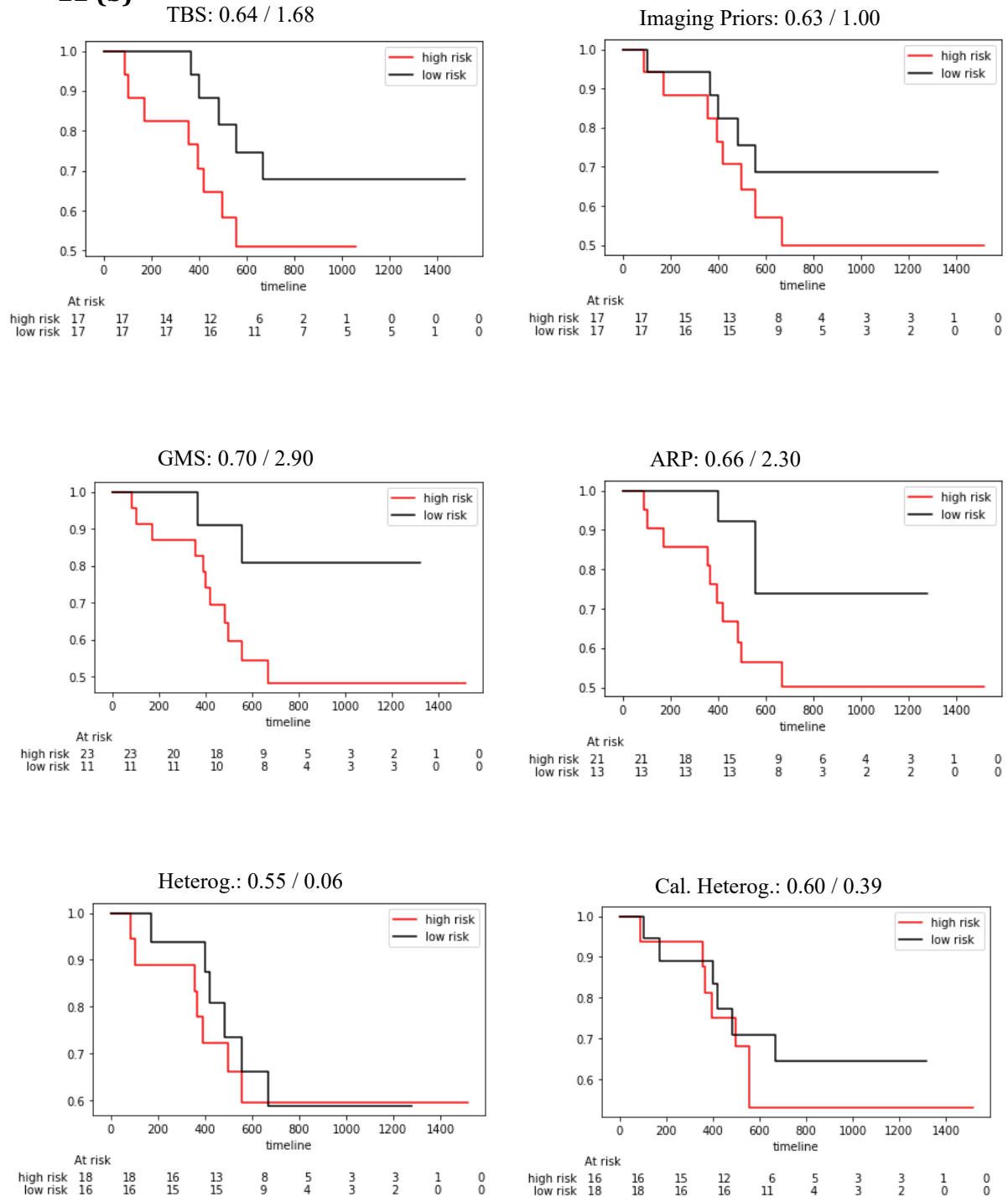


Machine-Learning-Analysen zur Prädiktion des 1-Jahres-Überlebens (1-YS) und der Gesamtüberlebenszeit für ungesehene Patienten.

(a): AUC bei 10x10-CV, sofern möglich (dunkelgrau), AUC (hellgrau) und C-index (rot) jeweils bei temporalem 2/3-Split. Zur besseren Anschaulichkeit wurde hier im Gegensatz zur statistischen Analyse (vgl. Abbildung 21 (b)) auf eine ausführliche Darstellung der zugrundeliegenden ROCs verzichtet und lediglich die AUCs der Modelle im Vergleich als Balkendiagramm dargestellt. \*: signifikant ( $p < 0.05$ ) besser als informiertes Raten.

## Ergebnisse

22 (b)



(b): Prädiktive Modellperformance anhand ausgewählter Kaplan-Meier-Kurven für ungesehene Patienten. Die Ergebnisse in (b) sind dargestellt als C-index des Modells bei temporalem 2/3-Split / als Ergebnis des log-rank Tests ( $p < 0.05$ : \*,  $p < 0.01$ : \*\*). #: 1-YS-Prädiktion basierend auf einem Modell logistischer Regression. QIB: Quantitative Imaging Biomarker, PTS: Primary tumor sidedness, TBS: Tumor burden score, GMS: Geometric metastatic spread model, ARP: Aerts radiomics prior model. Heterog.: Aerts heterogeneity model vor interner Kalibrierung, Cal. Heterog.: o.g. Modell nach interner Kalibrierung.

## 5 Diskussion

In der vorliegenden Arbeit wurde untersucht, ob eine radiomische und geometrische Analyse der Gesamttumorlast der Leber (WLTB), welche aus kontrastmittelverstärkten Baseline-CT-Daten vor Therapiebeginn vermessen wurde, zu einem neuen prognostischen Biomarker für die Vorhersage des 1-Jahres-Überlebens von Patienten mit metastasiertem kolorektalen Karzinom (mKRK) führt. Dafür wurden etablierte und neuartige Ansätze in 5 verschiedenen Modellen gruppiert:

- i. *Imaging prior model*
- ii. *Clinical priors model*
- iii. *Aerts radiomics prior (ARP) model*
- iv. *Geometric metastatic spread (GMS) model*
- v. *Naive model*

Es wurden final 103 mKRK-Patienten retrospektiv eingeschlossen, die geeignete CT-Daten aufwiesen und einer First-Line-Therapie zugeführt werden sollten.

Die Ergebnisse der Datenauswertung zeigten, dass das auf die WLTB angewandte GMS- und das ARP-Modell präzisere quantitative Bildmarker in der Überlebensvorhersage darstellten als bisherige etablierte Biomarker wie beispielsweise der TBS oder das Volumen der Gesamttumorlast der Leber (WLTB-Volumen) alleine.

Derzeit existieren vor allem Studien zur Gewinnung von Informationen zur Tumorerheterogenität mittels radiomischer Analyse für Zielläsionen (*target lesions*) diverser Tumorentitäten. In einer vorangegangenen Studie eigneten sich Radiomics beispielsweise erfolgreich für die Vorhersage des 1- und 2-Jahres-Überlebens von Patienten mit Lungenkrebs aus Baseline-CT-Daten (189). Ein ähnlicher Ansatz wurde auch verfolgt, um das Überleben und Therapieansprechen für Patienten mit Malignomen im Nasen-Rachen-Bereich oder Patienten mit Glioblastom anhand von MRT-Daten vorherzusagen (190, 191).

Es gibt allerdings auch im Falle des kolorektalen Karzinoms bereits einige Studien, die das Potential von Radiomics und den bisher noch nicht erschlossenen Informationsgehalt in CT-Daten verdeutlichen. So konnte eine Studie zeigen, dass anhand der radiomischen Tumorsignatur in der CT zwischen UICC-Tumorstadien I/II und III/IV unterschieden werden kann (149) und eine weitere, in welcher ein Nomogramm klinisch erfolgreich getestet werden konnte, das in Kombination aus radiomischer Signatur, Lymphknotenstatus in der CT und CEA-Wert des Tumors fragliche Lymphknotenmetastasen vorhersagen kann, welche gemäß den UICC-Kriterien den Übergang von Stadium II zu III darstellen (150) In der radiomischen Signatur von Tumoren steckt also ebenso Information über die Wahrscheinlichkeit, zu metastasieren. Das ist nicht nur von Bedeutung für die Wahl der Tumorthherapie, also ob primär kurative Chirurgie oder palliative Chemotherapie, sondern auch insbesondere für die Prognose der Patienten, die sich von Tumorstadium zu Tumorstadium verschlechtert (87).

Sowohl im Primarius des KRK als auch in dessen Metastasen steckt prognostische Information. So konnten in einer Studie von *Ng et al.* bestimmte CT-Textur-Features des Primarius von KRK-Patienten mit einem geringeren 5-Jahres-Überleben in Verbindung gebracht werden (151), ebenso konnte in kolorektalen Lebermetastasen die Korrelation von radiomischer Signatur mit Differenzierungsgrad und Gesamtüberleben gezeigt werden (152).

Es lassen sich aber auch Rückschlüsse auf genetische Mutationen der Tumore und sogar a priori auf das Therapieansprechen anhand radiomischer Bildmarker ziehen. So konnten *Yang et al.* in ihrer Arbeit RAS/BRAF-Mutationen anhand der radiomischen Signatur des Primarius bei KRK-Patienten vorhersagen (153). *Ahn et al.* konnten außerdem bei hepatisch metastasierten KRK-Patienten, die mit den gängigen Schemata FOLFOX und FOLFIRI therapiert wurden, anhand der Baseline-CT-Daten das Therapieansprechen unabhängig von den vier in der Studie verwendeten CT-Scannern vorhersagen (154).

Ob sich auch zukünftige maligne Veränderungen in bisher gesundem Gewebe durch quantitative Bilddatenanalyse vorhersagen lassen, ist umstritten; hier werden die Grenzen von Radiomics deutlich. So konnten *Becker et al.* zwar zeigen, dass sich bei radiomischer Texturanalyse der gesamten Leber aus CT-Daten von bisher nicht hepatisch metastasierten Patienten eine frühe hepatische Fernmetastasierung innerhalb der nächsten 6 Monate vorhersagen lässt (155). In einer weiteren Studie

von *Lee et al.* fand sich jedoch kein signifikanter Zusammenhang zwischen Gesamtlebertextur und der späteren Entwicklung dortiger Metastasen (156), ebenso konnten *Rao et al.* keinen signifikanten Unterschied in der Texturanalyse zwischen dem gesunden Leberparenchym von KRK-Patienten mit und ohne Lebermetastasen entdecken (157). Es scheint also so, als steckt vor allem in krankem Gewebe die meiste nicht mit bloßem Auge sichtbare Information.

Parallel dazu zeigte sich in einigen Studien auch die Wichtigkeit der Morphologie der gesamten Tumormasse, wie die Tumorgöße und Anzahl der Metastasen, für einen möglichen Biomarker in der Überlebensvorhersage beim mKRK (112, 176, 192, 193). Es wurde daraufhin der TBS entwickelt, der sich als ein verlässliches Instrument in der Überlebenszeitanalyse erwies, das Überleben anhand der Größe und absoluter Anzahl der Metastasen besser als der etablierte Fong-Score vorhersagen konnte und damit die Wichtigkeit der Tumormorphologie für die Prognose des Patienten herausstellte (112).

Diese Studie beschreibt und vergleicht deshalb die Vorhersagekraft des neuartigen Modells geometrischer Metastasenausbreitung (*GMS model*) und weiterer radiomischer Analysen mit etablierten klinischen Labor- und Bildmarkern sowie insbesondere dem TBS, um auf mögliche neue quantitative Bildgebungs-Biomarker (*quantitative imaging biomarker*, QIB) zu stoßen.

Als Ergebnis ließ sich zunächst mit einem statistischen Ansatz mit logistischer Regression die gute Vorhersagekraft des TBS in der Diskrimination von Patienten anhand des 1-Jahres-Überlebens sowie in der Überlebenszeitanalyse reproduzieren.

Das im Folgenden getestete *GMS-Modell* (iv) sowie das *ARP-Modell* (iii) auf Basis eines *Machine-Learning-Ansatzes* mit einem *Random Forest* als Klassifikator konnte den TBS in der Vorhersage des 1-Jahres-Überlebens anhand ungesehener Daten jedoch übertreffen.

Die aufgestellte Forschungshypothese dieser Studie kann somit als zutreffend beantwortet werden.

Das Abschneiden der beiden erfolgreichsten getesteten Modelle der vorliegenden Arbeit erscheint plausibel, da zum einen das *ARP-Modell* bereits in mehreren Studien sein hohes prädiktives Potential über verschiedene Tumorentitäten hinweg beweisen konnte (126, 139, 180, 181) und zum anderen die räumliche Ausbreitung der Lebermetastasen, welche u.a. in das *GMS-Modell* mit einfließt, von

diagnostischer Bedeutung hinsichtlich der möglicherweise dadurch beeinflussten Resektabilität von Lebermetastasen sein könnte.

Des Weiteren schnitten alle auf Bildgebungs-Biomarkern basierenden Modelle besser ab als das *clinical priors model (ii)*, welches auf klinischen Baseline-Parametern mit Prognoserelevanz wie beispielsweise Laborwerten, Mutationsstatus und der ursprünglichen Lage des Primarius (*primary tumor sidedness*) (194, 195) aufbaut.

Auch erbrachten die klinischen Bildmarker in Kombination mit den im *naive model (v)* verwendeten QIB keinen Mehrwert im Gegensatz zu denen in einer Studie von *Kim et al.*, die an Patienten mit hepatozellulärem Karzinom durchgeführt wurde (196).

In der Machine-Learning-Analyse brachte zwar keines der Modelle in den Kaplan-Meier-Kurven eine ausreichende Trennschärfe für ein signifikantes Diskriminieren zwischen Niedrigrisiko- und Hochrisikopatienten auf, die entsprechend der Teststatistik größte Zunahme der Trennschärfe hatten jedoch ebenfalls das ARP- und das GMS-Modell.

All dies verdeutlicht, dass die Informationen, die die onkologische Bildgebung liefert, mitunter stärker mit der Prognose des Patienten assoziiert sein könnten als klinische Biomarker. Da die traditionellen Prognose-Scores, wie zum Beispiel der Fong-Score, derzeit allerdings vor allem auf klinischen und pathologischen Parametern des Primarius und der Metastasen von kolorektalen Karzinomen aufbauen, unterstreicht dies das klinische Potential neuartiger Bildgebungs-basierter Modelle und Biomarker für die Prognose-Vorhersage.

Das WLTV-Volumen alleine war weder im statistischen noch im Machine-Learning-Ansatz ein guter Bildmarker. Das heißt auch, dass die Tumorprognose eventuell unabhängig vom Gesamttumorvolumen alleine ist, da die Anzahl und räumliche Ausdehnung der Metastasen eine größere Rolle zu spielen scheint als das Tumolvolumen an sich. Dies könnte durch eine schwierigere Ausgangslage für eine chirurgische Metastasenexzision begründet sein sowie durch mehr Komplikationen durch multilokales Wachstum, die Infiltration wichtiger Strukturen oder ein möglicherweise gemischtes Ansprechen bei mehreren Metastasen, sodass ein Teil der Läsionen zwar auf die Chemotherapie anspricht, die übrigen jedoch weiter größtenprogreredient sind.

Die erfolgreiche Unterscheidung zwischen 1-Jahres-Überlebenden und Nicht-Überlebenden in dieser Studie könnte zudem darauf hindeuten, dass mithilfe von Radiomics auch beim Darmkrebs zwischen aggressiv wachsenden und weniger invasiven Tumoren unterschieden werden kann, ähnlich wie beispielsweise für das Glioblastoma multiforme bereits nachgewiesen wurde (197).

Unter Betrachtung aller beeinflussenden Faktoren wie der Variabilität durch verschiedene Eigenschaften unterschiedlicher CT-Scanner, der praktischen Interpretierbarkeit des Modells und der Modellstärke über alle verwendeten statistischen und Machine-Learning-Ansätze hinweg, scheint das GMS-Modell das vielversprechendste unter allen getesteten Modellen zu sein, da geometrische Messungen durch die Verwendung verschiedener CT-Scanner im Gegensatz zu den im ARP-Modell verwendeten höhergradigen radiomischen Textur-Features nicht beeinflusst werden.

Das ARP-Modell könnte dahingehend unter dem Einfluss unterschiedlicher CT-Scanner auf seine Textur-Features abweichende Ergebnisse erzielen und leidet somit unter einer potenziell geringeren Validität (vgl. Tabelle A6 (s. Anhang)), obwohl es bereits vielversprechende Ansätze zur internen Kalibrierung an die individuellen Gegebenheiten gibt.

Dem Problem der Reproduzierbarkeit insbesondere von Textur-Features, bedingt durch unterschiedliche CT-Scanner, Spannungen, Ströme, Protokolle und Schichtdicken, wurde in dieser Studie versucht, mit der Technom-Methode nach *Mühlberg et al.* durch interne Kalibrierung zu begegnen, welche in der Originalpublikation die Inter-Scanner-Variabilität um mehr als 90% reduzieren konnte (177). Das in dieser Arbeit beispielhaft angeführte, aus den Textur-Features des ARP-Modells bestehende und infolgedessen *Aerts heterogeneity model* getaufte Modell, enthält als rein radiomisches Texturmodell die potenziell am stärksten von unterschiedlichen CT-Scannerprotokollen beeinflussbaren Features und ließ sich in seiner Performance signifikant durch diese interne Kalibrierung verbessern (vgl. Tabelle A5 (s. Anhang)).

Das GMS-Modell zeigte auch ohne eine nötige Kalibrierung eine gute prädiktive Stärke, ebenso war die räumliche Metastasen-Ausbreitung im statistischen Ansatz größtenteils signifikant. Vereinfacht gesagt erweitert das GMS-Modell damit den TBS um eine dritte Dimension und integriert auch feinere, nicht lineare

Messmethoden wie die *surface-area-to-volume ratio* ( $SA/V$ ) und die *compactness* zur Beschreibung der Metastasenverteilung und der Tumormorphologie. Es versteht sich deshalb als eine Machine-Learning-Erweiterung des TBS analog zur Erweiterung der Möglichkeiten zur Beschreibung der Tumorheterogenität durch Radiomics nach *Aerts et al.* (126).

Zur Einordnung der Qualität der vorliegenden Arbeit ist auch der 2017 von *Lambin et al.* entworfene radiomische Qualitätsscore (*radiomic quality score*, RQS) (141) hilfreich. Die eigenen radiomischen Ansätze erzielten in diesem Score ein RQS-Ergebnis von 19/36 Punkten, dabei gab es den höchsten Abzug von 7 Punkten in dieser Studie für den fehlenden Punkt der Prospektivität. Das Testergebnis entspricht einem Wert von 52,7% der Gesamtpunktzahl (mit Prospektivität wären es 72,2%); damit ist es besser als das Durchschnittsergebnis vieler weiterer radiomischer Publikationen in der Onkologie, deren RQS-Ergebnis laut einem systematischen Review von *Park et al.* aus dem Jahre 2020 bei nur 9,4/36 Punkten, also 26,1%, liegt (198).

## 5.1 Stärken und Limitationen der Studie

Diese Studie ist die erste, in welcher das neuartige GMS-Modell sowie das ARP-Modell (welches eigentlich auf Patientendaten mit Lungenkrebs und Kopf-Hals-Tumoren trainiert wurde) auf die Gesamttumorlast der Leber (WLTB) von Patienten mit Lebermetastasen eines kolorektalen Karzinoms angewandt wurde.

Die Studie konnte zeigen, dass diese beiden Modelle dem etablierten TBS in der Vorhersage des 1-Jahres-Überlebens überlegen sind und damit die Forschungshypothese bestätigen.

Die Vorhersagekraft des ARP-Modells ließ sich anhand des mKRRK somit auch mit einer anderen Tumorentität, als ursprünglich angedacht, reproduzieren und das GMS-Modell zeigte sich ebenso als vorhersagekräftiges neues Modell, welches vielversprechende Ergebnisse auch für andere Tumorentitäten mit häufiger hepatischer Metastasierung zeigen könnte, wie beispielsweise dem Pankreaskarzinom.



In der vorliegenden Arbeit werden als ROI (*regions of interest*) Lebermetastasen von KRK-Patienten herangezogen, da sie zum einen aufgrund ihrer exponierten Lage und der Organgröße bestens geeignet für eine semiautomatische Segmentierung sind und zum anderen die Leber der häufigste Metastasierungsort des KRK ist (45). Dies macht sie hinsichtlich des Therapieansprechens am ehesten repräsentativ für die Gesamttumorlast im Körper und lässt sich so gegebenenfalls auch auf andere kleine oder sich der Bildgebung entziehende Metastasen übertragen, also am besten generalisieren.

Eine weitere Stärke der Studie ist die Entscheidung für WLTB-Segmentierungen, da diese nach aktueller Studienlage repräsentativer sind, somit bessere Ergebnisse in der Prädiktion erzielen können (176) und kein *selection bias* wie bei manuell eingezeichneten ROI entsteht, die eine größere Inter-Observer-Variabilität aufweisen (136). Die Ergebnisse dieser Arbeit bestätigen die oben zitierten Vorarbeiten, sodass ganzheitliche Ansätze wie die automatische Segmentierung der WLTB als effektiver angesehen werden können als die manuelle Segmentierung einzelner Läsionen (*single lesion assessments*).

Zudem wurde durch eine vielschichtige statistische Analyse inklusive multipler Testkorrektur sowie durch verschiedene interne Validierungsverfahren im Machine-Learning-Ansatz darauf geachtet, dass die dargestellten Ergebnisse eine größtmögliche Validität aufweisen.

Dennoch hat die Aussagekraft dieser Studie Grenzen.

Da es sich um eine retrospektive Studie handelt, hat dies im gesammelten Kollektiv eine heterogene Datenlage sowie Datenlücken zur Folge, die sich nicht nachträglich schließen lassen.

Eine andere Beschränkung der Studie sind die vergleichsweise geringen Patientenzahlen, da aufgrund mangelhafter Bilddaten sowie mangelnder Verfügbarkeit klinischer Daten immer wieder Patienten ausgeschlossen werden mussten, was mit 103 Patienten zu einem relativ kleinen Kollektiv führt. Dieses kann insbesondere bei Klassifikatoren, die auf maschinellem Lernen basieren, zum Problem werden, da ein solcher Klassifikator wie bspw. der Random Forest nur bei großen Datenmengen besser funktioniert als die logistische Regression der konventionellen Statistik (132, 133) und das Kollektiv dadurch für ein valides Machine Learning eventuell nicht groß genug sein könnte.

Ein weiteres aussagelimitierendes Problem für die Arbeit ergibt sich aus der fehlenden externen Validierungskohorte. Obwohl verschiedene interne Validierungsmethoden zum Vergleich angewandt wurden, da allgemein je nach Fragestellung die eine oder andere Methode zum Trainieren des Machine-Learning-Modells bessere Ergebnisse bietet (199), fehlt ein komplett unabhängiges externes Kontrollkollektiv. Zudem unterscheiden sich Ergebnisse, obwohl ohne Auswirkung auf das ordinale Abschneiden der Modelle, teilweise signifikant je nach gewählter interner Validierungsmethode – also je nach zeitlichem 2/3 Split oder 10x10 CV – und verdeutlichen so die Notwendigkeit prospektiver Folgestudien mit externer Validierung der Ergebnisse. Dennoch stellen die beiden verwendeten internen Validierungen etablierte Trainingsmethoden im Machine Learning dar (199) und liefern auch im Vergleich mit dem Training anhand externer Validierungskohorten ähnlich stabile Ergebnisse (200).

## 5.2 Ausblick

Eine hochautomatisierte radiomische Analyse der WLTB und im Speziellen das GMS-Modell könnten zu umfassenderen QIB für die Überlebensvorhersage von mKRRK-Patienten führen als der herkömmliche TBS, sie schnitten in dieser Studie zudem besser als klinische Biomarker ab.

Obwohl die Komplexität und der Aufwand mit einer solchen klinischen WLTB-Analyse des GMS größer ist als die Berechnung des TBS oder die Vermessung einzelner Läsionen wie bspw. anhand der RECIST-Kriterien, könnte die gute prädiktive Modellstärke, die plausible Interpretierbarkeit und die voraussichtliche Robustheit gegenüber abweichenden CT-Scannern den Aufwand wert sein. Eine weiterführende Möglichkeit wäre deshalb, insbesondere das GMS-Modell in multizentrischen Studien mit anderen Tumorentitäten weitergehend zu untersuchen.

Außerdem könnte die automatische Segmentierung der WLTB dabei helfen, die Subjektivität vom bisherigen RECIST-basierten Vermessen einzelner Läsionen in möglichen weiteren Longitudinalstudien zur Bestimmung des Therapieansprechens zu überwinden. Das könnte potenziell auf lange Sicht zu einer verbesserten klinischen Entscheidungsfindung führen. Zu guter Letzt könnten ganzheitliche

Ansätze wie die Bestimmung der WLTB und stabile prädiktive Parameter wie der GMS direkt zu einem verbesserten Patienten-Management führen.

Eine mögliche Überlebensvorhersage sowie eine Vorhersage des Therapieansprechens bedeutet, dass Patienten schneller der richtigen Therapie zugeführt werden könnten und so potenziell einen Überlebensvorteil davon haben. So könnte eine Zweitlinien- einer Erstlinientherapie vorgezogen werden, wenn man, ohne den Tumor direkt vermessen zu müssen, bereits allein anhand der Tumortextur Aussagen über die Effektivität dieser treffen könnte.

Eine mögliche Folgestudie zur KRAS/CEA-Vorhersage als Endpunkt allein anhand der Bilddaten könnte Biopsien und Blutentnahmen überflüssig machen oder zumindest eine Alternative mit geringerem Risiko, z.B. Blutungsrisiko bei Biopsien, darstellen.

Neben der Möglichkeit prospektiver Folgestudien sind auch externe Validierungskohorten im Rahmen multizentrischer Studien wünschenswert, um die Validität der Ergebnisse zu verbessern, da insbesondere beim Machine Learning das „Overfitting“ und damit die fehlende Flexibilität bei Anwendung auf ungesehene Patientendaten problematisch sein kann (201).

Des Weiteren könnte auch das ARP-Modell als Ganzes von einer internen Kalibrierung wie dem in dieser Arbeit nur isoliert für dessen Textur-Features im *Aerts heterogeneity model* verwendeten Technom nach *Mühlberg et al.* profitieren und in Folgestudien noch bessere Ergebnisse über verschiedene CT-Scanner hinweg erzielen.

Ebenso könnte man versuchen, die Ergebnisse des GMS-Modells auch anhand der Lebermetastasen noch weiterer Tumorentitäten zu reproduzieren oder gar in Anlehnung an dieses auch für andere mögliche Metastasierungsorte wie die Lunge oder das Gehirn ähnliche geometrisch-orientierte radiomische Modelle entwickeln und untersuchen.

Als weitere Studie könnte zudem die Gesamttumorlast der Leber von Darmkrebspatienten neben den hier vorliegenden CT-Daten auch anhand von MRT-Daten analysiert werden, da diese laut aktueller S3-Leitlinie durch die Möglichkeit, auch in der CT okkulte Metastasen zu erkennen, für die Darstellung der Gesamttumorlast noch sensitiver als die CT ist, was im Endeffekt zu noch besseren Ergebnissen führen könnte (14, 57, 58).

Doch auch psychologische Effekte für die Patienten sind nicht zu unterschätzen. Wenn stärkere Gewissheit bestünde, wie gut die Prognose des Patienten schon vor Therapiebeginn tatsächlich ist, könnte so gegebenenfalls auf belastende Therapien verzichtet - und am Ende für die verbliebene Zeit ein Maximum an Lebensqualität erhalten werden. In jedem Falle führt eine bessere Diagnostik zu einer besseren Ausgangslage für den Patienten, sie verbessert die Therapieplanung und unter Umständen auch dessen Prognose.

## Abbildungsverzeichnis

<b>Abbildung 1:</b> Weltweite altersstandardisierte Inzidenzraten beim kolorektalen Karzinom pro 100.000 Einwohner pro Jahr, übernommen aus (2).....	10
<b>Abbildung 2:</b> Schema des Dickdarms mit Häufigkeit der Tumorlokalisationen in den Darm-Abschnitten, modifiziert nach (25).....	13
<b>Abbildung 3:</b> Prävalenz und Lebenszeitrisiko der Entwicklung eines KRK bei verschiedenen Risikogruppen.....	14
<b>Abbildung 4:</b> Die klassische Adenom-Karzinom-Sequenz nach (32), übernommen aus (33).....	15
<b>Abbildung 5:</b> Der serratierte Karzinogeneseweg, übernommen aus (37).....	16
<b>Abbildung 6:</b> Bildhafte Darstellung der KRK-Stadien nach der UICC-Klassifikation, modifiziert nach (41).....	18
<b>Abbildung 7:</b> Schema zur Koloskopie, übernommen aus (56).....	21
<b>Abbildung 8:</b> Therapieschema bei lokal fortgeschrittenem KRK, übernommen aus (31).....	25
<b>Abbildung 9:</b> Therapiestruktur im Stadium IV, übernommen aus (14).....	26
<b>Abbildung 10:</b> Auswirkungen chirurgischer und onkologischer Kriterien auf die Therapie des „fitten“ KRK-Patienten, modifiziert nach (86).....	27
<b>Abbildung 11:</b> Der Fong-Score, übernommen aus (87).....	28
<b>Abbildung 12:</b> Systemtherapie nach molekularer Einordnung des mKRK.....	30
<b>Abbildung 13:</b> Auszug ausgewählter Studien zum KRK, übernommen aus (95).....	31
<b>Abbildung 14:</b> Revidierte ESMO-Patientengruppen mit mKRK und nicht initial resektablen Metastasen, übernommen aus (86).....	33
<b>Abbildung 15:</b> ESMO-Toolbox ablativer Verfahren, übernommen aus (86).....	34

<b>Abbildung 16:</b> Der „Tumor Burden Score“, übernommen aus (112).....	38
<b>Abbildung 17:</b> Die Fehlbarkeit des Radiologen nach (118), übernommen aus (119).....	41
<b>Abbildung 18:</b> Radiomischer „Workflow“, übernommen aus (115).....	43
<b>Abbildung 19:</b> Ausschlusskriterien des zusammengestellten Patientenkollektivs, übernommen aus (186).....	57
<b>Abbildung 20:</b> Ergebnisse der durchgeführten Tumorsegmentierungen.....	59
<b>Abbildung 21:</b> Boxplots zur univariaten statistischen Analyse, ROC der 1-Jahres-Überlebensvorhersage und Kaplan-Meier-Kurve zur Überlebenszeitanalyse im multivariaten statistischen Ansatz, modifiziert nach (186).....	61
<b>Abbildung 22:</b> Grafiken zur 1-Jahres-Überlebensvorhersage und zur Überlebenszeitanalyse im Machine-Learning-Ansatz, modifiziert nach (186).....	65
<b>Abbildung A1:</b> Durchgeführter Radiomics-Workflow und Modellerstellung, übernommen aus (186).....	95
<b>Abbildung A2:</b> Zusammenhänge zwischen den klinischen Features und Imaging Features, übernommen aus (186).....	96

## Tabellenverzeichnis

<b>Tabelle 1:</b> Vereinfachte UICC-Klassifikation des kolorektalen Karzinoms und Dukes-Stadien nach (14) und (43).....	19
<b>Tabelle 2:</b> Die RECIST-Kriterien, Version 1.1, nach (97).....	36
<b>Tabelle 3:</b> Für die Bilddatenakquise des finalen Patientenkollektivs verwendete CT-Scanner.....	58
<b>Tabelle A1:</b> Basisdaten des untersuchten Patientenkollektivs.....	97
<b>Tabelle A2:</b> Ergebnisse der univariaten statistischen Analyse, übernommen aus (186).....	99
<b>Tabelle A3:</b> Multivariate Ein-Feature-Modelle, übernommen aus (186).....	101
<b>Tabelle A4:</b> Ergebnisse der Machine-Learning-Analyse, übernommen aus (186).....	102
<b>Tabelle A5:</b> Ergebnisse der Subgruppenanalyse des <i>Aerts radiomics prior model</i> (iii), modifiziert nach (186).....	104
<b>Tabelle A6:</b> Der Einfluss des CT-Scanner-Typs auf die Radiomics Features, übernommen aus (186).....	105

**Literaturverzeichnis**

1. Krebs in Deutschland für 2017/2018. 13. Ausgabe. Robert Koch-Institut (Hrsg) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg). Berlin.
2. Safiri S, Sepanlou SG, Ikuta KS, Bisignano C, Salimzadeh H, Delavari A, et al. The global, regional, and national burden of colorectal cancer and its attributable risk factors in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Gastroenterology & Hepatology*. 2019;4(12):913-33.
3. Brown WR, Ahnen DJ. The international health care burden of cancers of the gastrointestinal tract and liver. *Cancer Research Frontiers*. 2014;1(1):1-9.
4. Beltran-Sanchez H, Soneji S, Crimmins EM. Past, Present, and Future of Healthy Life Expectancy. *Cold Spring Harbor perspectives in medicine*. 2015;5(11).
5. Vaupel JW. Biodemography of human ageing. *Nature*. 2010;464(7288):536-42.
6. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut*. 2017;66(4):683-91.
7. Global Health Estimates 2016: Disease burden by Cause A, Sex, by Country and by Region, 2000-2016. Geneva, World Health Organization; 2018.
8. World Health Organization. Metrics: Disability-Adjusted Life Year (DALY) 2019 [cited 2019 Feb 25]. Available from: [https://www.who.int/healthinfo/global\\_burden\\_disease/metrics\\_daly/en/](https://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/).
9. Islami F, Miller KD, Siegel RL, Zheng Z, Zhao J, Han X, et al. National and State Estimates of Lost Earnings From Cancer Deaths in the United States. *JAMA oncology*. 2019.
10. Mariotto AB, Yabroff KR, Shao Y, Feuer EJ, Brown ML. Projections of the cost of cancer care in the United States: 2010-2020. *Journal of the National Cancer Institute*. 2011;103(2):117-28.
11. An Z. Monoclonal antibodies - a proven and rapidly expanding therapeutic modality for human diseases. *Protein & cell*. 2010;1(4):319-30.
12. Ayyar BV, Arora S, O'Kennedy R. Coming-of-Age of Antibodies in Cancer Therapeutics. *Trends in pharmacological sciences*. 2016;37(12):1009-28.
13. Buss NA, Henderson SJ, McFarlane M, Shenton JM, de Haan L. Monoclonal antibody therapeutics: history and future. *Current opinion in pharmacology*. 2012;12(5):615-22.
14. Leitlinienprogramm Onkologie. S3-Leitlinie Kolorektales Karzinom, Langversion 2.1. AWMF-Registernummer: 021/007OL. Januar 2019:p. 27-38, 198-216, 94 ff., 56, 138, 11, 37-05, 42, 11-20, 84-205, 117, 201, 163, 6.
15. Nadeem H, Jayakrishnan TT, Rajeev R, Johnston FM, Gamblin TC, Turaga KK. Cost Differential of Chemotherapy for Solid Tumors. *Journal of oncology practice*. 2016;12(3):e299-307, 251.
16. Khalil DN, Smith EL, Brentjens RJ, Wolchok JD. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. *Nature reviews*



- Clinical oncology. 2016;13(5):273-90.
17. Martin-Liberal J, Ochoa de Olza M, Hierro C, Gros A, Rodon J, Tabernero J. The expanding role of immunotherapy. *Cancer treatment reviews*. 2017;54:74-86.
  18. Geynisman DM, Chien CR, Smieliauskas F, Shen C, Shih YC. Economic evaluation of therapeutic cancer vaccines and immunotherapy: a systematic review. *Human vaccines & immunotherapeutics*. 2014;10(11):3415-24.
  19. Deutsche Krebsgesellschaft. OncoMap 2019 [cited 2019 October 22]. Available from: <https://www.oncomap.de/>.
  20. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology*. 2010;138(6):2044-58.
  21. Stoffel E, Mukherjee B, Raymond VM, Tayob N, Kastrinos F, Sparr J, et al. Calculation of risk of colorectal and endometrial cancer among patients with Lynch syndrome. *Gastroenterology*. 2009;137(5):1621-7.
  22. Herold G. Innere Medizin: Eine Vorlesungsorientierte Darstellung; Unter Berücksichtigung Des Gegenstandskataloges Für Die Ärztliche Prüfung; Mit ICD 10-Schlüssel Im Text Und Stichwortverzeichnis. Köln: Herold, 2015. 2015:488-96.
  23. Yamauchi M, Morikawa T, Kuchiba A, Imamura Y, Qian ZR, Nishihara R, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. *Gut*. 2012;61(6):847-54.
  24. Steinke VE, Christoph; Büttner, Reinhard; Schackert, Hans Konrad; Schmiegel, Wolff H.; Propping, Peter. Hereditary Nonpolyposis Colorectal Cancer (HNPCC), Lynch-Syndrom. *Dtsch Arztebl Int*. 2013;110(3):32-8.
  25. Iacopetta B. Are there two sides to colorectal cancer? *International journal of cancer*. 2002;101(5):403-8.
  26. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *The American journal of gastroenterology*. 2001;96(10):2992-3003.
  27. Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, et al. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA: a cancer journal for clinicians*. 2018;68(1):31-54.
  28. Kim S-E, Paik HY, Yoon H, Lee JE, Kim N, Sung M-K. Sex- and gender-specific disparities in colorectal cancer risk. *World journal of gastroenterology*. 2015;21(17):5167-75.
  29. Sloan JA, Goldberg RM, Sargent DJ, Vargas-Chanes D, Nair S, Cha SS, et al. Women experience greater toxicity with fluorouracil-based chemotherapy for colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2002;20(6):1491-8.
  30. Yang Y, Wang G, He J, Ren S, Wu F, Zhang J, et al. Gender differences in colorectal cancer survival: A meta-analysis. *International journal of cancer*. 2017;141(10):1942-9.
  31. Hofheinz R-D, Arnold D, Borner M, Folprecht G, Graeven U, Hebart H, et al. Kolonkarzinom, Leitlinie ICD-10 C18.- Empfehlungen der Fachgesellschaft zur Diagnostik und Therapie hämatologischer und onkologischer Erkrankungen: DGHO Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e.V.;

- Oktober 2018 [cited 2019 April 11]. Available from: <https://www.onkopedia.com/de/onkopedia/guidelines/kolonkarzinom/@@view/html/index.html>.
32. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990;61(5):759-67.
  33. Cardoso J, Boer J, Morreau H, Fodde R. Expression and genomic profiling of colorectal cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*. 2007;1775(1):103-37.
  34. Hamilton SR, Aaltonen LA, World Health Organization, Cancer IAfRo. Pathology and genetics of tumours of the digestive system. France: Lyon : Oxford : IARC Press ; Oxford University Press (distributor), 2000.
  35. Tannapfel A, Neid M, Aust D, Baretton G. Entstehung kolorektaler Karzinome. *Dtsch Arztebl International*. 2010;107(43):760-6.
  36. O'Brien MJ, Winawer SJ, Zauber AG, Gottlieb LS, Sternberg SS, Diaz B, et al. The National Polyp Study. Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. *Gastroenterology*. 1990;98(2):371-9.
  37. Yamane L, Scapulatempo-Neto C, Reis RM, Guimarães DP. Serrated pathway in colorectal carcinogenesis. *World journal of gastroenterology*. 2014;20(10):2634-40.
  38. Ijspeert JEG, Vermeulen L, Meijer GA, Dekker E. Serrated neoplasia—role in colorectal carcinogenesis and clinical implications. *Nature Reviews Gastroenterology & Hepatology*. 2015;12:401.
  39. Sebolt-Leopold JS, Herrera R. Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nature Reviews Cancer*. 2004;4:937.
  40. Bechtler MN, Eickhoff A, Riemann JF. Das Intervallkarzinom und mögliche Ursachen. *Endo heute* 2009; 22(1): 13-16.
  41. American Cancer Society. Colorectal Cancer Facts & Figures 2020-2022. Atlanta: American Cancer Society. 2020.
  42. Zlobec I, Lugli A. Prognostic and predictive factors in colorectal cancer. *Postgraduate Medical Journal*. 2008;84(994):403.
  43. Dukes-Klassifikation: DocCheck Medical Services GmbH; 2019 [cited 2019 Jun 25]. Available from: <https://flexikon.doccheck.com/de/Dukes-Klassifikation>.
  44. Hatch QM, Kniery KR, Johnson EK, Flores SA, Moeil DL, Thompson JJ, et al. Screening or Symptoms? How Do We Detect Colorectal Cancer in an Equal Access Health Care System? *Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract*. 2016;20(2):431-8.
  45. Michl M. Prognostische und prädiktive Biomarker sowie seltene Metastasierungswege beim metastasierten kolorektalen Karzinom. Habilitationsschrift, LMU München: Medizinische Fakultät 2017.
  46. Bhattacharya R, Rao S, Kowdley KV. Liver involvement in patients with solid tumors of nonhepatic origin. *Clin Liver Dis*. 2002;6(4):1033-43, x.
  47. Tischhoff I, Tannapfel A. CUP in der Leber. *Der Pathologe*. 2020;41(5):495-504.
  48. Mitry E, Guiu B, Coscinea S, Jooste V, Faivre J, Bouvier AM. Epidemiology, management and prognosis of colorectal cancer with lung metastases: a 30-year population-based study. *Gut*. 2010;59(10):1383-8.

49. Suthanathan AE, Bhandari M, Platell C. Influence of primary site on metastatic distribution and survival in stage IV colorectal cancer. *ANZ journal of surgery*. 2018;88(5):445-9.
50. Scheele J, Altendorf-Hofmann A, Stangl R, Gall FP. Pulmonary resection for metastatic colon and upper rectum cancer. Is it useful? *Diseases of the colon and rectum*. 1990;33(9):745-52.
51. Weiss L, Bronk J, Pickren JW, Lane WW. Metastatic patterns and target organ arterial blood flow. *Invasion & metastasis*. 1981;1(2):126-35.
52. Tamas K, Walenkamp AME, de Vries EGE, van Vugt MATM, Beets-Tan RG, van Etten B, et al. Rectal and colon cancer: Not just a different anatomic site. *Cancer treatment reviews*. 2015;41(8):671-9.
53. Horton KM, Abrams RA, Fishman EK. Spiral CT of colon cancer: imaging features and role in management. *Radiographics : a review publication of the Radiological Society of North America, Inc.* 2000;20(2):419-30.
54. Gemeinsamer Bundesausschuss. Der iFOBT im Darmkrebs-Screening: Ergebnisse der medizinischen Laboratorien für das Jahr 2018. 2018 [cited 2019 May 14]. Available from: <https://www.g-ba.de/richtlinien/17/>.
55. Gemeinsamer Bundesausschuss. Pressemitteilung: Darmkrebs-Screening künftig als organisiertes Programm 2018 [cited 2020 Sep 8]. Available from: <https://www.g-ba.de/presse/pressemitteilungen/756/>.
56. Prof. Dr. med. Robert Ehehalt. Was ist eine Koloskopie (Darmspiegelung)? : Deutscher Verlag für Gesundheitsinformation; 2018 [cited 2021 Aug 31]. Available from: <https://www.leading-medicine-guide.de/diagnostik/koloskopie-darmspiegelung>.
57. Floriani I, Torri V, Rulli E, Garavaglia D, Compagnoni A, Salvolini L, et al. Performance of imaging modalities in diagnosis of liver metastases from colorectal cancer: a systematic review and meta-analysis. *Journal of magnetic resonance imaging : JMRI*. 2010;31(1):19-31.
58. Niekel MC, Bipat S, Stoker J. Diagnostic imaging of colorectal liver metastases with CT, MR imaging, FDG PET, and/or FDG PET/CT: a meta-analysis of prospective studies including patients who have not previously undergone treatment. *Radiology*. 2010;257(3):674-84.
59. Quaia E, D'Onofrio M, Palumbo A, Rossi S, Bruni S, Cova M. Comparison of contrast-enhanced ultrasonography versus baseline ultrasound and contrast-enhanced computed tomography in metastatic disease of the liver: diagnostic performance and confidence. *European radiology*. 2006;16(7):1599-609.
60. Rafaelsen SR, Jakobsen A. Contrast-enhanced ultrasound vs multidetector-computed tomography for detecting liver metastases in colorectal cancer: a prospective, blinded, patient-by-patient analysis. *Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland*. 2011;13(4):420-5.
61. Schmoll HJ, Van Cutsem E, Stein A, Valentini V, Glimelius B, Haustermans K, et al. ESMO Consensus Guidelines for management of patients with colon and rectal cancer. a personalized approach to clinical decision making. *Annals of oncology : official journal of the European Society for Medical Oncology*. 2012;23(10):2479-516.
62. Pickhardt PJ, Hassan C, Laghi A, Kim DH. CT colonography to screen for

- colorectal cancer and aortic aneurysm in the Medicare population: cost-effectiveness analysis. *AJR American journal of roentgenology*. 2009;192(5):1332-40.
63. Lessler DS, Sullivan SD, Stergachis A. Cost-effectiveness of unenhanced MR imaging vs contrast-enhanced CT of the abdomen or pelvis. *AJR American journal of roentgenology*. 1994;163(1):5-9.
  64. Computertomographie (CT): Klinikum der Universität München, Klinik und Poliklinik für Radiologie; [cited 2019 May 15]. Available from: <http://www.klinikum.uni-muenchen.de/Klinik-und-Poliklinik-fuer-Radiologie/de/Leistungen/CT/index.html>.
  65. Siebeck M. Darmspiegelung (Koloskopie): Klinikum der Universität München, Klinik für Allgemeine-, Unfall- und Wiederherstellungschirurgie; [cited 2019 May 15]. Available from: <http://www.klinikum.uni-muenchen.de/Klinik-fuer-Allgemeine-Unfall-und-Wiederherstellungschirurgie/de/fach/viszeralchirurgie/proktologie/untersuchungsmethoden/koloskopie/index.html>.
  66. Magnetresonanztomographie (MRT / Kernspin): Klinikum der Universität München, Klinik und Poliklinik für Radiologie; [cited 2019 May 15]. Available from: <http://www.klinikum.uni-muenchen.de/Klinik-und-Poliklinik-fuer-Radiologie/de/Leistungen/MRT/index.html>.
  67. Glazer ES, Beaty K, Abdalla EK, Vauthey JN, Curley SA. Effectiveness of positron emission tomography for predicting chemotherapy response in colorectal cancer liver metastases. *Archives of surgery (Chicago, Ill : 1960)*. 2010;145(4):340-5; discussion 5.
  68. Wicherts DA, de Haas RJ, van Kessel CS, Bisschops RH, Takahara T, van Hillegersberg R, et al. Incremental value of arterial and equilibrium phase compared to hepatic venous phase CT in the preoperative staging of colorectal liver metastases: an evaluation with different reference standards. *European journal of radiology*. 2011;77(2):305-11.
  69. Krebsinformationsdienst, Deutsches Krebsforschungszentrum; 0800-420 30 40 oder [krebsinformationsdienst@dkfz.de](mailto:krebsinformationsdienst@dkfz.de). Computertomographie: Technik, Forschung, Fachinformationen 2019 [cited 2019 May 17]. Available from: [www.krebsinformationsdienst.de/untersuchung/computertomographie-technik.php](http://www.krebsinformationsdienst.de/untersuchung/computertomographie-technik.php).
  70. Keil S. Computertomographie: Uniklinik RWTH Aachen, Klinik für Diagnostische und Interventionelle Radiologie; 2019 [cited 2019 May 17]. Available from: <https://www.ukaachen.de/kliniken-institute/klinik-fuer-diagnostische-und-interventionelle-radiologie/fuer-patienten/computertomographie.html>.
  71. Holch JW, Ricard I, Stintzing S, Fischer von Weikersthal L, Decker T, Kiani A, et al. Relevance of baseline carcinoembryonic antigen for first-line treatment against metastatic colorectal cancer with FOLFIRI plus cetuximab or bevacizumab (FIRE-3 trial). *European journal of cancer (Oxford, England : 1990)*. 2019;106:115-25.
  72. Douillard JY, Oliner KS, Siena S, Tabernero J, Burkes R, Barugel M, et al. Panitumumab-FOLFOX4 treatment and RAS mutations in colorectal cancer. *The New England journal of medicine*. 2013;369(11):1023-34.
  73. Van Cutsem E, Lenz HJ, Kohne CH, Heinemann V, Tejpar S, Melezinek I, et al. Fluorouracil, leucovorin, and irinotecan plus cetuximab treatment and RAS mutations in colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2015;33(7):692-700.

74. Bokemeyer C, Kohne CH, Ciardiello F, Lenz HJ, Heinemann V, Klinkhardt U, et al. FOLFOX4 plus cetuximab treatment and RAS mutations in colorectal cancer. *European journal of cancer (Oxford, England : 1990)*. 2015;51(10):1243-52.
75. Hsu HC, Thiam TK, Lu YJ, Yeh CY, Tsai WS, You JF, et al. Mutations of KRAS/NRAS/BRAF predict cetuximab resistance in metastatic colorectal cancer patients. *Oncotarget*. 2016;7(16):22257-70.
76. Venderbosch S, Nagtegaal ID, Maughan TS, Smith CG, Cheadle JP, Fisher D, et al. Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: a pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2014;20(20):5322-30.
77. Lasota J, Kowalik A, Wasag B, Wang ZF, Felisiak-Golabek A, Coates T, et al. Detection of the BRAF V600E mutation in colon carcinoma: critical evaluation of the immunohistochemical approach. *The American journal of surgical pathology*. 2014;38(9):1235-41.
78. Stintzing S, Jung A, Rossius L, Modest DP, Weikersthal LFv, Decker T, et al. Mutations within the EGFR signaling pathway: Influence on efficacy in FIRE-3—A randomized phase III study of FOLFIRI plus cetuximab or bevacizumab as first-line treatment for wild-type (WT) KRAS (exon 2) metastatic colorectal cancer (mCRC) patients. *Journal of Clinical Oncology*. 2014;32(3\_suppl):445-.
79. Pietrantonio F, Petrelli F, Coinu A, Di Bartolomeo M, Borgonovo K, Maggi C, et al. Predictive role of BRAF mutations in patients with advanced colorectal cancer receiving cetuximab and panitumumab: A meta-analysis. *European Journal of Cancer*. 2015;51(5):587-94.
80. O'Connor ES, Greenblatt DY, LoConte NK, Gangnon RE, Liou JI, Heise CP, et al. Adjuvant chemotherapy for stage II colon cancer with poor prognostic features. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2011;29(25):3381-8.
81. Gray R, Barnwell J, McConkey C, Hills RK, Williams NS, Kerr DJ. Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. *Lancet (London, England)*. 2007;370(9604):2020-9.
82. Ding HH, Wu WD, Jiang T, Cao J, Ji ZY, Jin JH, et al. Meta-analysis comparing the safety and efficacy of metastatic colorectal cancer treatment regimens, capecitabine plus irinotecan (CAPIRI) and 5-fluorouracil/leucovorin plus irinotecan (FOLFIRI). *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2015;36(5):3361-9.
83. Madi A, Fisher D, Wilson RH, Adams RA, Meade AM, Kenny SL, et al. Oxaliplatin/capecitabine vs oxaliplatin/infusional 5-FU in advanced colorectal cancer: the MRC COIN trial. *British journal of cancer*. 2012;107(7):1037-43.
84. Nordlinger B, Van Cutsem E, Gruenberger T, Glimelius B, Poston G, Rougier P, et al. Combination of surgery and chemotherapy and the role of targeted agents in the treatment of patients with colorectal liver metastases: recommendations from an expert panel. *Annals of oncology : official journal of the European Society for Medical Oncology*. 2009;20(6):985-92.
85. Alberts SR. Update on the optimal management of patients with colorectal liver metastases. *Critical reviews in oncology/hematology*. 2012;84(1):59-70.
86. Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D, et al.

- ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Annals of Oncology*. 2016;27(8):1386-422.
87. Fong Y, Fortner J, Sun RL, Brennan MF, Blumgart LH. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. *Annals of surgery*. 1999;230(3):309-18; discussion 18-21.
  88. Nordlinger B, Sorbye H, Glimelius B, Poston GJ, Schlag PM, Rougier P, et al. Perioperative FOLFOX4 chemotherapy and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC 40983): long-term results of a randomised, controlled, phase 3 trial. *The Lancet Oncology*. 2013;14(12):1208-15.
  89. Heinemann V, von Weikersthal LF, Decker T, Kiani A, Vehling-Kaiser U, Al-Batran SE, et al. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial. *The Lancet Oncology*. 2014;15(10):1065-75.
  90. Fransen K, Klintenas M, Osterstrom A, Dimberg J, Monstein HJ, Soderkvist P. Mutation analysis of the BRAF, ARAF and RAF-1 genes in human colorectal adenocarcinomas. *Carcinogenesis*. 2004;25(4):527-33.
  91. Maughan TS, Adams RA, Smith CG, Meade AM, Seymour MT, Wilson RH, et al. Addition of cetuximab to oxaliplatin-based first-line combination chemotherapy for treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial. *Lancet (London, England)*. 2011;377(9783):2103-14.
  92. Tveit KM, Guren T, Glimelius B, Pfeiffer P, Sorbye H, Pyrhonen S, et al. Phase III trial of cetuximab with continuous or intermittent fluorouracil, leucovorin, and oxaliplatin (Nordic FLOX) versus FLOX alone in first-line treatment of metastatic colorectal cancer: the NORDIC-VII study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2012;30(15):1755-62.
  93. Ulivi P, Scarpi E, Chiadini E, Marisi G, Valgiusti M, Capelli L, et al. Right- vs. Left-Sided Metastatic Colorectal Cancer: Differences in Tumor Biology and Bevacizumab Efficacy. *International journal of molecular sciences*. 2017;18(6).
  94. Cremolini C, Antoniotti C, Lonardi S, Bergamo F, Cortesi E, Tomasello G, et al. Primary tumor sidedness and benefit from FOLFOXIRI plus bevacizumab as initial therapy for metastatic colorectal cancer. Retrospective analysis of the TRIBE trial by GONO. *Annals of oncology : official journal of the European Society for Medical Oncology*. 2018.
  95. Guerrero RM, Labajos VA, Ballena SL, Macha CA, Lezama MS, Roman CP, et al. Targeting BRAF V600E in metastatic colorectal cancer: where are we today? *Ecancelmedicalscience*. 2022;16:1489.
  96. Tournigand C, Cervantes A, Figer A, Lledo G, Flesch M, Buyse M, et al. OPTIMOX1: a randomized study of FOLFOX4 or FOLFOX7 with oxaliplatin in a stop-and-Go fashion in advanced colorectal cancer--a GERCOR study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2006;24(3):394-400.
  97. E.A. Eisenhauer, P. Therasse, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford J, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *EUROPEAN JOURNAL OF CANCER*. 2009;45:228 – 47.
  98. Schwartz LH, Litière S, de Vries E, Ford R, Gwyther S, Mandrekar S, et al. RECIST 1.1-Update and clarification: From the RECIST committee. *European journal of*

- cancer (Oxford, England : 1990). 2016;62:132-7.
99. Pinker K, Riedl C, Weber WA. Evaluating tumor response with FDG PET: updates on PERCIST, comparison with EORTC criteria and clues to future developments. *European journal of nuclear medicine and molecular imaging*. 2017;44(Suppl 1):55-66.
  100. Kim JH, Kim BJ, Jang HJ, Kim HS. Comparison of the RECIST and EORTC PET criteria in the tumor response assessment: a pooled analysis and review. *Cancer chemotherapy and pharmacology*. 2017;80(4):729-35.
  101. Schwartz LH, Seymour L, Litière S, Ford R, Gwyther S, Mandrekar S, et al. RECIST 1.1 - Standardisation and disease-specific adaptations: Perspectives from the RECIST Working Group. *European Journal of Cancer*. 2016;62:138-45.
  102. Rustin GJ, Quinn M, Thigpen T, du Bois A, Pujade-Lauraine E, Jakobsen A, et al. Re: New guidelines to evaluate the response to treatment in solid tumors (ovarian cancer). *Journal of the National Cancer Institute*. 2004;96(6):487-8.
  103. Bubley GJ, Carducci M, Dahut W, Dawson N, Daliani D, Eisenberger M, et al. Eligibility and response guidelines for phase II clinical trials in androgen-independent prostate cancer: recommendations from the Prostate-Specific Antigen Working Group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 1999;17(11):3461-7.
  104. Scher HI, Halabi S, Tannock I, Morris M, Sternberg CN, Carducci MA, et al. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: recommendations of the Prostate Cancer Clinical Trials Working Group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2008;26(7):1148-59.
  105. Vergote I, Rustin GJ, Eisenhauer EA, Kristensen GB, Pujade-Lauraine E, Parmar MK, et al. Re: new guidelines to evaluate the response to treatment in solid tumors [ovarian cancer]. *Gynecologic Cancer Intergroup. Journal of the National Cancer Institute*. 2000;92(18):1534-5.
  106. Taieb J, Rivera F, Siena S, Karthaus M, Valladares-Ayerbes M, Gallego J, et al. Exploratory analyses assessing the impact of early tumour shrinkage and depth of response on survival outcomes in patients with RAS wild-type metastatic colorectal cancer receiving treatment in three randomised panitumumab trials. *J Cancer Res Clin Oncol*. 2018;144(2):321-35.
  107. Chun YS, Vauthey JN, Boonsirikamchai P, Maru DM, Kopetz S, Palavecino M, et al. Association of computed tomography morphologic criteria with pathologic response and survival in patients treated with bevacizumab for colorectal liver metastases. *Jama*. 2009;302(21):2338-44.
  108. Boonsirikamchai P, Asran MA, Maru DM, Vauthey JN, Kaur H, Kopetz S, et al. CT findings of response and recurrence, independent of change in tumor size, in colorectal liver metastasis treated with bevacizumab. *AJR American journal of roentgenology*. 2011;197(6):W1060-6.
  109. Moltz JH, D'Anastasi M, Kiessling A, Pinto dos Santos D, Schulke C, Peitgen HO. Workflow-centred evaluation of an automatic lesion tracking software for chemotherapy monitoring by CT. *European radiology*. 2012;22(12):2759-67.
  110. Wulff AM, Fabel M, Freitag-Wolf S, Tepper M, Knabe HM, Schafer JP, et al. Volumetric response classification in metastatic solid tumors on MSCT: initial results in a whole-body setting. *European journal of radiology*. 2013;82(10):e567-

- 73.
111. Zinsser D, Baumann AB, Winter KS, Bamberg F, Lange P, Nikolaou K, et al. Semi-automatic CT-angiography based evaluation of the aortic annulus in patients prior to TAVR: interchangeability with manual measurements. *The international journal of cardiovascular imaging*. 2018;34(10):1657-67.
  112. Sasaki K, Morioka D, Conci S, Margonis GA, Sawada Y, Ruzzenente A, et al. The Tumor Burden Score: A New “Metro-ticket” Prognostic Tool For Colorectal Liver Metastases Based on Tumor Size and Number of Tumors. *Annals of surgery*. 2018;267(1):132-41.
  113. Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Science translational medicine*. 2011;3(75):75ra26.
  114. Burnside ES, Sickles EA, Bassett LW, Rubin DL, Lee CH, Ikeda DM, et al. The ACR BI-RADS experience: learning from history. *Journal of the American College of Radiology : JACR*. 2009;6(12):851-60.
  115. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016;278(2):563-77.
  116. Bomers JG, Barentsz JO. Standardization of multiparametric prostate MR imaging using PI-RADS. *BioMed research international*. 2014;2014:431680.
  117. Kazerooni EA, Armstrong MR, Amorosa JK, Hernandez D, Liebscher LA, Nath H, et al. ACR CT Accreditation Program and the Lung Cancer Screening Program Designation. *Journal of the American College of Radiology : JACR*. 2016;13(2 Suppl):R30-4.
  118. Lambin P. The use of radiomics in medical imaging: Speech No. A-0267, held on Feb 28 at 25th European Congress of Radiology (ECR) in Vienna, 2019; 2019 [cited 2019 Aug 8]. Available from: <https://ecronline.myesr.org/ecr2019/index.php?p=recording&t=recorded&lecture=the-use-of-radiomics-in-medical-imaging>.
  119. Optical illusion with checkerboard [cited 2020 July 4]. Vector image of 3d checkerboard and a cylinder. Available from: <https://publicdomainvectors.org/en/free-clipart/Optical-illusion-with-checkerboard/33938.html>.
  120. Giger ML. Update on the potential of computer-aided diagnosis for breast cancer. *Future Oncology*. 2010;6(1):1-4.
  121. Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*. 2007;31(4-5):198-211.
  122. Clauson J, Hsieh Y, Acharya S, Rademaker AW, Morrow M. Results of the Lynn Sage Second - Opinion Program for local therapy in patients with breast carcinoma: Changes in management and determinants of where care is delivered. *Cancer*. 2002;94(4):889-94.
  123. Nguyen PL, Schultz D, Renshaw AA, Vollmer RT, Welch WR, Cote K, et al., editors. The impact of pathology review on treatment recommendations for patients with adenocarcinoma of the prostate. *Urologic Oncology: Seminars and Original Investigations*; 2004: Elsevier.
  124. Staradub VL, Messenger KA, Hao N, Wiely EL, Morrow M. Changes in breast



- cancer therapy because of pathology second opinions. *Annals of surgical oncology*. 2002;9(10):982-7.
125. Kronz JD, Westra WH, Epstein JI. Mandatory second opinion surgical pathology at a large referral hospital. *Cancer: Interdisciplinary International Journal of the American Cancer Society*. 1999;86(11):2426-35.
  126. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature communications*. 2014;5:4006.
  127. Segal E, Sirlin CB, Ooi C, Adler AS, Gollub J, Chen X, et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. *Nature biotechnology*. 2007;25(6):675.
  128. Diehn M, Nardini C, Wang DS, McGovern S, Jayaraman M, Liang Y, et al. Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. *Proceedings of the National Academy of Sciences*. 2008;105(13):5213-8.
  129. Kuo MD, Gollub J, Sirlin CB, Ooi C, Chen X. Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. *Journal of Vascular and Interventional Radiology*. 2007;18(7):821-30.
  130. Grossmann P, Gutman DA, Dunn WD, Holder CA, Aerts HJ. Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in Glioblastoma. *BMC cancer*. 2016;16(1):611.
  131. Crispin-Ortuzar M, Apte A, Grkovski M, Oh JH, Lee NY, Schöder H, et al. Predicting hypoxia status using a combination of contrast-enhanced computed tomography and [(18)F]-Fluorodeoxyglucose positron emission tomography radiomics features. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*. 2018;127(1):36-42.
  132. Couronné R, Probst P, Boulesteix AL. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC bioinformatics*. 2018;19(1):270.
  133. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for Quantitative Radiomic Biomarkers. *Scientific Reports*. 2015;5:13087.
  134. Velazquez ER, Aerts HJ, Gu Y, Goldgof DB, De Ruyscher D, Dekker A, et al. A semiautomatic CT-based ensemble segmentation of lung tumors: comparison with oncologists' delineations and with the surgical specimen. *Radiotherapy and Oncology*. 2012;105(2):167-73.
  135. van Dam IE, de Koste JRvS, Hanna GG, Muirhead R, Slotman BJ, Senan S. Improving target delineation on 4-dimensional CT scans in stage I NSCLC using a deformable registration tool. *Radiotherapy and Oncology*. 2010;96(1):67-72.
  136. Chalian H, Tochetto SM, Tore HG, Rezai P, Yaghmai V. Hepatic tumors: region-of-interest versus volumetric analysis for quantification of attenuation at CT. *Radiology*. 2012;262(3):853-61.
  137. Haralick RM, Shanmugam K. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*. 1973(6):610-21.
  138. Wibmer A, Hricak H, Gondo T, Matsumoto K, Veeraraghavan H, Fehr D, et al. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. *European radiology*. 2015;25(10):2840-50.

139. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer research*. 2017;77(21):e104-e7.
140. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *European journal of cancer (Oxford, England : 1990)*. 2012;48(4):441-6.
141. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nature reviews Clinical oncology*. 2017;14(12):749-62.
142. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC medicine*. 2015;13(1):1.
143. Van Rossum PS, Fried DV, Zhang L, Hofstetter WL, Van Vulpen M, Meijer GJ, et al. The incremental value of subjective and quantitative assessment of 18F-FDG PET for the prediction of pathologic complete response to preoperative chemoradiotherapy in esophageal cancer. *Journal of Nuclear Medicine*. 2016;57(5):691-700.
144. Cunliffe A, Armato III SG, Castillo R, Pham N, Guerrero T, Al-Hallaq HA. Lung texture in serial thoracic computed tomography scans: correlation of radiomics-based features with radiation therapy dose and radiation pneumonitis development. *International Journal of Radiation Oncology\* Biology\* Physics*. 2015;91(5):1048-56.
145. Leijenaar RT, Carvalho S, Hoebbers FJ, Aerts HJ, Van Elmpt WJ, Huang SH, et al. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. *Acta oncologica*. 2015;54(9):1423-9.
146. Cui Y, Song J, Pollom E, Alagappan M, Shirato H, Chang DT, et al. Quantitative analysis of 18F-Fluorodeoxyglucose positron emission tomography identifies novel prognostic imaging biomarkers in locally advanced pancreatic cancer patients treated with stereotactic body radiation therapy. *International Journal of Radiation Oncology\* Biology\* Physics*. 2016;96(1):102-9.
147. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays. *Radiology*. 2016;281(2):382-91.
148. Teruel JR, Heldahl MG, Goa PE, Pickles M, Lundgren S, Bathen TF, et al. Dynamic contrast - enhanced MRI texture analysis for pretreatment prediction of clinical and pathological response to neoadjuvant chemotherapy in patients with locally advanced breast cancer. *NMR in Biomedicine*. 2014;27(8):887-96.
149. Liang C, Huang Y, He L, Chen X, Ma Z, Dong D, et al. The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage I-II and stage III-IV colorectal cancer. *Oncotarget*. 2016;7(21):31401-12.
150. Huang Y-q, Liang C-h, He L, Tian J, Liang C-s, Chen X, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *Journal of Clinical Oncology*. 2016;34(18):2157-64.
151. Ng F, Ganeshan B, Kozarski R, Miles KA, Goh V. Assessment of primary colorectal

- cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. *Radiology*. 2013;266(1):177-84.
152. Lubner MG, Stabo N, Lubner SJ, del Rio AM, Song C, Halberg RB, et al. CT textural analysis of hepatic metastatic colorectal cancer: pre-treatment tumor heterogeneity correlates with pathology and clinical outcomes. *Abdominal Imaging*. 2015;40(7):2331-7.
  153. Yang L, Dong D, Fang M, Zhu Y, Zang Y, Liu Z, et al. Can CT-based radiomics signature predict KRAS/NRAS/BRAF mutations in colorectal cancer? *European radiology*. 2018;28(5):2058-67.
  154. Ahn SJ, Kim JH, Park SJ, Han JK. Prediction of the therapeutic response after FOLFOX and FOLFIRI treatment for patients with liver metastasis from colorectal cancer using computerized CT texture analysis. *European journal of radiology*. 2016;85(10):1867-74.
  155. Beckers RCJ, Lambregts DMJ, Schnerr RS, Maas M, Rao S-X, Kessels AGH, et al. Whole liver CT texture analysis to predict the development of colorectal liver metastases: a multicentre study. *European journal of radiology*. 2017;92:64-71.
  156. Lee SJ, Zea R, Kim DH, Lubner MG, Deming DA, Pickhardt PJ. CT texture features of liver parenchyma for predicting development of metastatic disease and overall survival in patients with colorectal cancer. *European radiology*. 2018;28(4):1520-8.
  157. Rao SX, Lambregts DM, Schnerr RS, van Ommen W, van Nijnatten TJ, Martens MH, et al. Whole-liver CT texture analysis in colorectal cancer: Does the presence of liver metastases affect the texture of the remaining liver? *United European gastroenterology journal*. 2014;2(6):530-8.
  158. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RTH, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*. 2015;114(3):345-50.
  159. Huynh E, Coroller TP, Narayan V, Agrawal V, Hou Y, Romano J, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. *Radiotherapy and Oncology*. 2016;120(2):258-66.
  160. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting malignant nodules from screening CT scans. *Journal of Thoracic Oncology*. 2016;11(12):2120-8.
  161. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology*. 2016;281(3):947-57.
  162. Plautz TE, Zheng C, Noid G, Li XA. Time stability of delta-radiomics features and the impact on patient analysis in longitudinal CT images. *Medical physics*. 2019;46(4):1663-76.
  163. Jeon SH, Song C, Chie EK, Kim B, Kim YH, Chang W, et al. Delta-radiomics signature predicts treatment outcomes after preoperative chemoradiotherapy and surgery in rectal cancer. *Radiat Oncol*. 2019;14(1):43-.
  164. Fave X, Zhang L, Yang J, Mackin D, Balter P, Gomez D, et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Scientific reports*. 2017;7(1):588-.

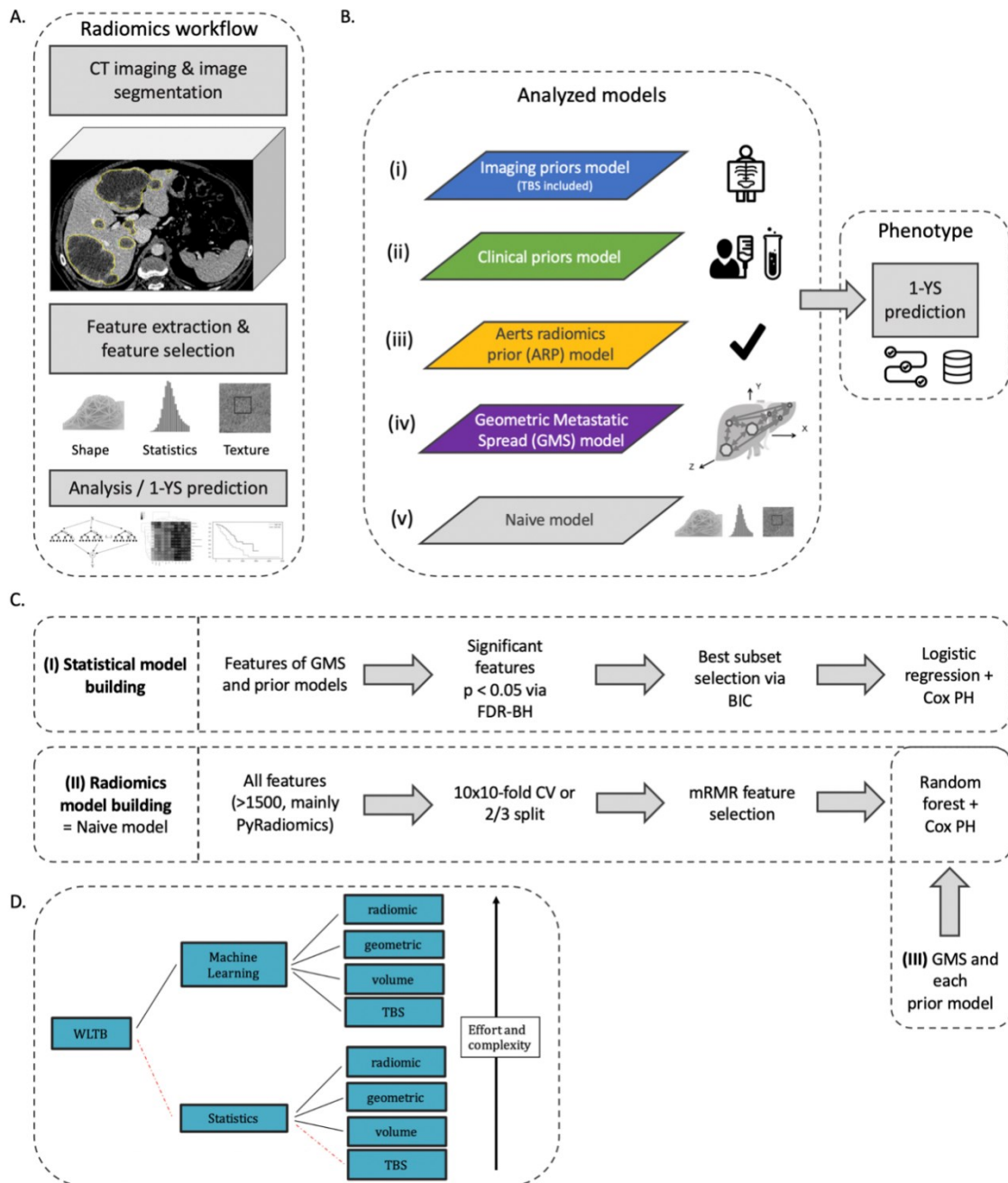
165. Boldrini L, Cusumano D, Chiloiro G, Casà C, Masciocchi C, Lenkowicz J, et al. Delta radiomics for rectal cancer response prediction with hybrid 0.35 T magnetic resonance-guided radiotherapy (MRgRT): a hypothesis-generating study for an innovative personalized medicine approach. *La Radiologia medica*. 2019;124(2):145-53.
166. Cherezov D, Hawkins SH, Goldgof DB, Hall LO, Liu Y, Li Q, et al. Delta radiomic features improve prediction for lung cancer incidence: A nested case-control analysis of the National Lung Screening Trial. *Cancer Med*. 2018;7(12):6340-56.
167. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet (London, England)*. 2009;374(9683):86-9.
168. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med*. 2010;8:24.
169. Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? *Nature reviews Drug discovery*. 2011;10(9):712.
170. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. *Nature*. 2012;483(7391):531-3.
171. Miles KA, Ganeshan B, Griffiths MR, Young RC, Chatwin CR. Colorectal cancer: texture analysis of portal phase hepatic CT images as a potential marker of survival. *Radiology*. 2009;250(2):444-52.
172. Siemens Healthineers Deutschland. syngo.via: Powerful reading. Actionable results. Reading as it should be. [cited 2020 May 15]. Available from: <https://www.siemens-healthineers.com/de/medical-imaging-it/advanced-visualization-solutions/syngovia>.
173. MeVisLab: MeVis Medical Solutions AG, Bremen, Germany; [Available from: <https://www.mevislab.de/>].
174. Ritter F, Boskamp T, Homeyer A, Laue H, Schwier M, Link F, et al. Medical Image Analysis. *IEEE Pulse*. 2011;2(6):60-70.
175. Chlebus G, Schenk A, Moltz JH, van Ginneken B, Hahn HK, Meine H. Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. *Sci Rep*. 2018;8(1):15497.
176. Sahu S, Scherthaner R, Ardon R, Chapiro J, Zhao Y, Sohn JH, et al. Imaging Biomarkers of Tumor Response in Neuroendocrine Liver Metastases Treated with Transarterial Chemoembolization: Can Enhancing Tumor Burden of the Whole Liver Help Predict Patient Survival? *Radiology*. 2017;283(3):883-94.
177. Muhlberg A, Katzmann A, Heinemann V, Kargel R, Wels M, Taubmann O, et al. The Technome - A Predictive Internal Calibration Approach for Quantitative Imaging Biomarker Research. *Sci Rep*. 2020;10(1):1103.
178. Froelich MF, Heinemann V, Sommer WH, Holch JW, Schoeppe F, Hesse N, et al. CT attenuation of liver metastases before targeted therapy is a prognostic factor of overall survival in colorectal cancer patients. Results from the randomised, open-label FIRE-3/AIO KRK0306 trial. *European radiology*. 2018;28(12):5284-92.
179. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust Radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One*. 2014;9(7):e102107.

180. Parmar C, Leijenaar RT, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Sci Rep.* 2015;5:11044.
181. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RT, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology.* 2015;114(3):345-50.
182. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology.* 2005;3(2):185-205.
183. Lu L, Ehmke RC, Schwartz LH, Zhao B. Assessing Agreement between Radiomic Features Computed for Multiple CT Imaging Settings. *PLoS One.* 2016;11(12):e0166550.
184. Kim H, Park CM, Lee M, Park SJ, Song YS, Lee JH, et al. Impact of Reconstruction Algorithms on CT Radiomic Features of Pulmonary Tumors: Analysis of Intra- and Inter-Reader Variability and Inter-Reconstruction Algorithm Variability. *PLoS One.* 2016;11(10):e0164924.
185. Ger RB, Zhou S, Chi PM, Lee HJ, Layman RR, Jones AK, et al. Comprehensive Investigation on Controlling for CT Imaging Variabilities in Radiomics Studies. *Sci Rep.* 2018;8(1):13047.
186. Mühlberg A, Holch JW, Heinemann V, Huber T, Moltz J, Maurus S, et al. The relevance of CT-based geometric and radiomics analysis of whole liver tumor burden to predict survival of patients with metastatic colorectal cancer. *European radiology.* 2020.
187. Dziak JJ, Coffman DL, Lanza ST, Li R, Jermin LS. Sensitivity and specificity of information criteria. *Briefings in Bioinformatics.* 2020;21(2):553-65.
188. Cameron Davidson-Pilon JK, Paul Zivich, Ben Kuhn, Mike Williamson, AbdealiJK, ... André F. Rendeiro. *CamDavidsonPilon/lifelines: v0.23.7* [cited 2020 Feb 7]. Available from: [www.zenodo.org](http://www.zenodo.org).
189. Wang L, Dong T, Xin B, Xu C, Guo M, Zhang H, et al. Integrative nomogram of CT imaging, clinical, and hematological features for survival prediction of patients with locally advanced non-small cell lung cancer. *European radiology.* 2019;29(6):2958-67.
190. Zhao L, Gong J, Xi Y, Xu M, Li C, Kang X, et al. MRI-based radiomics nomogram may predict the response to induction chemotherapy and survival in locally advanced nasopharyngeal carcinoma. *European radiology.* 2020;30(1):537-46.
191. Ingrisch M, Schneider MJ, Norenberg D, Negrao de Figueiredo G, Maier-Hein K, Suchorska B, et al. Radiomic Analysis Reveals Prognostic Information in T1-Weighted Baseline Magnetic Resonance Imaging in Patients With Glioblastoma. *Investigative radiology.* 2017;52(6):360-6.
192. Fleckenstein FN, Scherthaner RE, Duran R, Sohn JH, Sahu S, Zhao Y, et al. 3D Quantitative tumour burden analysis in patients with hepatocellular carcinoma before TACE: comparing single-lesion vs. multi-lesion imaging biomarkers as predictors of patient survival. *Eur Radiol.* 2016;26(9):3243-52.
193. Sasaki K, Margonis GA, Andreatos N, Zhang X-F, Buettner S, Wang J, et al. The prognostic utility of the “Tumor Burden Score” based on preoperative radiographic features of colorectal liver metastases. *Journal of Surgical Oncology.*

- 2017;116(4):515-23.
194. Kamran SC, Clark JW, Zheng H, Borger DR, Blaszkowsky LS, Allen JN, et al. Primary tumor sidedness is an independent prognostic marker for survival in metastatic colorectal cancer: Results from a large retrospective cohort with mutational analysis. *Cancer Med.* 2018;7(7):2934-42.
  195. Modest DP, Stintzing S, von Weikersthal LF, Decker T, Kiani A, Vehling-Kaiser U, et al. Exploring the effect of primary tumor sidedness on therapeutic efficacy across treatment lines in patients with metastatic colorectal cancer: analysis of FIRE-3 (AIOKRK0306). *Oncotarget.* 2017;8(62):105749-60.
  196. Kim J, Choi SJ, Lee SH, Lee HY, Park H. Predicting Survival Using Pretreatment CT for Patients With Hepatocellular Carcinoma Treated With Transarterial Chemoembolization: Comparison of Models Using Radiomics. *AJR American journal of roentgenology.* 2018;211(5):1026-34.
  197. Zhou M, Hall L, Goldgof D, Russo R, Balagurunathan Y, Gillies R, et al. Radiologically defined ecological dynamics and clinical outcomes in glioblastoma multiforme: preliminary results. *Transl Oncol.* 2014;7(1):5-13.
  198. Park JE, Kim D, Kim HS, Park SY, Kim JY, Cho SJ, et al. Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement. *European radiology.* 2020;30(1):523-36.
  199. Mitchell JBO. Machine learning methods in chemoinformatics. *Wiley Interdisciplinary Reviews: Computational Molecular Science.* 2014;4(5):468-81.
  200. Adelabu S, Mutanga O, Adam E. Testing the reliability and stability of the internal accuracy assessment of random forest for classifying tree defoliation levels using different validation methods. *Geocarto International.* 2015;30(7):810-21.
  201. Koprowski R, Foster KR. Machine learning and medicine: book review and commentary. *BioMedical Engineering OnLine.* 2018;17(1).

Anhang

**Abbildung A1:** Durchgeführter Radiomics-Workflow und Modellerstellung, übernommen aus (186)

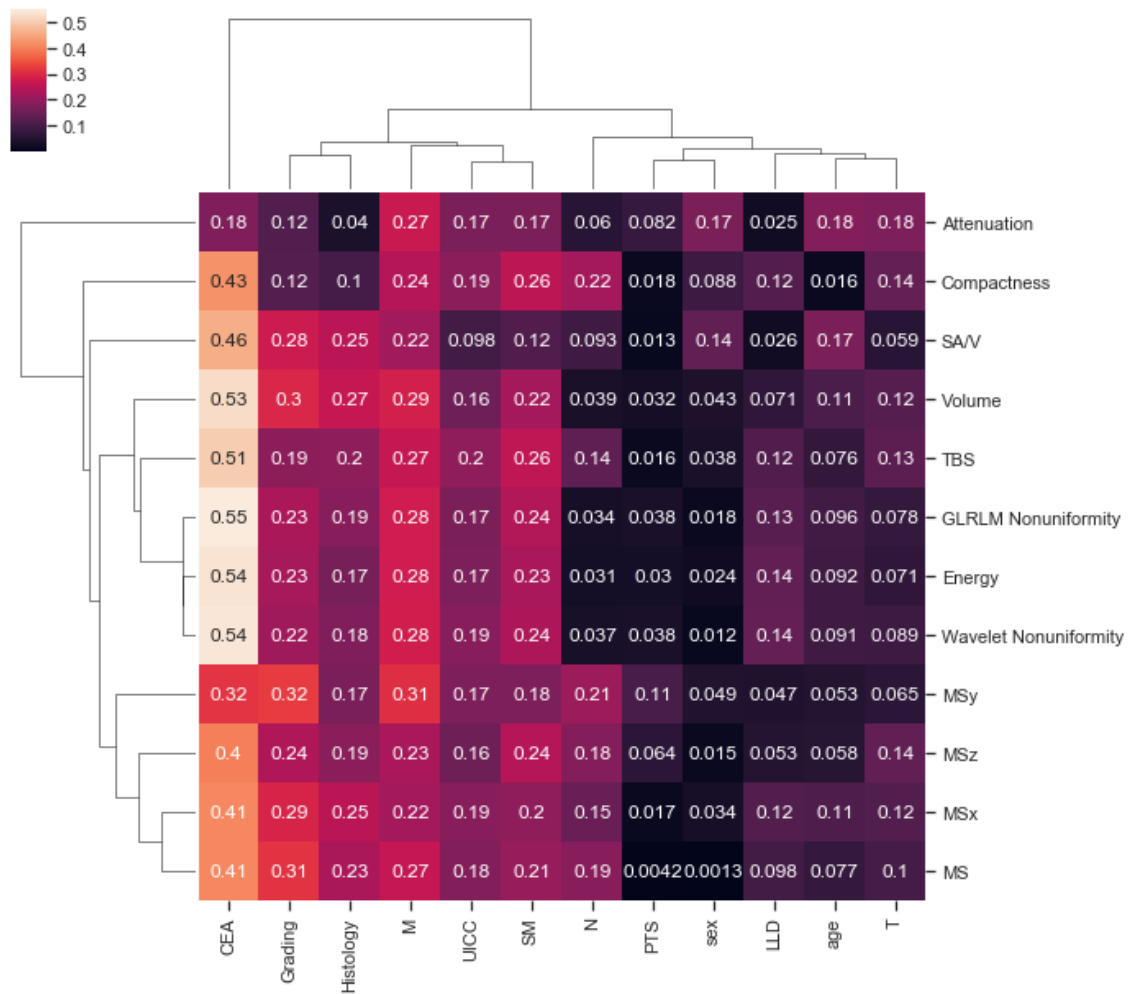


Schemata der durchgeführten Analysen. (A) zeigt den generellen Radiomics-Workflow dieser Studie. In (B) sind die analysierten Modelle aufgelistet. (C) stellt die Schritte zum Modellbau über (I) den statistischen Ansatz, (II und III) und den Machine-Learning-Ansatz zu den klinischen und Imaging Biomarker-Modellen (i-ii), zum bereits etablierten Radiomics-Modell (iii), zur GMS-Hypothese (iv) sowie zum naiven Modell (v), welches alle klassischen Radiomics Features enthält. (D) beschreibt den zusätzlichen prädiktiven Nutzen einer radiomischen und geometrischen WLTB-Analyse. Die Hauptvergleiche sind konventionelle Statistik vs. Machine Learning und der etablierte quantitative Imaging Biomarker (QIB) TBS vs. radiomische und geometrische Analyse. Grob unterteilt, steigt die

## Anhang

Komplexität (*complexity*) und der Aufwand (*effort*) von unten nach oben aufgrund des höheren Aufwandes, die WLTB-Segmentierungen zu erstellen, der höheren Modellkomplexität eines nicht linearen Machine-Learning-Ansatzes im Vergleich zu regulären statistischen Modellen und der Komplexität, bei radiomischer Texturanalyse den Einfluss der verschiedenen im Patientenkollektiv verwendeten CT-Scanner zu kontrollieren.

**Abbildung A2:** Zusammenhänge zwischen den klinischen Features und Imaging Features, übernommen aus (186)



*Spearman correlation heatmap* mit absoluten Werten und Dendrogramm, um die Assoziation zwischen klinischen und Imaging Features zu visualisieren. SM: syn-/metachronous, PTS: primary tumor sidedness, TBS: Tumor burden score, MS: Metastatic spread, MSx: Metastatic spread entlang der CT-Scanner x-Achse (etc.).



**Tabelle A1:** Basisdaten des untersuchten Patientenkollektivs

	<b>Charakteristika</b>	<b>Gesamtkollektiv</b>
<b>Demographische und klinische Patientendaten</b>	<b>Durchschnittsalter bei Erstdiagnose in Jahren n</b> (Standardabweichung (SD); Spannweite)	<b>n=103</b>  61,0 (11,2; 33,0-81,0)
	<b>Geschlecht n (%)</b>	<b>n=103</b>
	männlich	55 (53,4)
	weiblich	48 (46,6)
	<b>Lokalisation des Primarius (PTS) n (%)</b>	<b>n=102</b>
	rechtsseitig	34 (41,2)
	linksseitig	68 (58,8)
	<b>Liver limited Disease (LLD) n (%)</b>	<b>n=103</b> 44(42,7)
	<b>KRAS n (%)</b>	<b>n=95</b>
	Wildtyp	53 (55,8)
	mutiert	42 (44,2)
	<b>BRAF n (%)</b>	<b>n=36</b>
	Wildtyp	32 (88,9)
	mutiert	4(11,1)
	<b>Durchschnitts-CEA n (SD)</b>	<b>n=87</b> 552,4 (1392,8)
	<b>Durchschnitts-Ca19-9 n (SD)</b>	<b>n=77</b> 1418,5 (2418,8)
	<b>Durchschnitts-LDH n (SD)</b>	<b>n=87</b> 531,8 (439,4)
	<b>syn- oder metachrone Metastasierung n (%)</b>	<b>n=103</b>
	synchron	81 (78,6)
	metachron	22 (21,4)
	<b>Gesamtumorvolumen der Leber in cm<sup>3</sup> n (SD)</b>	<b>n=103</b> 332,7 (469,7)
	<b>Betroffene anatomische Leberlappen n (%)</b>	<b>n=103</b>
	links	0 (0,0)
rechts	26 (25,2)	
beide	77 (74,8)	
<b>Betroffene chirurgische Leberlappen n (%)</b>	<b>n=103</b>	
links	4 (3,9)	
rechts	13 (12,6)	
beide	82 (79,6)	
<b>Lobus caudatus mitbetroffen (Segment I) n (%)</b>	<b>n=103</b>  31 (30,1)	

Anhang

	<b>Lebersegment mit Haupttumorlast n (%)</b>	<b>n=48</b>
	I	0 (0,0)
	II	4 (8,3)
	III	5 (10,4)
	IVa	8 (16,6)
	IVb	3 (19,6)
	V	5 (10,4)
	VI	7 (14,6)
	VII	5 (10,4)
	VIII	11 (22,9)
<b>Patientendaten zur Therapie</b>	<b>Resektionsstatus des Primarius n (%)</b>	<b>n=60</b>
	R0	47 (78,3)
	R1	8 (13,3)
	R2	3 (5,0)
	Rx	2 (3,3)
	<b>Art der Chemotherapie n (%)</b>	<b>n=91</b>
	FOLFOX/CAPOX	44 (48,4)
	FOLFIRI/CAPIRI	33 (36,3)
	FOLFOXIRI/CAPOXIRI	12 (13,2)
	nur Fluorpyrimidin und/oder Folsäure-Analoga (5-FU/LV; Capecitabine; Raltitrexed)	2 (2,2)
	<b>Antikörperinsatz n (%)</b>	<b>n=83</b>
	keine	17 (20,5)
	Cetuximab	18 (21,7)
	Bevacicumab	43 (51,8)
	Panitumomab	5 (6,0)
	<b>Durchschnittlicher Untersuchungszeitraum von <i>baseline</i> bis <i>follow-up</i> in Wochen n (SD)</b>	<b>n=88</b>
		12,1 (7,7)
	<b>Durchschnittsdauer der Erstlinientherapie in Monaten n (SD)</b>	<b>n=66</b>
		5,5 (4,0)
	<b>Bestes Ansprechen im Verlauf nach RECIST n %</b>	<b>n=92</b>
	CR	4 (4,3)
	PR	50 (54,3)
	SD	29 (31,5)
	PD	9 (9,8)
	<b>Durchschnittliche Überlebenszeit in Monaten n (SD)</b>	<b>n=103</b>
		21,5 (12,6)
	<b>1-Jahres-Überleben n (%)</b>	<b>n=103</b>
ja	82 (79,6)	
nein	21 (20,4)	

Anhang

**Tabelle A2:** Ergebnisse der univariaten statistischen Analyse, übernommen aus (186)

	N	mean no 1-YS	SD no 1-YS	mean 1-YS	SD 1-YS	Test	p	p_bh	C- index
CT variables									
Attenuation [HU]	103	68.1	13.1	63.4	13.6	t-test	0.16	0.19	0.52
Volume [cm <sup>3</sup> ]	103	438.4	581.2	305.7	432.5	u-test	0.11	0.14	0.58
TBS	103	28.4	18.7	16.0	15.0	u-test	0.005	0.022	0.62*
<i>MSx [mm]</i>	103	<i>167.5</i>	<i>61.3</i>	<i>143.5</i>	<i>67.8</i>	<i>u-test</i>	<i>0.17</i>	<i>0.19</i>	<i>0.59</i>
<i>MSy [mm]</i>	103	<i>153.9</i>	<i>52.4</i>	<i>125.0</i>	<i>54.0</i>	<i>u-test</i>	<i>0.015</i>	<i>0.045</i>	<i>0.62*</i>
<i>MSz [mm]</i>	103	<i>151.1</i>	<i>61.7</i>	<i>111.9</i>	<i>54.5</i>	<i>t-test</i>	<i>0.006</i>	<i>0.022</i>	<i>0.65**</i>
<i>MS [mm]</i>	103	<i>472.5</i>	<i>161.3</i>	<i>380.5</i>	<i>163.3</i>	<i>u-test</i>	<i>0.027</i>	<i>0.065</i>	<i>0.62*</i>
<i>SA/V</i>	103	<i>0.31</i>	<i>0.15</i>	<i>0.33</i>	<i>0.19</i>	<i>u-test</i>	<i>0.99</i>	<i>0.99</i>	<i>0.50</i>
<i>Compactness</i>	103	<i>0.34</i>	<i>0.15</i>	<i>0.44</i>	<i>0.15</i>	<i>u-test</i>	<i>0.004</i>	<i>0.022</i>	<i>0.65**</i>
GLRLM Nonuniformity	103	59007	81891	48180	89114	u-test	0.076	0.14	0.58
Wavelet Nonuniformity	103	104 x 10 <sup>3</sup>	145 x 10 <sup>3</sup>	84 x 10 <sup>3</sup>	156 x 10 <sup>3</sup>	u-test	0.085	0.14	0.58
Energy	103	439 x 10 <sup>9</sup>	621 x 10 <sup>9</sup>	390 x 10 <sup>9</sup>	744 x 10 <sup>9</sup>	u-test	0.11	0.14	0.59
Clinical variables									
PTS	102	r:11, l:10	---	r:23, l:58	---	Fisher	0.067	0.80	0.58*
Sex	103	w:11, m:10	---	w:37, m:45	---	Fisher	0.63	0.83	0.51
Age	103	60.9	13.0	61.0	10.7	u-test	0.83	0.83	0.55
Grading	93	G2:14, G3:3, Gx:1	---	G1:2, G2:55, G3:18	---	Chi2	0.48	0.83	0.51
Histology	93	1:20, 12:1	---	1:64, 2:6, 15:1, 10:1	---	Chi2	0.75	0.83	0.54

Anhang

Syn./Met.	103	met:4, syn:18	---	met:18, syn:63	---	Chi2	0.72	0.83	0.54
CEA	87	441.4	1268.5	574.7	1417.3	t-test	0.71	0.83	0.51
UICC	99	38.9	5.4	37.5	8.190	Chi2	0.63	0.83	0.60
LLD	103	n:13, y:8	---	n:46, y:36	---	Chi2	0.81	0.93	0.50
T	86	T1:1, T2:1, T3:9, T4:6, Tx:1	---	T1:3, T2:2, T3:45, T4:12, Tx:6	---	Chi2	0.60	0.83	0.46
M	94	M0:3, M1:15, M2:1	---	M0:15, M1:58, M2:2	---	Chi2	0.79	0.83	0.53
N	84	N0:2, N1:4, N2:10, Nx:2	---	N0:15, N1:22, N2:24, Nx:5	---	Chi2	0.39	0.83	0.61*

Univariate statistische Analyse. CT-Variablen des *imaging priors model (i)*, des *GMS model (iv)* (kursiv), des *Aerts radiomics prior model (iii)* (hellgrau) und klinische Variablen. P-Werte mit (p\_bh) und ohne (p) multipler Testkorrektur, C-index und Signifikanz im Log-likelihood-Test ( $p < 0.05$ : \*,  $p < 0.01$ : \*\*). Histologie: 1: adenocarcinoma (n.s = not specified), 2: mucinous adenocarcinoma, 3: adenocarcinoma in villous adenoma, 4: adenocarcinoma in tubulo-villous adenoma, 5: mucinous carcinoma, 6: adenocarcinoma in tubular adenoma, 7: medullar adenocarcinoma, 8: tubular adenocarcinoma, 9: mixed-cell adenocarcinoma, 10: adenocarcinoma with neuroendocrine differentiation, 11: intraductal papillary-mucinous neoplasia associated with invasive carcinoma, 12: Mixed adeno-neuroendocrine carcinoma (MANEC), 13: malignant gastrointestinal stromal tumor (GIST), 14: Signet ring cell carcinoma (SRCC), 15: malignant tumor (n.s.) 16: papillary adenocarcinoma (n.s.), 17: mixed-cell carcinoid-adenocarcinoma

**Tabelle A3:** Multivariate Ein-Feature-Modelle, übernommen aus (186)

Model	BIC	Multivariable Model	OR [CI]	p	AUC
N1	105.0	TBS	1.9 [1.2, 3.1]	0.004	0.70* [0.57, 0.83]
N2	105.5	MSz	2.1 [1.2, 3.6]	0.008	0.69* [0.55, 0.82]
N3	106.0	Compactness	2.1 [1.2, 3.9]	0.012	0.71* [0.58, 0.83]
N4	107.4	MS	1.8 [1.1, 3.0]	0.02	0.66* [0.53, 0.79]
N5	108.5	MSy	1.8 [1.0, 3.2]	0.04	0.67* [0.53, 0.81]

Multivariate statistische Modelle (logistische Regression) mit Odds Ratio (OR) inklusive 95%-CI, p-Wert und diskriminativer AUC mit 95%-CI. TBS: Tumor burden score. \*: signifikant ( $p < 0.05$ ) besser als informiertes Raten. Anmerkung: Alle 103 Patienten wurden für die Modellkonstruktion benutzt, d.h. keine Daten blieben für die anstehende 1-YS-Diskrimination ungesehen.

**Tabelle A4:** Ergebnisse der Machine-Learning-Analyse, übernommen aus (186)

Feature	Variable importance [%] CV	AUC CV	C-index SPLIT	
			Train (N = 69)	Test (N = 34)
mCRC QIBs				
TBS#	---	0.68* [0.54, 0.79]	0.66	0.64
PTS#	---	0.51 [0.37, 0.66]	0.58	0.59
WLTB volume#	---	0.53 [0.40, 0.66]	0.62	0.50
Attenuation#	---	0.57 [0.44, 0.69]	0.51	0.41
Imaging priors model				
TBS	31.4	0.67* [0.54, 0.79]	0.69	0.63
PTS	7.1			
WLTB volume	31.4			
Attenuation	30.1			
Clinical priors model				
PTS	5.2	0.56 [0.43, 0.67]	0.65	0.49
Sex	3.5			
Age	25.1			
CEA	20.1			
Grading	8.1			
Histology	1.8			
Syn-/metachronous disease	2.4			
UICC	10.6			
LLD	2.6			
T	6.2			
M	5.3			

## Anhang

N	9.1			
Aerts radiomics prior (ARP) model				
Energy	21.1	0.76* [0.65,0.86]	0.66	0.67
Compactness	32.8			
GLRLM Nonuniformity	25.7			
Wavelet Nonuniformity	20.3			
Geometric metastatic spread (GMS) model				
MSx [mm]	16.4	0.73* [0.60, 0.84]	0.66	0.70
MSy [mm]	14.4			
MSz [mm]	18.0			
MS [mm]	15.5			
SA/V	14.3			
Compactness	21.4			
Naive model				
§		0.65* [0.55, 0.78]	0.68	0.48

Etablierte QIB und Modelle in der Machine-Learning-Analyse zur Prädiktion des 1-Jahres-Überlebens für ungesehene Patientendaten. Die Prädiktionsmodelle basieren auf einem Random Forest als Klassifikator. Die Ergebnisse sind, sofern verfügbar, nach Abschätzung der internen Validität durch 10x10CV und temporalen 2/3-Split (SPLIT) auf den Testdaten dargestellt. Die AUCs sind mit dem 95%-Konfidenzintervall angegeben. #: Logistische Regression wurde als Klassifikator für die Prädiktion benutzt, \*: signifikant ( $p < 0.05$ ) besser als informiertes Raten. §: Die *variable importance* ist hier aufgrund der Vielzahl verwendeter Features nicht von Bedeutung. QIB: quantitative imaging biomarker, PTS: primary tumor sidedness, TBS: tumor burden score, GMS: Geometric metastatic spread model, ARP: Aerts radiomics prior model.

**Tabelle A5:** Ergebnisse der Subgruppenanalyse des *Aerts radiomics prior model (iii)*, modifiziert nach (186)

Feature	AUC SPLIT Test (N = 34)
Aerts Heterogeneity	
Energy	0.67* [0.55, 0.77]
GLRLM NonUniformity	
Wavelet Nonuniformity	
Calibrated Aerts Heterogeneti (Aerts Heterogeneity & Calibration Features)	
	0.73* [0.59, 0.84]
Energy	
GLRLM NonUniformity	
<i>Trachea_wavelet_LHH_glrlm</i>	
<i>Trachea_wavelet_LHH_gldm_ LargeDependenceEmphasis</i>	
<i>ATStripe_logarithm_glszm_ GrayLevelNonUniformityNormalized</i>	
<i>AirStripe_wavelet_LHL_glcm_Cluster- Shade</i>	

Heterogenitätsfeatures des Aerts radiomics prior model (iii) ohne (oben) und mit (unten) interner Kalibrierung. Inkorporierte Kalibrierungsfeatures sind in kursiv dargestellt. Das Feature "Wavelet Nonuniformity" ist im untenstehenden kalibrierten Modell nicht mehr enthalten, da es im Rahmen der internen Kalibrierung aufgrund von fehlender Relevanz für das Modell automatisch aussortiert und durch die Kalibrierungsfeatures ersetzt wurde. Diskriminative AUC [95%-CI] für 1-Jahres-Überleben bei temporalem 2/3-Split (SPLIT) auf den Testdaten. \*: signifikant ( $p < 0.05$ ) besser als informiertes Raten.



**Tabelle A6:** Der Einfluss des CT-Scanner-Typs auf die Radiomics Features, übernommen aus (186)

Feature	Effect size [%]
Energy	47.1
GCRLM Nonuniformity	44.0
Wavelet Nonuniformity	41.0
MSx	24.5
MSy	18.1
MSz	18.2
MS	21.7
SA/V	28.2
Compactness	21.8

Um den Einfluss verschiedener CT-Scanner auf die verwendeten Features zu untersuchen, wurde deren Effektstärke über eine ANCOVA-Analyse geschätzt. Features mit hoher Effektstärke sind abhängiger vom CT-Scannertyp als Features mit niedriger Effektstärke. Hier ist es wichtig zu wissen, dass die Ergebnisse aufgrund der Vielzahl der verwendeten CT-Scanner (>20) bei einem Kollektiv von 103 Patienten nicht repräsentativ sind, dennoch wird aus der Analyse bei Textur-Features eine fast doppelt so hohe Varianz in Abhängigkeit vom verwendeten CT-Scanner deutlich wie bei nicht-texturabhängigen Features.

## Veröffentlichungen

Teile dieser Doktorarbeit wurden in folgenden Artikeln veröffentlicht:

### **The relevance of CT-based geometric and radiomics analysis of whole liver tumor burden to predict survival of patients with metastatic colorectal cancer**

A. Mühlberg, J. Holch, V. Heinemann, T. Huber, J. Moltz, S. Maurus, **N. Jäger**, L. Liu, M. Froelich, A. Katzmann, E. Gresser, O. Taubmann, M. Sühling, D. Nörenberg

Received: 11 March 2020 / Revised: 2 July 2020 / Accepted: 13 August 2020

European Radiology. 2020. <https://doi.org/10.1007/s00330-020-07192-y>

### **Deep Learning based Radiomics and its usage in prediction for metastatic colorectal cancer**

D. Nörenberg, T. Huber, S. Maurus, **N. Jäger**, A. Katzmann, A. Mühlberg, J. Moltz, V. Heinemann, J. Holch

Abstract (Nr. B-1586) und Kongressvortrag in: 25th European Congress of Radiology (ECR) 2019, 27. Februar bis 3. März, Wien/AT

## **Danksagung**

Ich möchte allen Menschen danken, die mich auf dem Weg zur Fertigstellung dieser Arbeit begleitet haben.

Vielen Dank an alle am PANTHER-Projekt mitwirkenden Radiologen, den Mitarbeitern von Fraunhofer MEVIS und Siemens Healthineers für ihr Engagement, die vielen produktiven Projekttreffen, die Unterstützung der eindrucksvollen Kongressreise nach Wien, die zahlreichen Analysen und innovativen Ideen.

Vielen Dank an meine Doktormutter Frau Prof. Dr. med. Birgit Ertl-Wagner für die freundliche Überlassung des Themas und die Unterstützung der Arbeit in ihren Anfängen.

Meinem Doktorvater Prof. Dr. med. Wolfgang Kunz möchte ich herzlich danken für die bereitwillige Übernahme der Betreuung im Verlauf. Danke für die freundliche und unkomplizierte Unterstützung in den zahlreichen Formalitäten auf dem Weg zur Fertigstellung und in den letzten Schritten.

Einen besonderen Dank möchte ich meinem Betreuer PD Dr. med. Dominik Nörenberg aussprechen für die gewissenhafte Betreuung dieser Arbeit über alle Arbeitsschritte hinweg. Für die viele Geduld in der Einarbeitung, die schnellen Rückmeldungen zu jeder Zeit, die vielen konstruktiven Vorschläge vom Beginn bis zum Schluss und seine über die Jahre unermüdliche Arbeitsmoral.

Zu guter Letzt gebührt ein ganz besonderer Dank meiner Familie, auf deren Unterstützung ich - sowohl vor als auch während dieser Arbeit und darüber hinaus - immer bauen konnte und können werde. Danke für die immer wieder guten Ratschläge, die zahlreichen Erinnerungen und Ermutigungen und kurzum - für die mentale Unterstützung von allen Seiten.

Und Kathrin möchte ich Danke sagen für alles, was sich nicht so einfach schreiben lässt.

## **Eidesstattliche Erklärung**

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

**„Detektion von Imaging Biomarkern zur Charakterisierung und  
Prognoseabschätzung bei Patienten mit metastasiertem kolorektalen  
Karzinom“**

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe. Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den 07.10.2022

---

Ort, Datum

Nils Jäger

---

Unterschrift Doktorand