

Aus der Klinik und Poliklinik für Diagnostische und Interventionelle Radiologie der  
Universitätsmedizin der Johannes Gutenberg-Universität Mainz

Machine learning zur Verbesserung der Vorhersagequalität  
nach TACE bei HCC-Patienten

Inauguraldissertation  
zur Erlangung des Doktorgrades der Medizin  
der Universitätsmedizin  
der Johannes Gutenberg-Universität Mainz

Vorgelegt von

Franziska Doris Carmen Wagner  
aus Bad Kreuznach

Mainz, 2022

Wissenschaftlicher Vorstand:

1. Gutachter:

2. Gutachter:

Tag der Promotion: 28.04.2023

# Inhaltsverzeichnis

Abkürzungsverzeichnis.....	I
Abbildungsverzeichnis.....	III
Tabellenverzeichnis.....	IV
1 Einleitung.....	1
1.1 Epidemiologie.....	1
1.2 Ätiologie des HCC und Risikofaktoren.....	1
1.2.1 Allgemeines.....	1
1.2.2 Zirrhose.....	2
1.2.3 Virale Hepatitiden.....	2
1.2.4 Nutritiv toxische Zirrhose.....	3
1.2.5 Weitere Ursachen.....	3
1.3 Diagnostik.....	4
1.3.1 Früherkennung.....	4
1.3.2 Sonographie.....	4
1.3.3 Computertomographie und Magnetresonanztomographie.....	5
1.3.4 Leberbiopsie.....	8
1.3.5 Labor, Child Pugh Score, Model for End-Stage Liver Disease, Eastern Cooperative Oncology Group Performance Status.....	9
1.3.6 Bildgebende Verfahren zum Tumorstaging.....	11
1.4 Einteilung der Stadien nach BCLC.....	12
1.5 Transarterielle Chemoembolisation.....	14
1.6 Statistische Definitionen.....	18
1.6.1 Harrell's C Index.....	18
1.6.2 Receiver Operating Characteristics Curve und Area under the Curve receiver Characteristics Curve.....	19
1.6.3 Brier Score.....	19
1.7 Vorhandene Scoring Systeme/Bewertungssysteme.....	19
1.7.1 Assessment for Retreatment with TACE-Score (ART-Score).....	20
1.7.2 ABCR-Score.....	22
1.7.3 SNACOR.....	24
1.8 Bewertung der Tumor response.....	26
1.8.1 Bewertung der Tumor response nach mRECIST.....	26

1.8.2	Bewertung der Tumor Response nach EASL.....	27
1.9	Artifizielle neuronale Netzwerke.....	29
1.10	Problem .....	30
1.11	Zielsetzung.....	34
2	Methoden und Material.....	35
2.1	Patientenkollektiv .....	35
2.2	Parameter .....	37
2.3	Aufbau des neuronalen Netzwerks .....	38
2.4	Training und Validierung des Neuronalen Netzwerks .....	40
2.5	Statistische Analyse .....	43
3	Ergebnis .....	43
3.1	Genauigkeit des KNN.....	44
3.2	Genauigkeit des KNN im Vergleich zu bekannten Scoring-Systemen.....	45
4	Diskussion.....	47
5	Zusammenfassung.....	52
6	Literaturverzeichnis .....	53
7	Danksagung .....	66

## Abkürzungsverzeichnis

AASLD	American Association for the Study of Liver Diseases
AFP	Alpha-Feto-Protein
ART-Score	Assessment for Retreatment with TACE-Score
AST	Aspartat-Aminotransferase
AUROC	Area Under the Curve Receiver Characteristics Curve
BCLC	Barcelona Clinic Liver Cancer
CEUS	Contrast Enhanced Ultrasound
CR	Complete Response
CT	Computertomographie
cTACE	conventional TACE
DEB	Drug Eluting Beads
DSA	Digitale Subtraktionsangiographie
EASL	European Association for the Study of the Liver
ECOG	Eastern Cooperative Oncology Group
HBV	Hepatitis B Virus
HCC	Hepatozelluläres Karzinom
HCV	Hepatitis C Virus
HR	Hazard Ratio
IBS	Integrated Briar Score
INR	International Normalized Ratio
KI	Konfidenzintervall
KNN	Künstliches Neuronales Netz
MELD	Model for End-Stage Liver Disease
mRECIST	modified Response Evaluation Criteria in Solid Tumors
MRT	Magnetresonanztomographie
MWA	Mikrowellenablation
NAFLD	Non-Alcoholic Fatty Liver Disease
NASH	Nicht-Alkoholische Steatohepatitis
PD	Progressive Disease
PR	Partial Response
RFA	Radiofrequenzablation
ROC	Receiver Operating Characteristics Curve

SD	Stable Disease
TACE	Transarterielle Chemoembolisation
TIPS	Transjugulärer Portosystemischer Shunt
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis

# Abbildungsverzeichnis

<b>Abbildung 1:</b> Diagnostikalgorithmus bei Verdacht auf ein HCC bei Patienten mit einer Leberzirrhose (23). .....	6
<b>Abbildung 2:</b> Arterielle Phase einer CT. Durch Hypervaskularisierung des HCC zeigt sich dieses hyperdens im Vergleich zum umliegenden Leberparenchym. ....	7
<b>Abbildung 3:</b> Spätphase einer CT. Es zeigt sich das für ein HCC charakteristische Wash-Out. Der Herd stellt sich hypodens im Vergleich zum umliegenden Leberparenchym dar. ....	8
<b>Abbildung 4:</b> ECOG Performance Status (44) .....	11
<b>Abbildung 5:</b> Therapiealgorithmus HCC bei vorliegender Leberzirrhose (23). ....	14
<b>Abbildung 6:</b> CT vor TACE mit arterieller Kontrastmittelphase. Es zeigt sich ein deutlich hypervaskularisierter Tumorherd. ....	16
<b>Abbildung 7:</b> Digitale Subtraktionsangiographie (DSA) intrahepatischer Gefäße mit Darstellung der tumorversorgenden Arterien. ....	17
<b>Abbildung 8:</b> DSA. Superselektive Applikation des Chemoembolisats. ....	17
<b>Abbildung 9:</b> CT nach TACE mit arterieller Kontrastmittelphase. Der Tumorherd zeigt sich deutlich hypovaskularisiert nach erfolgter Chemoembolisation. ....	18
<b>Abbildung 10:</b> Schema neuronales Netz (78) .....	29
<b>Abbildung 11:</b> Drop-Out Kriterien (86) .....	36
<b>Abbildung 12:</b> Struktur des KNN bestehend aus einem Input-Layer, 3 Hidden-Layer und dem Output-Layer (86). ....	39
<b>Abbildung 13:</b> Modell der 5-fachen Kreuzvalidierung (86).....	40
<b>Abbildung 14:</b> ROC der Trainingskohorte (86). ....	44
<b>Abbildung 15:</b> ROC der Validierungskohorte (86). ....	45
<b>Abbildung 16:</b> Vergleich der ROCs zwischen dem KNN, ART, ABRC und SNACOR (86). ....	46
<b>Abbildung 17:</b> Studienablauf und Endpunkte des KNN (86).....	49

## Tabellenverzeichnis

<b>Tabelle 1:</b> Kriterien Child Pugh Score (10) .....	10
<b>Tabelle 2:</b> Ergebnis der Regressions-Analyse der Parameter der Trainingskohorte des ART Score (71).....	21
<b>Tabelle 3:</b> Multivariate Analyse der Parameter des ABCR Score nach der ersten TACE (72). .....	23
<b>Tabelle 4:</b> Beta Koeffizient mit zugehörigem Risiko Score der Trainingskohorte des SNACOR (74).....	25
<b>Tabelle 5:</b> Definition der Tumor Response nach EASL und mRECIST (73).....	28
<b>Tabelle 6:</b> Ergebnisse der externen Validierung des ART und ABCR Score (79). ...	32
<b>Tabelle 7:</b> Vergleich des Gesamtüberlebens der Validierungsstudie im Vergleich zur ursprünglichen SNACOR Studie (80). .....	33
<b>Tabelle 8:</b> Proportionales Risikomodell zur Ermittlung unabhängiger Prädiktoren für das Überleben. Vergleich der Daten von Patienten mit hepatozellulärem Karzinom aus der ursprünglichen Studie mit den Daten dieser Studie (80).....	33
<b>Tabelle 9:</b> Detaillierte Aufstellung der Ausgangsparameter einschließlich entsprechender P-Werte. (86). .....	41



# 1 Einleitung

## 1.1 Epidemiologie

Primäre Karzinome der Leber sind weltweit die siebthäufigste Krebserkrankung und die zweithäufigste krebsbedingte Todesursache (1). Weltweit ist das hepatozelluläre Karzinom (HCC) mit etwa 75% die vorherrschende Form der Leberkarzinome (2). Im Jahr 2015 erkrankten weltweit 854.000 Menschen und 810.000 Patienten starben an einem Leberkarzinom. Von 1990 bis 2015 stieg die Inzidenz der Lebertumoren um 75% (3). Zu den Hauptrisikofaktoren gehören Infektionen mit Hepatitis B und C, Alkohol und Nikotinabusus, metabolische Faktoren und Aflatoxine (4,5). Die Inzidenz des HCCs nimmt mit zunehmendem Alter in sämtlichen Bevölkerungsgruppen zu. Der Erkrankungsspeak liegt bei 70 Jahren (6,7). Die meisten HCCs werden erst in späteren Stadien entdeckt. In etwa 80% der Fälle liegt begleitend eine Leberzirrhose vor. Aufgrund der dadurch eingeschränkten Leberfunktion sind auch die Therapiemöglichkeiten häufig eingeschränkt (8). Als potenziell kurative Therapien kommen eine Resektion, lokalablativ Verfahren oder eine Lebertransplantation in Frage. Zu den palliativen Behandlungskonzepten gehört die transarterielle Chemoembolisation (TACE), systemische Therapie und rein supportive Therapie (9). Für Patienten mit einem HCC im intermediären Stadium nach der Barcelona Clinic Liver Cancer Klassifikation (BCLC) ist die Transarterielle Chemoembolisation (TACE) die Therapie der Wahl (9).

## 1.2 Ätiologie des HCC und Risikofaktoren

### 1.2.1 Allgemeines

Ungefähr 90% aller hepatozellulären Karzinome sind mit einer bekannten Grunderkrankung assoziiert. Zu den häufigsten Grunderkrankungen gehören die Folgen des chronischen Alkoholkonsums, chronische Hepatitis B und C Infektionen und Aflatoxin-Exposition. Weltweit gesehen sind ca. 54% der HCC-Erkrankungen einer Hepatitis B Infektion zuzuschreiben, dies betrifft 400 Millionen Menschen. 31% (170 Millionen Menschen) können auf eine Hepatitis C Infektion zurückgeführt werden und 15% sind mit weiteren Ursachen assoziiert. Bei diesen Berechnungen handelt es sich jedoch um Schätzungen welche keinerlei Komorbiditäten berücksichtigen und

möglicherweise den Einfluss des metabolischen Syndroms und der nicht-Alkoholischen-Steatohepatitis unterschätzen (3).

## **1.2.2 Zirrhose**

Aufgrund von unbehandelten, chronischen Lebererkrankungen kann es zur irreversiblen Schädigung der Leber kommen. Bei der sogenannten Leberzirrhose kommt es durch Bindegewebsvermehrung zur irreversiblen Zerstörung der Läppchen und Gefäßstruktur. Durch die fibröse Umwandlung des Lebergewebes kommt es zur Bildung sogenannter Regeneratknoten. Dies führt zum Funktionsverlust des Parenchyms und somit zu einer eingeschränkten Syntheseleistung und Stoffwechselfunktion. Der fibröse Umbau des Parenchyms führt zur Erhöhung des Widerstands im hepatischen Stromgebiet und über diesen Pathway zu portaler Hypertension. Zu den klinischen Zeichen der portalen Hypertension zählen Aszites, eine Splenomegalie sowie ggfls. Omphalozele, und eine verstärkte Venenzeichnung der Bauchhaut. Durch die verminderte Syntheseleistung kommt es zu Ödemen (Albumin ↓) und zu einer erhöhten Blutungsneigung (Gerinnungsfaktoren ↓)(10). Unabhängig von der Ätiologie der Zirrhose stellt diese eine Prädisposition für ein hepatozelluläres Karzinom dar (11). Das Risiko ein HCC zu entwickeln ist allerdings am höchsten bei Patienten, die eine Zirrhose aufgrund einer viralen Hepatitis aufweisen. Insgesamt entwickeln ein Drittel aller Patienten mit einer Zirrhose ein HCC (12).

## **1.2.3 Virale Hepatitiden**

### **1.2.3.1 Hepatitis C**

Das 1989 entdeckte Hepatitis C Virus (HCV) stellt ein weltweites Gesundheitsproblem dar und betrifft über 170 Millionen Menschen (13). Das HCV ist ein Einzelstrang RNA-Virus und gehört zur Familie der Flaviridae, welche wiederum zur Gattung der Hepaciviren gehört (14). Es werden 7 Genotypen unterschieden. Metaanalysen haben ergeben, dass das Risiko für die Entwicklung eines HCCs erhöht ist bei Patienten mit Genotyp 1b (15) und Genotyp 3 (16).

### **1.2.3.2 Hepatitis B**

Bei dem Hepatitis B Virus (HBV) handelt es sich um ein hepatotropes DNA-Virus der Gruppe der Hepadnaviren (17). Patienten mit einer Hepatitis B Infektion zeigen ein erhöhtes Risiko für ein HCC mit steigender Viruslast, Dauer der Infektion und Schwere der Lebererkrankung (18). Sowohl eine Infektion mit Hepatitis B als auch C Virus begünstigen die Entstehung einer Leberzirrhose. In Abhängigkeit von der Ätiologie, Region beziehungsweise ethnischer Zugehörigkeit und dem Stadium der Zirrhose variiert das Risiko (19). 0,6- 6,5 % aller Patienten mit einer HCV Infektion entwickeln in Abhängigkeit des Ausmaßes der Leberfibrose ein HCC (20). Patienten mit einer chronischen HBV-Infektion können auch ohne eine Leberzirrhose an einem HCC erkranken. Das Risiko bei einer unbehandelten chronischen HBV Infektion, ohne Zirrhose an einem HCC zu erkranken liegt bei 0,12 % (21). Wird die chronische HBV Infektion behandelt, liegt das Risiko zwischen 0,01-1,4 % (22).

### **1.2.4 Nutritiv toxische Zirrhose**

Neben einer Hepatitis C Infektion gehört in Deutschland der Alkoholkonsum zu den häufigsten Risikofaktoren für die Entwicklung eines HCC (23). Die berechnete 5-Jahresinzidenz eines HCC auf dem Boden einer äthyltoxischen Leberzirrhose liegt bei 8% (19). Differentialdiagnostisch sollte die nicht-alkoholische Fettlebererkrankung (non-alcoholic fatty liver disease=NAFLD) berücksichtigt werden. Besonders die nicht-alkoholische Steatohepatitis (NASH), bei der es sich um einen histologischen Phänotyp einer NAFLD handelt, kann potenziell zu fortgeschrittenen Lebererkrankungen, Zirrhose und schließlich einem hepatozellulären Karzinom führen (24). Diese wird, unter anderem, mit Adipositas, Diabetes Mellitus Typ 2, Hyperlipidämie und dem metabolischen Syndrom in Verbindung gebracht (25).

### **1.2.5 Weitere Ursachen**

Seltenere Risikofaktoren für die Entwicklung einer Leberzirrhose und somit auch für die Entstehung eines HCCs, sind Alpha-1-Antitrypsinmangel (26), Hämochromatose, M. Wilson, Primär biliäre Cholangitis, primär sklerosierende Cholangitis sowie Autoimmunhepatitiden (27). Patienten mit einer Hämochromatose entwickeln in bis zu

45% der Fälle ein HCC (28).

### **1.3 Diagnostik**

Zur Diagnostik eines HCCs können sowohl bildgebende Verfahren als auch eine Biopsie mit anschließender histo-pathologischer Untersuchung angewandt werden (23).

#### **1.3.1 Früherkennung**

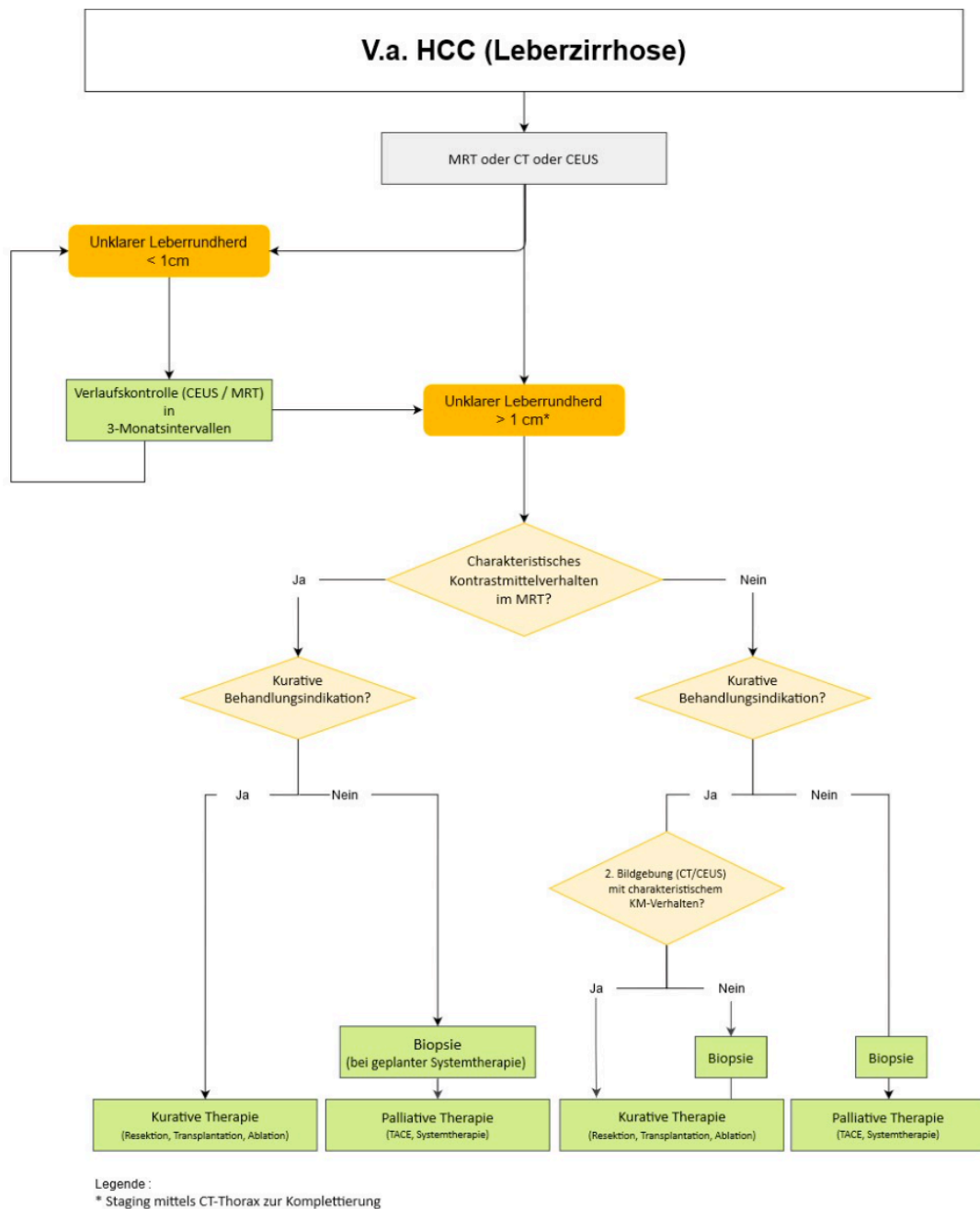
Bei Patienten mit einer Zirrhose, einer NASH oder einer chronischen HCV Infektion mit fortgeschrittener Leberfibrose sollte eine regelmäßige Früherkennung erfolgen (23,29). Dies gilt auch für Patienten mit einer chronischen Hepatitis B Infektion ab einem bestimmten Schweregrad (30,31). Eine Sonographie der Leber wird im Rahmen der Früherkennung eines HCCs alle 6 Monate empfohlen (9). Dieses Intervall wurde anhand von Studien zur Tumorverdopplungszeit ausgewählt (32). Durch Bestimmung des Alpha-Feto-Proteins (AFP) kann die Früherkennung ergänzt werden (33).

#### **1.3.2 Sonographie**

Die Sonographie ist eine kostengünstige und gut verfügbare Untersuchungsmethode. Da sie keinerlei Strahlenbelastung aufweist, wird die Sonographie insbesondere im Rahmen der Früherkennung eingesetzt. Eine weitere Möglichkeit ist der kontrastmittelverstärkte Ultraschall (CEUS). Wird eine neue Läsion bei Patienten mit einem erhöhten Risiko für ein HCC entdeckt, sollte dieser eingesetzt werden (34). Als Malignitätsmarker kann durch das Kontrastmittel die Hypervaskularisierung einer Läsion dargestellt werden. Der kontrastmittelverstärkte Ultraschall ersetzt allerdings nicht die Computertomographie (CT) beziehungsweise die Magnetresonanztomographie (MRT) zur weiteren Charakterisierung und zum Staging der Erkrankung (35).

### 1.3.3 Computertomographie und Magnetresonanztomographie

Ein HCC stellt sich in der CT und auch der MRT charakteristisch dar. In der arteriellen Phase kommt es im Vergleich zum umliegenden Gewebe zu einer verstärkten Kontrastmittelaufnahme (wash-in). Gleichzeitig zeigt sich die Läsion in den später akquirierten Phasen hypodens (CT) beziehungsweise hypointens (MRT) im Vergleich zum umliegenden Leberparenchym (wash-out). Dieses charakteristische Kontrastmittelverhalten hat zwar lediglich eine Sensitivität von 60%, allerdings eine Spezifität von 96-100%. Als Untersuchung zur Früherkennung von Patienten mit einem erhöhten Risiko wird die CT nicht empfohlen da sie mit Strahlenbelastung und höheren Kosten als der Ultraschall verbunden ist. Die MRT wird zur Früherkennung auch nicht eingesetzt da diese ebenfalls nicht als kosteneffizient angesehen wird (35). Bei Raumforderungen  $\geq 1\text{cm}$  in einer zirrhotischen Leber sollte zur Diagnostik primär eine MRT der Leber mit arterieller und portalvenöser Phase erfolgen (36). Bei unklaren Befunden in der Bildgebung oder Kontraindikationen für eine MRT sollte eine weitere Diagnostik mittels triphasischer CT und/oder CEUS erfolgen (23). Ist ein Leberherd allerdings  $< 1\text{cm}$  sollte dieser mit der am besten geeigneten Kontrastmittelverstärkten Schnittbildgebung in einem Intervall von 3 Monaten kontrolliert werden (23,37).



**Abbildung 1:** Diagnostikalgorithmus bei Verdacht auf ein HCC bei Patienten mit einer Leberzirrhose (23).



**Abbildung 2:** Arterielle Phase einer CT. Durch Hypervaskularisierung des HCC zeigt sich dieses hyperdens im Vergleich zum umliegenden Leberparenchym.



**Abbildung 3:** Spätphase einer CT. Es zeigt sich das für ein HCC charakteristische Wash-Out. Der Herd stellt sich hypodens im Vergleich zum umliegenden Leberparenchym dar.

#### 1.3.4 Leberbiopsie

Kann die Diagnose eines HCC anhand des typischen Musters mittels CT oder MRT gestellt werden, können Patienten direkt mit einer Therapie beginnen. Eine Bestätigung mittels Biopsie und anschließender histopathologischer Untersuchung ist nicht nötig (37–39). Ein HCC im frühen Stadium kann sich allerdings histologisch und bildmorphologisch von einem fortgeschrittenen HCC unterscheiden. Meist zeigen sich frühe Stadien aufgrund noch unreifer Neoangiogenese eher hypovaskularisiert. Außerdem sind fettgewebige Veränderungen innerhalb des Tumors häufig. Wenn das typische vaskuläre Muster in CT oder MRT nicht zu sehen ist, muss eine Biopsie durchgeführt werden, um die Diagnose eines HCCs zu sichern. Hierbei sollte aber



nicht außer Acht gelassen werden, dass die histopathologische Untersuchung ebenfalls falsch negativ sein kann. In frühen Stadien ist es zudem schwierig zwischen einem HCC im Frühstadium und einem dysplastischen Knoten zu differenzieren. Außerdem kann es immer zu einer Fehlpunktion kommen (35). Bei Patienten mit einer Leberherd  $\geq 1$  cm mit untypischen Kontrastmittelverhalten, bei denen kein kurativer Ansatz besteht jedoch eine Indikation zur Tumor gerichteten Therapie, sollte ebenfalls eine Biopsie erfolgen (23).

### **1.3.5 Labor, Child Pugh Score, Model for End-Stage Liver Disease, Eastern Cooperative Oncology Group Performance Status**

Um eine individuelle Therapieentscheidung für HCC Patienten treffen zu können sollte die Leberfunktion, die Tumorlast und der Leistungsstatus mit einbezogen werden (23). Die folgenden Scores wurden primär entwickelt, um die Leberfunktion bei Zirrhosepatienten abzuschätzen. Da die Leberfunktion allerdings ein wichtiger Faktor bei der Therapieempfehlung für HCC ist, werden diese auch in diesem Zusammenhang berücksichtigt. Die folgenden Daten zum Gesamtüberleben beziehen sich somit zunächst auf eine Leberzirrhose. Der sogenannte Child Pugh Score wird genutzt, um das Überleben von Patienten mit chronischen Lebererkrankungen oder einer Leberzirrhose einzuschätzen. Dieser Score berücksichtigt das Serum Albumin, die International Normalized Ratio (INR), Aszites und den Grad der Enzephalopathie. Je nach Ausprägung der hier genannten Parameter werden drei Gruppen von A-C unterschieden (40).

**Tabelle 1:** Kriterien Child Pugh Score (10)

Parameter	1 Punkt	2 Punkte	3 Punkte
Gesamtbilirubin (mg/dl)	< 2	2-3	>3
Serum-Albumin (g/dl)	>3,5	2,8-3,5	<2,8
INR	<1,7	1,7-2,2	>2,2
Aszites im Ultraschall	nein	leicht	mittelgradig
Hepatische Enzephalopathie	nein	Grad 1-2	Grad 3-4

Eine gute Leberfunktion liegt bei 5 und 6 Punkten vor. Dies entspricht einer Leberzirrhose im Stadium Child A. 7 bis 9 Punkte beschreiben eine mäßige Leberfunktion und entsprechen dem Stadium B. Das Stadium C entspricht einem Child-Pugh Score von 10- 15 Punkten. Hier liegt eine schlechte Leberfunktion vor. Die perioperative Mortalität liegt bei Child A bei 10%, Child B 30% und bei Child C bei >80% (10).

Der wichtigste Tumormarker für ein HCC ist das Alpha-Fetoprotein. Das AFP ist als Verlaufsparemeter ein außerordentlich nützlicher Marker. Hohe AFP Werte weisen auf ein fortgeschrittenes Tumorstadium und somit auf eine schlechte Prognose hin (35).

Der heute als Model for End-Stage Liver Disease Score (MELD-Score) bekannte Score fand seinen Anfang in der Vorhersage der postinterventionellen Überlebenszeit nach einem Transjugulären Portosystemischen Shunt (TIPS) bei Patienten mit einer Leberzirrhose. Der Score beinhaltet Prädiktoren für den Verlauf einer Lebererkrankung und bezieht sich auf ein Überlebensintervall von drei Monaten. Es wird der Serum Kreatininwert, Serum Bilirubinwert, die INR sowie die Ätiologie der Leberzirrhose berücksichtigt (41). Mittlerweile wird der MELD-Score in abgewandelter Form unter Berücksichtigung von Serum Kreatinin, Serum Bilirubin und INR in vielen Therapieentscheidungen bei Patienten mit einer Lebererkrankung angewandt. Anhand des errechneten Punktwerts, kann die Wahrscheinlichkeit, dass der Patient innerhalb von 3 Monaten ohne eine Lebertransplantation verstirbt, abgeschätzt werden (10). Im Jahr 2003 wurde der MELD Score auch von Eurotransplant eingeführt, um die

Dringlichkeit der Verteilung der Spenderorgane zu verbessern (42). Als weiterer wichtiger prognostischer Faktor wurde das Serum Natrium Level zum MELD-Score hinzugefügt zum sogenannten MELD-NA Score (43).

Zur Bestimmung des Allgemeinzustandes kann der sogenannte Eastern Cooperative Oncology Group (ECOG) Performance Status zu Rate gezogen werden. Dieser umfasst die Grade 0 bis 5. Wobei Grad 5 dem Tod entspricht und Grad 0 eine uneingeschränkten Gesundheitsleistungsstatus beschreibt (44). Eine genau Beschreibung der einzelnen Grade kann Abbildung 4 entnommen werden.

<b>GRADE</b>	<b>ECOG PERFORMANCE STATUS</b>
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all selfcare but unable to carry out any work activities; up and about more than 50% of waking hours
3	Capable of only limited selfcare; confined to bed or chair more than 50% of waking hours
4	Completely disabled; cannot carry on any selfcare; totally confined to bed or chair
5	Dead

**Abbildung 4:** ECOG Performance Status (44)

### **1.3.6 Bildgebende Verfahren zum Tumorstaging**

Nach Diagnosestellung ist der nächste entscheidende Schritt die prognostische Einschätzung der Erkrankung. Dies erfolgt über eine weitere Bildgebung. Hierbei sollte entweder ein kontrastverstärktes CT des Thorax und des Abdomens erfolgen, beziehungsweise eine native CT des Thorax wenn die kontrastverstärkte MRT das gesamte Abdomen bereits diagnostisch erfasst hat (23).

## 1.4 Einteilung der Stadien nach BCLC

Anhand der BCLC-Stadien können sowohl eine Prognose als auch die bestmögliche Therapie bestimmt werden. Die Therapieempfehlung wurde seit der ersten Beschreibung im Jahr 1999 (45) stetig aktualisiert, zuletzt im Jahr 2022 (46). In der Barcelona Clinic Liver Cancer Klassifikation werden 5 Stadien unterschieden (46). In dieser Klassifikation wird der Allgemeinzustand mittels ECOG Performance Status, das Child Pugh Stadium (Leberfunktion), die Tumorgröße, sowie vaskuläre Infiltration und extrahepatische Manifestation berücksichtigt (45).

Das BCLC-Stadium 0 beschreibt ein sehr frühes HCC-Stadium. Dies ist definiert als das Vorhandensein eines einzelnen Tumors mit einem Durchmesser von weniger als 2 cm ohne vaskuläre Invasion. Patienten mit gutem Allgemeinzustand (ECOG-0) und gut erhaltener Leberfunktion bei kompensierter oder nicht vorhandener Leberzirrhose, fehlender portaler Hypertension und normwertigem Bilirubin (Child-Pugh A) werden zu dieser Gruppe gezählt (46). Neue Studien zeigen bei Patienten mit einem solitären HCC <2cm die mit einer Resektion behandelt wurden eine Fünf-Jahres-Überlebensrate von > 80 % (47). Allerdings kann mit einer Radiofrequenz- oder Mikrowellenablation (RFA bzw. MWA) ein ähnliches Ergebnis erzielt werden. (23).

Das „**early Stage**“ HCC entspricht dem Stadium BCLC A und beschreibt Patienten in gutem Allgemeinzustand (ECOG-0 PS0) mit einem Tumor größer als 2cm oder drei Tumoren kleiner als 3 cm und erhaltener Leberfunktion (9). Diese Patienten könnten von einer radikalen Therapie profitieren. Das BCLC Stadium A wird weiter in vier Subgruppen eingeteilt.

Stadium A1: Ein Tumor, keine relevante portale Hypertension, Bilirubin normal.

Stadium A2: Ein Tumor, relevante portale Hypertension, Bilirubin normal

Stadium A3: Ein Tumor, mit einer relevanten portalen Hypertension und einem erhöhtem Bilirubinwert.

Stadium A4: Drei Tumore kleiner als 3 cm ungeachtet der Leberfunktion.

In diesem Stadium sollte eine kurative Therapie angestrebt werden. Dazu zählen die chirurgische Resektion, Lebertransplantation oder perkutane Verfahren (45).

Anhand der erfassten Kriterien sollte individuell für jeden Patienten die beste dieser

drei Therapieoptionen ermittelt werden.

Folgenden Patienten im Stadium BCLC A wird eine Resektion empfohlen: Einzelner Tumor Herd < 3cm, guter Gesundheitszustand (PS 0) und erhaltener Leberfunktion, kompensierte Leberzirrhose (Child Pugh A), MELD-Score <10, ausreichend verbleibendem Leberparenchym und der Möglichkeit für einen laparoskopischen oder minimal invasiven Zugangsweg (9).

Relevante Parameter für das Überleben nach einer Resektion sind portale Hypertension und der Bilirubinwert. Patienten ohne relevante portale Hypertension (definiert als HVPg  $\leq$  10 mmHg) und mit normwertigem Bilirubin zeigten ein gutes Überleben nach Resektion (48,49).

Eine Ablation wird Patienten mit 2-3 Tumorkläsionen  $\leq$  3 cm empfohlen, die nicht für eine Lebertransplantation in Frage kommen (9).

Die stärksten prognostischen Variablen bei lokaler Ablation sind das Child Pugh Stadium A (50–52) sowie die Tumorgröße und das Ansprechen des Tumors auf die Behandlung (53).

Eine Lebertransplantation wird nach der BCLC Klassifikation empfohlen bei Patienten mit 1-3 HCC Herden  $\leq$  3 cm (9). Die Abschätzung der Erfolgsaussichten einer Lebertransplantation erfolgt anhand der Milan-Kriterien welche sich aus den folgenden Faktoren zusammensetzen: Ein Tumorherd  $\leq$  5cm oder drei Herde  $\leq$  3cm, keine vaskuläre Infiltration und keine extrahepatische Manifestation (54,55).

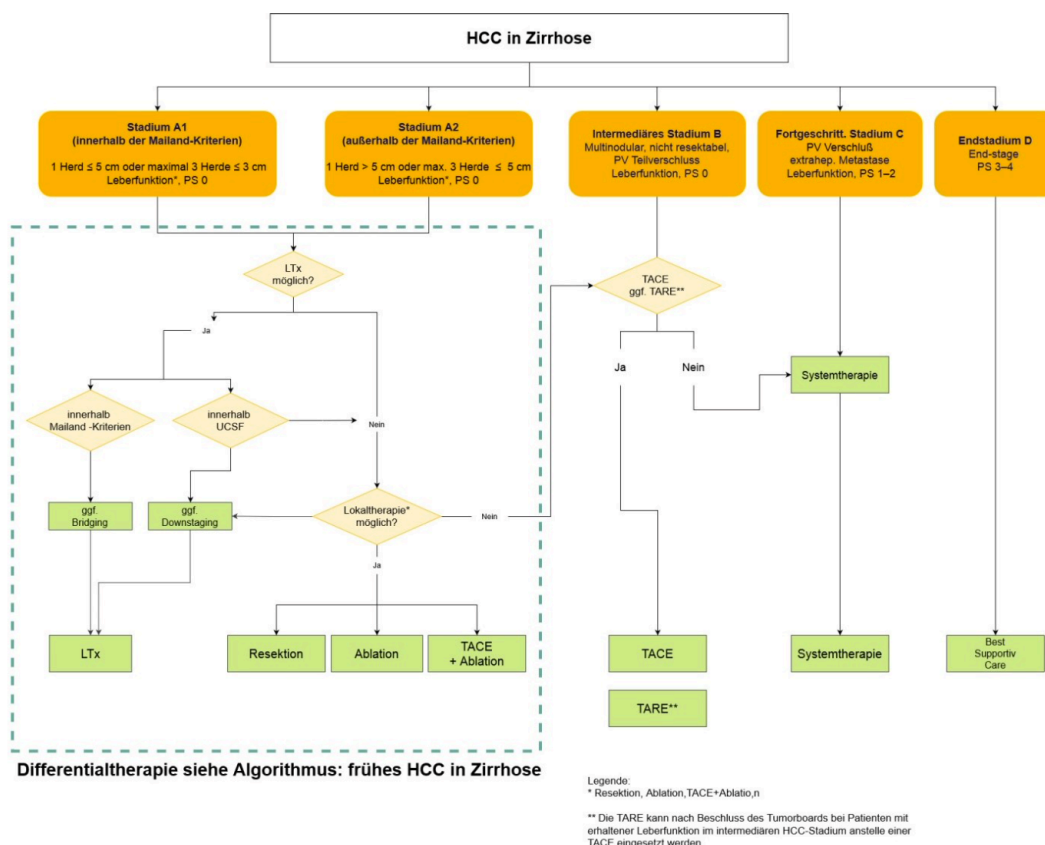
Das Stadium B wird auch als intermediate Stadium bezeichnet und beinhaltet asymptotische Patienten mit einem multinodalen Tumorleiden ohne eine vaskuläre Infiltration oder extrahepatische Manifestation. Patienten mit einem HCC im Stadium B nach BCLC sollten palliativ, oder im Rahmen randomisierter klinischer Studien behandelt werden (45). Die Methode der Wahl in diesem Stadium ist die transarterielle Chemoembolisation (TACE) (9).

Fortgeschrittene Tumorleiden werden im Stadium C zusammengefasst. Dazu zählen symptomatische Patienten mit einer Gefäßinfiltration oder extrahepatischen Metastasen und mit einem eingeschränkten Allgemeinzustand (PS1-2) (9). In diesem Stadium stehen als Erstlinientherapie eine Kombination aus Atezolizumab und Bevacizumab (56) sowie die beiden Tyrosinkinase-Inhibitoren Sorafenib (57,58) und

Lenvatinib (59) zu Verfügung.

Das Stadium BCLC D beschreibt ein HCC im Endstadium. Patienten, aus dieser Gruppe, zeigen schwere HCC-bedingte Krankheitssymptome, einen schlechten Allgemeinzustand (PS3-4), sowie eine stark eingeschränkte Leberfunktion (45).

In diesem Stadium sollte nur noch eine symptomatische Therapie erfolgen (9).



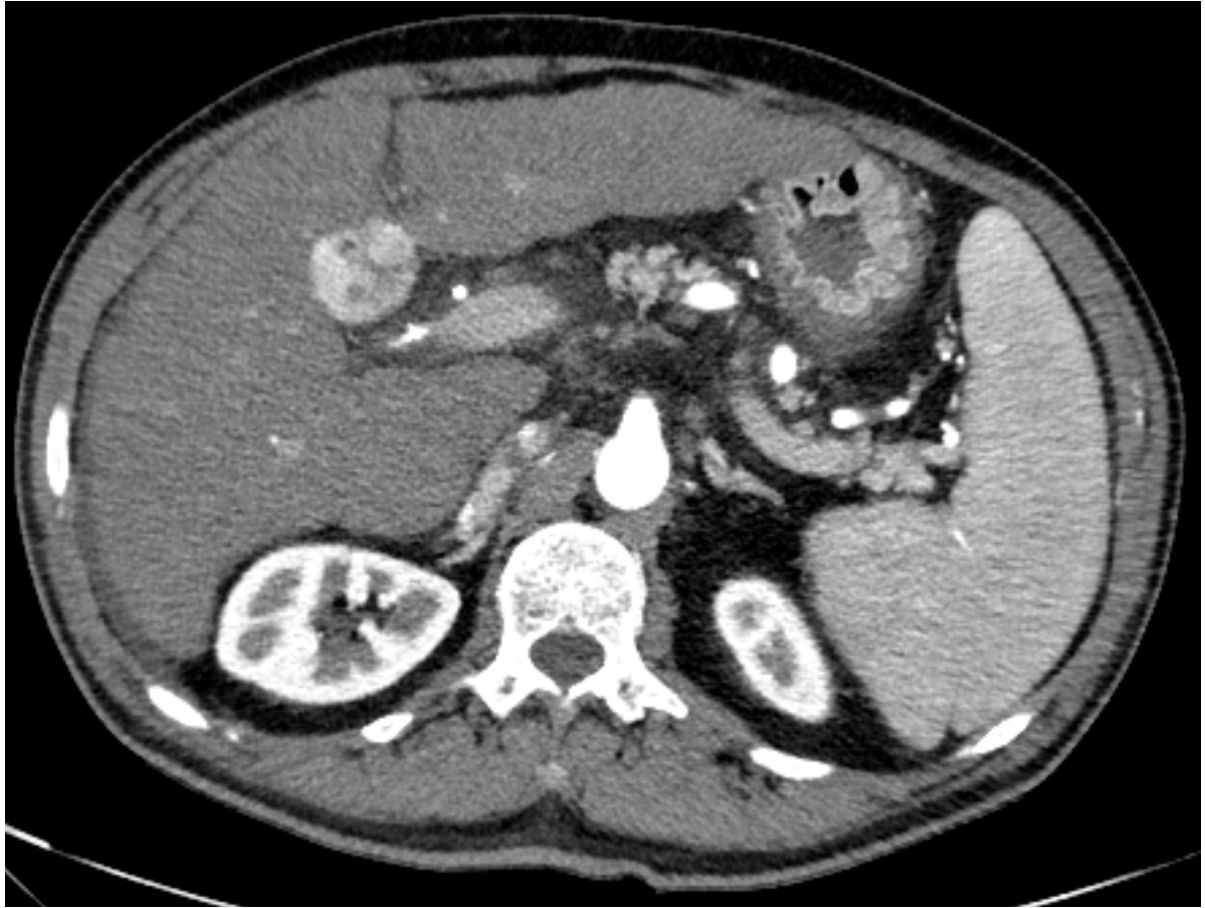
**Abbildung 5:** Therapiealgorithmus HCC bei vorliegender Leberzirrhose (23).

### 1.5 Transarterielle Chemoembolisation

Die Blutversorgung der Leber stellt eine Besonderheit im menschlichen Körper da. Das normale Leberparenchym wird zu 2/3 über die Pfortader und zu 1/3 über die Leberarterie versorgt (60). Ein HCC hingegen wird vornehmlich über die Leberarterien mit Blut versorgt (61).

Eine Transarterielle Chemoembolisation ist eine Kombination aus einer regionalen, selektiven Chemotherapie und einer transarteriellen Embolisation des Tumors. Durch die Doppelversorgung kann das Lebergewebe geschont werden, obwohl eine Ischämie des Tumorgewebes durch Embolisation über die Arterie erfolgt. Durch die Embolisation ist es außerdem möglich, dass die Chemotherapeutika länger und in höherer Dosis im Tumorgewebe verbleiben. Da durch die Embolisation nur ein geringerer Teil des Chemotherapeutikums systemisch wirken kann, ist ein weiterer Vorteil eine geringere systemische Toxizität. Um eine segmentale oder subsegmentale TACE durchführen zu können wird eine möglichst exakte anatomische Beurteilung des HCCs und seiner Gefäßversorgung benötigt. Deshalb sollte vor jeder TACE eine Bildgebung mittels einer CT oder MRT-Untersuchung erfolgen. Hierbei sollte eine möglichst detaillierte Gefäßdarstellung angestrebt werden um alle tumorversorgenden Gefäße, einschließlich der möglichen extrahepatischen Gefäße zu lokalisieren (60).

Als konventionelle TACE (cTACE) bezeichnet man die Kombination eines Chemotherapeutikums, Embolisats sowie Lipiodol. Hierbei wird Lipiodol als weiteres Embolisat, zur Tumordemarkierung und als Transportmedium für den Wirkstoff eingesetzt (62). Wird Lipiodol in die Leberarterie injiziert, verbleibt es selektiv über Wochen bis hin zu einem Jahr im Tumorknoten. Dieser „Siphoneffekt“ wird durch die Hypervaskularisierung des Tumorgewebes und das Fehlen von Kupfer-Zellen begünstigt (63,64). Ein etwas neuerer Ansatz für die TACE-Therapie sind die sogenannten Drug-Eluting-Beads (DEB). Hierbei handelt es sich um Chemotherapiebeladene Partikel welche selektiv in die tumorversorgenden Arterien injiziert werden und das Chemotherapeutikum über einen längeren Zeitraum hinweg freisetzen (65).

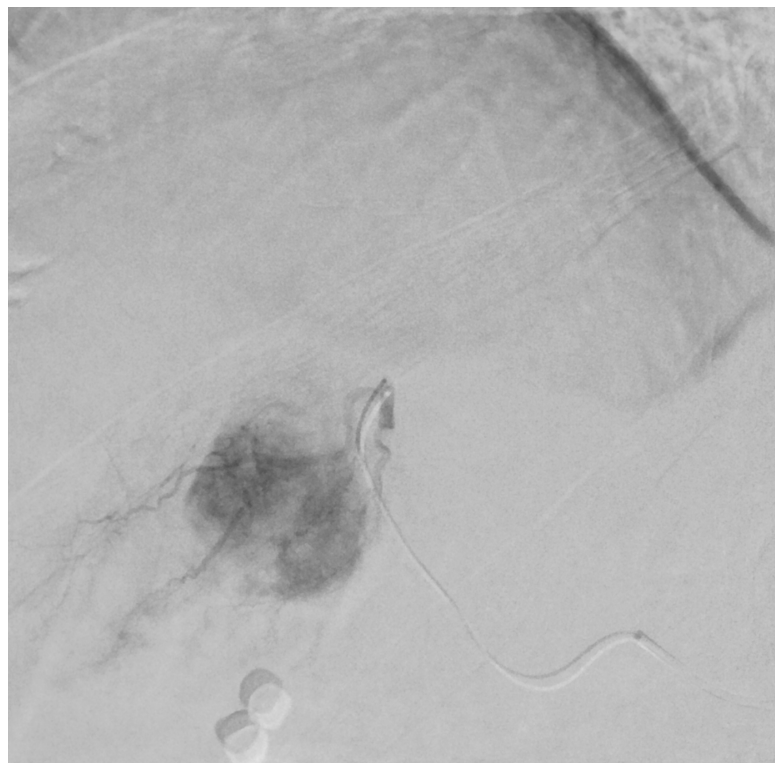


**Abbildung 6:** CT vor TACE mit arterieller Kontrastmittelphase. Es zeigt sich ein deutlich hypervaskularisierter Tumorherd.

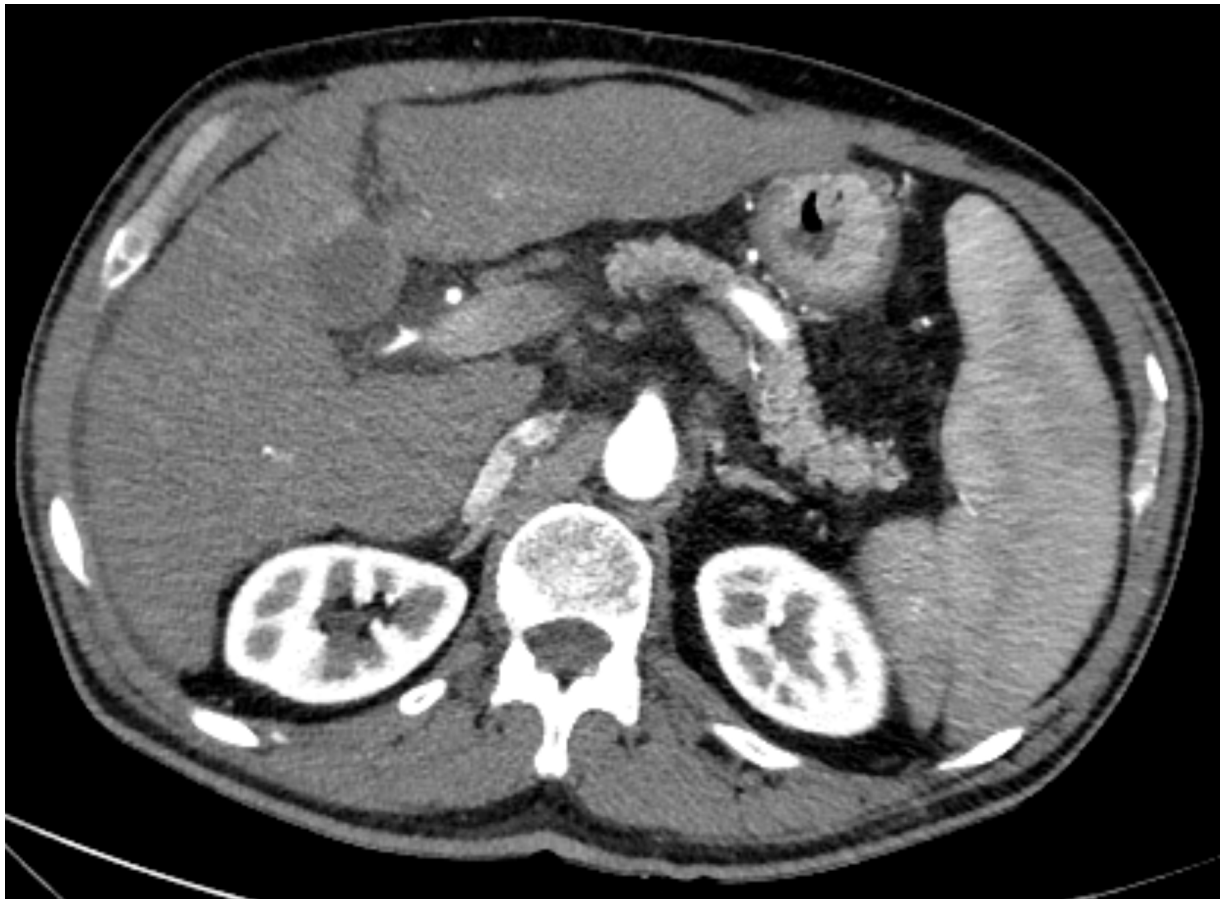




**Abbildung 7:** Digitale Subtraktionsangiographie (DSA) intrahepatischer Gefäße mit Darstellung der tumorversorgenden Arterien.



**Abbildung 8:** DSA. Superselektive Applikation des Chemoembolisats.



**Abbildung 9:** CT nach TACE mit arterieller Kontrastmittelphase. Der Tumorherd zeigt sich deutlich hypovaskularisiert nach erfolgter Chemoembolisation.

## 1.6 Statistische Definitionen

### 1.6.1 Harrell's C Index

Der 1982 von Harrell et. al eingeführte Harrell's C Index wird verwendet, um die Güte von Risikomodellen zu bestimmen, welche in der Überlebensanalyse zum Einsatz kommen. Der Harrell's C Index kann auch angewendet werden, wenn Daten zensiert wurden. In einem guten Risikomodell sollten Patienten, die mit einer hohen Wahrscheinlichkeit zeitnah erkranken beziehungsweise versterben einen hohen Risikoscore aufweisen. C-Werte können von 0 bis 1 reichen (66).

Ein Wert kleiner als 0,5 weist auf ein sehr schlechtes Modell hin. Ein Wert von 0,5 zeigt, dass ein Modell ein Ereignis nicht besser vorhersagen kann als es per Zufall möglich ist. Werte über 0,7 weisen auf ein gutes Modell hin, während Werte über 0,8 bei starken Risikomodellen erreicht werden. Bei einem Wert von 1 kann ein Modell

perfekt vorhersagen, ob bei einem Patienten ein bestimmtes Ereignis eintreffen wird oder nicht (67).

### **1.6.2 Receiver Operating Characteristics Curve und Area under the Curve receiver Characteristics Curve**

Mit Hilfe einer sogenannten Receiver Operating Characteristic Curve (ROC) können diagnostische Tests verglichen werden. Hierbei wird die Rate der richtig positiven Ergebnisse den falsch positiven Ergebnissen graphisch gegenübergestellt. Die Genauigkeit des diagnostischen Tests wird als Area under the Curve receiver Characteristics Curve (AUROC) dargestellt. Je größer die Fläche unter der Kurve ist, desto akkurater ist der Test. Die Werte der AUROC können von 0 bis 1 reichen. Ein perfekter diagnostischer Test hat eine Fläche unter der ROC Kurve von 1 (68).

### **1.6.3 Brier Score**

Der Brier Score wird genutzt um die Genauigkeit einer Wahrscheinlichkeitsvorhersage zu bestimmen. Der Brier Score kann nur für binäre beziehungsweise kategoriale Systeme genutzt werden, solange die Ereignisse als binäre Ergebnisse beispielsweise in „wahr“ und „falsch“ eingeteilt werden können. Der Brier Score reicht von 0 bis 1. Eine perfekte Wahrscheinlichkeitsvorhersage hat einen Brier Score von 0. Bei einer völlig unzutreffenden Vorhersage liegt der Score bei 1 (69). Als Integrated Brier Score (IBS) bezeichnet man das Gesamtmaß für die Genauigkeit eines Modells zu allen Zeiten (70).

## **1.7 Vorhandene Scoring Systeme/Bewertungssysteme**

Um die Entscheidung zu objektivieren, ob die TACE-Therapie weitergeführt werden sollte, wurden verschiedene Bewertungssysteme entwickelt.

### 1.7.1 Assessment for Retreatment with TACE-Score (ART-Score)

In einer Studie der Universitäten Wien und Innsbruck wurden retrospektiv Daten von 222 Patienten mit einem HCC nach multipler TACE-Therapie ausgewertet. Beobachtet wurde der Einfluss der ersten TACE auf die Tumorgröße, auf Parameter der Leberfunktion und deren Auswirkung auf das Gesamtüberleben. Anhand einer Analyse dieser Daten wurde ein Punktesystem entwickelt. Folgende Parameter wurden berücksichtigt: der Anstieg der Aspartat-Aminotransferase (AST) und des Child-Pugh-Scores sowie das Therapieansprechen (Tumor response). Diese Parameter zeigten sich als einzige unabhängige negative Faktoren.

Als Überlebenszeit wurde das Intervall ausgehend vom Tag vor der zweiten TACE bis zum Tod beziehungsweise zum letzten Follow up definiert.

Anhand einer statistischen Analyse mittels Log Rank Test wurde der Einfluss der einzelnen Parameter auf das Gesamtüberleben ermittelt. Der Anstieg der Aspartat - Aminotransferase um 25% gegenüber des Ausgangswerts ergab eine Hazard Ratio von 8,4 ( $p < 0.001$ ). Die mediane Überlebenszeit von Patienten ohne AST-Anstieg lag bei 20,2 Monaten versus 6,4 Monaten bei Patienten mit einem AST-Anstieg  $> 25\%$ . Ein Anstieg des Child Pugh um einen Punkt ergab ein Hazard Ratio (HR) von 4,4 ( $p < 0,001$ ). Ein fehlendes Ansprechen des Tumors auf die Therapie war hingegen nur mit einer Hazard Ratio von 1,7 assoziiert ( $P = 0.026$ ). Die mediane Überlebenszeit lag bei 18,8 Monaten versus 9,3 Monaten bei Patienten mit fehlender Tumor response. Diese, durch den Regressionskoeffizient errechneten, B-Werte der einzelnen Variablen, wurden mit 2 multipliziert. Dadurch ergaben sich die Punkte des ART-Scores (71).

**Tabelle 2:** Ergebnis der Regressions-Analyse der Parameter der Trainingskohorte des ART Score (71).

Variable		Overall Survival			ART Score Points*	P-value (Cox Regression)
		HR	95% CI	B		
Child-Pugh score increase	Absent	1			–	
	+ 1 point	2.0	1.2-3.5	0.71	1.5	
	+ ≥2 points	4.4	2.0-9.6	1.49	3	<0.001
AST increase >25%	Absent	1			–	
	Present	8.4	4.5-15.5	2.13	4	<0.001
Radiologic tumor response	Present	1			–	
	Absent	1.7	1.1-2.6	0.51	1	0.026

Abbreviations: HCC, hepatocellular carcinoma; AST, aspartate aminotransferase.

\*The regression coefficients (B) were multiplied by 2 and rounded in order to facilitate the bedside calculation of the ART score.

Ein Anstieg der Aspartat-Aminotransferase um 25% nach der ersten TACE im Vergleich zum Ausgangswert ergibt 4 Punkte.

Ist nach der ersten TACE der Child Pugh Score um einen Punkt angestiegen, werden 1,5 Punkte beim ART-Score vergeben. Ein Anstieg um 2 Punkte oder mehr entsprechen drei Punkten beim ART-Score.

Die Trainingskohorte bestand aus 97 Patienten. Patienten mit einem ART-Score von 0-1,5 lebten im Median 23,7 Monate (95% KI 16,2-32,2 Monate), während Patienten mit einem ART-Score  $\geq 2,5$  insgesamt 6,6 Monate (95% KI 4,5-8,8 Monate) lebten ( $p < 0,001$ ). Das Ansprechen des Tumors auf die Therapie wurde mittels CT oder MRT-Bildgebung nach den European Association for the study of the Liver (EASL) -Kriterien bestimmt. Die externe Validierung des Scores mit 107 Patienten zeigte folgende Werte: Patienten mit einem ART-Score zwischen 0 und 1,5 zeigten eine mediane Überlebenszeit von 27,6 Monaten (95% CI, 22,5-33,5 Monate), während Patienten mit 2,5 Punkten oder mehr nur 8,1 Monate im Median überlebten (95% CI, 5,7-10,5 Monate). Werden 1,5 Punkte oder weniger erreicht, sollte eine weitere TACE in Betracht gezogen werden. Patienten mit einem ART-Score von 2,5 oder mehr profitieren möglicherweise nicht von einer weiteren TACE (71).

### 1.7.2 ABCR-Score

Zur Entwicklung des ABCR-Scores wurden Daten von 139 Patienten ausgewertet, welche wegen eines HCCs von Januar 2007 bis April 2012 mittels TACE behandelt wurden. Mit Hilfe eines Regressions-Modells wurden Parameter erfasst, welche eine Prognose bezüglich des Gesamtüberlebens erlauben. Als Überlebenszeit wurde die Zeit zwischen dem Tag der zweiten TACE und dem Todesdatum beziehungsweise dem letzten Follow-Up definiert.

Zuerst wurden mittels einer univariaten Analyse 7 Variablen erfasst, welche einen signifikanten Einfluss auf die Überlebenszeit zeigten. Dazu gehörten: BCLC, AFP ( $\geq 200\text{ng/ml}$ ) vor Therapiebeginn, portale Hypertension (Ösophagusvarizen), sowie vor der zweiten TACE die Tumor response, Verschlechterung des Child Pugh Scores, und ein Anstieg des AST um mehr als 25%. Danach wurden mit Hilfe eines multivariaten Cox Regressionsmodells schrittweise unabhängige Parameter bestimmt. Parameter die nach einem Log Rank Test ein  $P < 0.05$  zeigten, wurden als signifikant angesehen. Nachdem so schrittweise Parameter entfernt wurden, konnten vier signifikante Einflussfaktoren identifiziert werden. Zwei davon werden zu Beginn der Therapie erfasst (AFP  $\geq 200\text{ng/ml}$  und BCLC-Stadium) sowie zwei im Verlauf (Zunahme des Child Pugh Score von  $\geq 2$  Punkten und die Tumor Response).

Die Punkte des ABCR-Scores ergaben sich aus einer Duplikation des Beta-Koeffizienten und Rundung auf die nächste ganze Zahl (72).

**Tabelle 3:** Multivariate Analyse der Parameter des ABCR Score nach der ersten TACE (72).

Variable	Overall survival			Score*	p value (Cox)
	Hazard ratio	95% CI	B		
<b>AFP (ng/ml)</b>					
<200	1	-	-	0	-
≥200	2.04	1.22-3.41	0.71	1	0.007
<b>BCLC</b>					
A	1	-	-	0	-
B	2.60	1.49-4.54	0.95	2	0.0008
C	3.73	1.76-7.89	1.32	3	0.0006
<b>Child-Pugh increase</b>					
Absent	1	-	-	0	-
+1 point	0.92	0.52-1.66	-0.08	0	0.79
≥2 points	3.03	1.62-5.65	1.11	2	0.0005
<b>Tumour response</b>					
Absent	1	-	-	0	-
Present	0.28	0.15-0.53	-1.26	-3	<0.0001

BCLC, Barcelona Clinic Liver Cancer; AFP, alpha-fetoprotein.

Es wurde ein Punktesystem entwickelt, welches von -3 bis +6 reicht. Drei Gruppen mit Schwellenwerten von  $\leq 0$  und  $\geq 4$  wurden unterschieden, die eine signifikant unterschiedliche Überlebenszeit aufwiesen. In der Trainingskohorte zeigten Patienten mit einem ABCR-Score von  $\leq 0$  eine mediane Überlebenszeit von 34,5 Monaten, bei 1-3 Punkten waren es 13,2 Monate und bei  $\geq 4$  Punkten nur 4,6 Monate. Es konnte ein linearer Zusammenhang zwischen der Abnahme des Gesamtüberlebens und der Zunahme des ABCR-Scores gezeigt werden. (Linearer Korrelationskoeffizient von  $R^2$  0,94)

Eine Validierung fand in zwei Gruppen mit 178 Patienten statt. Die interne Validierungskohorte mit 78 Patienten wurde auch in drei Gruppen eingeteilt. Es wurde eine mediane Überlebenszeit von 26 Monaten bei  $\leq 0$  Punkten, 12 Monaten bei 1-3 Punkten und 7 Monaten bei 4 oder mehr Punkten beobachtet.

Die externe Validierung erfolgte anhand von 100 Patienten. Hier ergab die mediane Überlebenszeit der einzelnen Gruppen: 37,8 Monate bei einem ABCR-Score  $\leq 0$ , insgesamt 17,1 Monate bei einem Score von 1-3 Punkten und 7,5 Monate bei 4 oder mehr Punkten. Ein ABCR-Score von 4 Punkten oder mehr, identifiziert Patienten, welche von einer weiteren TACE nicht profitieren (72).

Als Zielläsionen wurden Herde definiert, die eine intratumorale Kontrastmittelanreicherung in der arteriellen Phase einer CT beziehungsweise MRT zeigten. Die Tumor Response wurde anhand der EASL Kriterien erfasst (um eine Vergleichbarkeit zum ART Score zu gewähren) Dazu gehören „complete Response“ welches ein vollständiges Verschwinden von Kontrastmittelaufnehmenden beziehungsweise lipiodolspeichernden Tumoranteilen voraussetzt sowie „partial response“ mit Abnahme vitaler Tumoranteile von 50% (72,73).

### 1.7.3 SNACOR

Das Akronym „SNACOR“ steht für Tumor **S**ize, Tumor **N**umber, **A**lpha- fetoprotein, **C**hild-Pugh Stadium und **o**bjective Radiological **R**esponse. Auch in dieser Studie wurden Patienten mit einem HCC nach der ersten TACE analysiert. Als Gesamtüberleben wurde das Intervall zwischen dem Datum der ersten TACE und dem letzten Follow-Up beziehungsweise dem Todesdatum definiert. Die retrospektiv analysierten Patientendaten stammen aus Datenbanken zweier großer Leberzentren in Südkorea. Die Behandlung der Patienten fand zwischen 2006 und 2009 statt. Es wurden insgesamt 485 Patienten per Zufall in zwei Gruppen eingeteilt. 340 Patienten (70%) bildeten die Trainingskohorte, während 145 Patienten (30%) zur Validierungsgruppe zählten. Beide Gruppen zeigten eine ähnliche Verteilung der Parameter. Mittels einer Kaplan-Meier-Analyse wurde die Überlebenszeit ermittelt. Der Unterschied in den verschiedenen Gruppen wurde mit einem Log-Rank-Test berechnet. Anhand dieser Analysen wurde ein 10 Punkte Risk prediction model erstellt. Als Grenzwerte für ein erhöhtes Risiko ergaben sich folgende Werte: Eine Tumorgroße von mehr als 5cm, mehr als 4 Tumore sowie ein AFP-Level  $\geq 400$  ng/ml, Child Pugh B versus A sowie objective radiologic response nach den modified Response Evaluation Criteria in Solid Tumors (mRECIST). Variablen mit  $P < 0,10$  wurden in das Risk prediction model einbezogen. Der Einfluss der einzelnen



Parameter wurde durch ein multivariates Cox-Proportional-Hazard-Modell ermittelt. Der Regressionskoeffizient (Beta Koeffizient) wurde durch den niedrigsten Koeffizienten geteilt, welcher in diesem Fall dem der Tumorgroße (<5cm vs. ≥5cm) entsprach. Der sich daraus ergebende Wert ergab gerundet den Punktwert des Scores (74).

**Tabelle 4:** Beta Koeffizient mit zugehörigem Risiko Score der Trainingskohorte des SNACOR (74).

Variables	$\beta$ coefficient	P-value	Risk score
Tumour size		0.1000	
<5 cm			0
≥5 cm	0.257		1
Tumour number		0.001	
<4			0
≥4	0.522		2
Baseline AFP level		<0.001	
<400 ng/ml			0
≥400 ng/ml	0.783		3
Child–Pugh class		0.074	
A			0
B	0.362		1
Response evaluation		<0.001	
CR + PR			0
SD + PD	0.807		3

AFP, alpha-foetoprotein; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

Es wurden drei Risikogruppen unterschieden: „Low Risk“ 0 bis 2 Punkte, „intermediate Risk“ 3 bis 6 Punkte und „high risk“ 7-10 Punkte.

In der Gruppe der Low Risk Patienten lag die mediane Überlebenszeit bei 49,8 Monaten, gefolgt von der Intermediate Risk Gruppe mit 30,7 Monaten und der High Risk Gruppe mit 12,4 Monaten.

Um die Güte der Vorhersage des Modells zu bestimmen, wurden für jeweils 1, 3 und

6 Jahres Überleben Areas under the receiver operating curve erstellt. Es ergaben sich AUROCs bei einem Konfidenzintervall von 95% von 0,76 (0,687-0,829) für ein Jahr, 0,754 (0,696-0,810) für 3 Jahre und 0,742 (0,658-0,818) für 6 Jahre in der Trainingskohorte. In der Validierungsgruppe wurden AUROCs von 0,712 (0,598-0,820), 0,750 (0,651-0,839) und 0,812 (0,77-0,943) erreicht. Der SNACOR wurde in dieser Studie nur darauf ausgelegt zwischen TACE 1 und 2 eingesetzt zu werden. Es wurden nur Patienten in die Studie eingeschlossen, welche eine cTACE erhielten (74).

## **1.8 Bewertung der Tumor response**

Zur Erfassung der Tumor Response gibt es zwei anerkannte Systeme. Zum einen die Bewertung der Tumor response nach mRECIST und zum anderen die Kriterien der European Association for the Study of the Liver.

### **1.8.1 Bewertung der Tumor response nach mRECIST**

Die Erfassung der Tumor Response nach mRECIST funktioniert wie folgt.

Es sollte entweder eine CT oder eine MRT mit Kontrastmittel erfolgen. Um die Untersuchungen gut vergleichen zu können ist es wichtig, dass möglichst das gleiche Untersuchungsprotokoll verwendet wird. Dieses sollte mindestens zwei Phasen beinhalten. Eine arterielle sowie eine portalvenöse Phase. Außerdem ist auf eine lückenlose Rekonstruktion zu achten (75).

Als Zielläsionen werden nur bis zu zwei Herde in der Leber gezählt (73), hier ist der größte Diameter des vitalen Tumoranteils zu messen. Die Durchmesser sollten in einer Raumbene mindestens 1 cm betragen. Die Messungen sollten reproduzierbar sein.

Anhand der mRECIST Kriterien wird die Tumor Response der Zielläsionen wie folgt definiert:

**„Complete Response“ (CR):** Verschwinden jeglicher arterieller Kontrastmittelanreicherungen in den Zielläsionen.

**„Partial Response“ (PR):** Eine Abnahme der Summe der Durchmesser der vitalen, kontrastmittelaufnehmenden Zielläsionen in der arteriellen Phase um mindestens

30%. Als Referenz dient hier die Summe aller Durchmesser der Zielläsionen im Ausgangsbefund.

**„Progressive Disease“ (PD):** Zunahme der Summe der Durchmesser um mindestens 20% bezogen auf kontrastmittelaufnehmende Zielläsionen im Vergleich zum Ausgangsbefund.

**„Stable Disease“ (SD):** Alle Veränderungen die weder als Partial Response noch als Progressive Disease gewertet werden können.

Die Messung sollte jeweils in der arteriellen Phase erfolgen. In dieser Phase ist der Kontrast zwischen vitalem, vaskularisiertem Tumorgewebe und nicht anreicherndem nekrotischem Gewebe am größten.

Alle weiteren Läsionen sollten als nontarget lesions im Ausgangsbefund gezählt werden. Auch wenn diese Herde im Verlauf nicht gemessen werden müssen, sollte ein Fortbestehen beziehungsweise Verschwinden dieser Läsionen dokumentiert werden. Die Response der Nicht-Zielläsionen wird wie folgt definiert:

**„Complete Response“ (CR):** Verschwinden von intratumoralen arteriellen Kontrastmittelanreicherungen in den nontarget lesions

**„Incomplete Response/Stable Disease“:** Fortbestehen der intratumoralen arteriellen Kontrastmittelaufnahme einer oder mehrerer nontarget lesions

**Progressive Disease (PD):** Eine oder mehrere neuaufgetretene nontarget lesions (75).

Soll die Overall Response bewertet werden, müssen sowohl Zielläsionen als auch Nicht-Zielläsionen berücksichtigt werden (73).

### **1.8.2 Bewertung der Tumor Response nach EASL**

Laut der Kriterien der EASL werden Tumorkläsionen in der arteriellen Phase einer CT oder MRT erfasst. Gemessen wird das Produkt des bi-dimensionalen Durchmessers des kontrastmittelaufnehmenden Bereichs. Als Zielläsionen werden alle messbaren kontrastmittelaufnehmenden Herde in der Leber gewertet. Tumor Response wird nach EASL wie folgt definiert:

„Complete Response“ (CR): komplettes Verschwinden jeglicher Kontrastmittelanreicherungen in allen messbaren Leberläsionen.

„Partial Response“ (PR): Verringerung um 50% der Summe im Vergleich zum Ausgangswert. Als Referenz zählt die Summe der Produkte der bi-dimensional gemessenen Durchmesser der arteriell anreichernden Anteile der Herde.

„Progressive Disease“ (PD): Zunahme der Gesamtdurchmesser um mindestens 25%. Als Referenz zählt der kleinste Gesamtdurchmesser der Messungen seit Beginn der Therapie.

„Stable Disease“ (SD): Alle Messungen die weder zu Partial Response oder Progressive Disease zählen (73).

**Tabelle 5:** Definition der Tumor Response nach EASL und mRECIST (73)

	EASL	RECIST 1.1	mRECIST
Complete Response (CR)	Disappearance of any intratumoral arterial enhancement in all measurable arterially-enhancing liver lesions.	Disappearance of all target lesions (up to two measurable liver lesions).	Disappearance of any intratumoral arterial enhancement in all target lesions (up to two measurable liver lesions).
Partial Response (PR)	At least a 50% decrease in the sum of the product of bi-dimensional diameters of viable (enhancement in the arterial phase) target lesions, taking as reference the baseline sum of the diameters of target lesions.	At least a 30% decrease in the sum of the greatest unidimensional diameters of target lesions, taking as reference the baseline sum of the diameters of target lesions.	At least a 30% decrease in the sum of unidimensional diameters of viable (enhancement in the arterial phase) target lesions, taking as reference the baseline sum of the diameters of target lesions.
Stable Disease (SD)	Any cases that do not qualify for either partial response or progressive disease.	Any cases that do not qualify for either partial response or progressive disease.	Any cases that do not qualify for either partial response or progressive disease.
Progressive Disease (PD)	An increase of at least 25% in the sum of the diameters of viable (enhancing) target lesions, taking as reference the smallest sum of the diameters of viable (enhancing) target lesions recorded since treatment started.	An increase of at least 20% in the sum of the diameters of target lesions, taking as reference the smallest sum of the diameters of target lesions recorded since treatment started.	An increase of at least 20% in the sum of the diameters of viable (enhancing) target lesions, taking as reference the smallest sum of the diameters of viable (enhancing) target lesions recorded since treatment started.

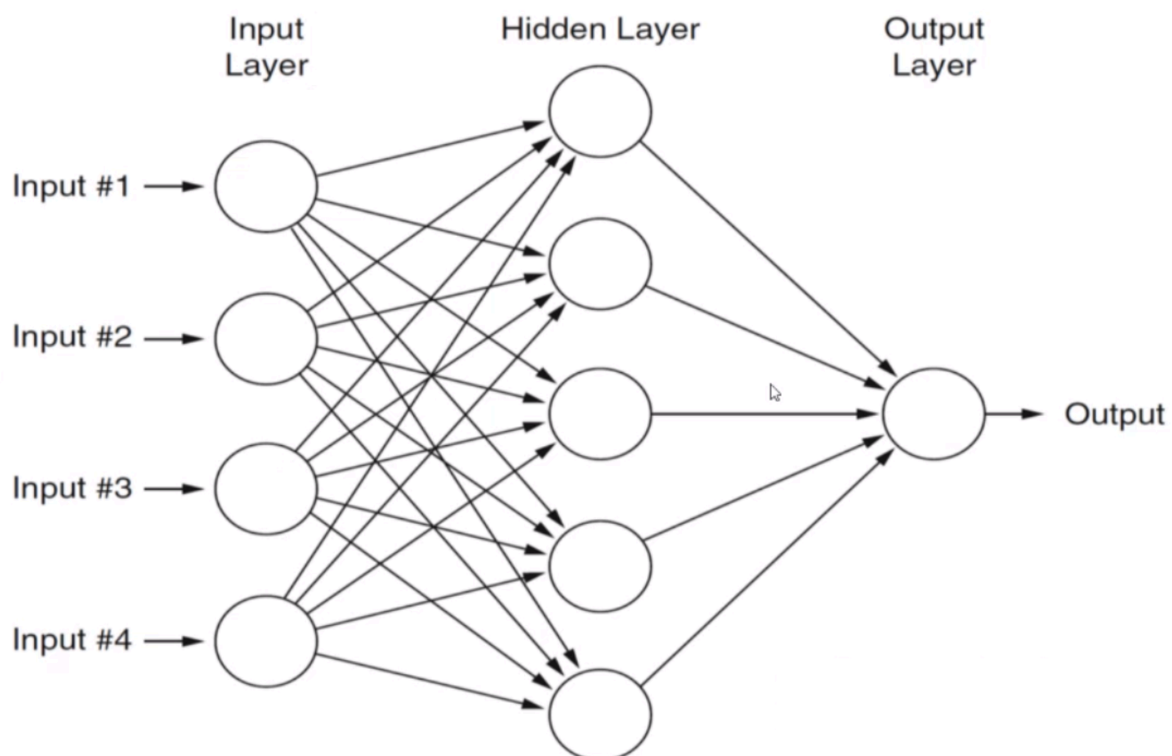
## 1.9 Artificielle neuronale Netzwerke

Bei dem Versuch, das menschliche Gehirn zu beschreiben, beziehungsweise zu verstehen auf welche Weise der Mensch etwas wahrnimmt, denkt, versteht und in der Lage ist, dieses Wissen wieder abzurufen, beschrieb Frank Rosenblatt 1958 das sogenannte Perzeptron (76).

Bei diesem Perzeptron handelt es sich um ein mathematisches Modell, welches ein künstliches Neuron beschreibt. Dieses Neuron besteht aus einem Ausgang, mehreren Eingängen und kann, je nach Eingabe, entweder aktiv oder inaktiv sein. Die Aktivität des Neurons wird hierbei über einen Schwellenwert und unterschiedliche Wichtungen der Eingänge bestimmt. Durch den Zusammenschluss multipler Perzeptrene, also Neurone, ergibt sich ein neuronales Netz. Dieses kann mit einem bekannten Datensatz trainiert werden. Der Lernvorgang beruht auf Veränderungen der Wichtungen der einzelnen Eingänge, welche so zu unterschiedlichen Ausgabewerten führen.

Weiterhin unterscheiden sich einlagige von mehrlagigen neuronalen Netzen.

Bei mehrlagigen neuronalen Netzen liegen die künstlichen Neurone in unterschiedlichen Schichten und sind untereinander über Knotenpunkte vernetzt (77).



**Abbildung 10:** Schema neuronales Netz (78)

## 1.10 Problem

Nach den Kriterien der BCLC Klassifikation wird Patienten mit einem HCC im Stadium B eine TACE als Therapie der Wahl empfohlen (45). Da eine Therapie mittels TACE bei HCC-Patienten meist aus mehreren Interventionen in einem Abstand von 6-8 Wochen besteht, ist es häufig komplex abzuschätzen, ob ein Patient von einer weiteren TACE profitiert.

Bei Patienten mit einer schlechten Prognose sollte über eine andere Therapieoption nachgedacht werden, da diese ggf. nicht mehr von einer erneuten TACE profitieren. In diesem Zusammenhang ist eine Vorhersage des weiteren Überlebens entscheidend.

Um diese Entscheidung zu erleichtern, wurden die oben genannten Systeme entwickelt.

Bei dem Versuch der externen Validierung des ART – sowie des ABCR-Score und des SNACOR zeigten sich, wie folgt, schlechte bis mäßige Vorhersagewahrscheinlichkeiten (79–81).

Für die externe Validierung des ART und ABCR-Scores wurden 176 Patienten, welche in den Jahren von 2000 bis 2015 eine TACE-Therapie aufgrund eines HCCs erhielten, analysiert. 34 der Patienten (19%) hatten ein Krankheitsstadium A nach BCLC und 81% (142) Patienten befanden sich im Stadium B. 39 Patienten (22%) hatten einen erhöhten AFP-Wert. Für diese Patienten wurde das Gesamtüberleben berechnet und so die Scores mittels Harrell's C Index, Integrated Brier score (IBS) und prediction error curves verglichen.

Vor der zweiten TACE zeigten 22 Patienten (12%) einen Anstieg des Child-Pugh Scores um einen Punkt. 51 Patienten (22%) hatten einen Anstieg um 2 Punkte. Somit war bei der Mehrheit der Patienten keine Veränderung des Child-Pugh-Scores zu beobachten. Einen Anstieg der Aspartat-Aminotransferase um 25 % zeigten 31 Patienten (23%). 114 Patienten (65%) zeigten ein Ansprechen des Tumors auf die Therapie. Demnach hatten vor der zweiten TACE 127 Patienten (72%) einen niedrigen ART-Score (0-1,5 Punkte) und 49 (28%) einen hohen ART-Score (2,5-8 Punkte). Patienten mit einem niedrigen ART-Score zeigten ein medianes Überleben von 20,8 Monaten, während sich für Patienten mit einem hohem ART-Score ein medianes Überleben von 15,3 Monaten ergab. Nach dem ABCR-Score hatten 158 Patienten

(90%) einen niedrigen ABCR-Score ( $\leq 2$  Punkte) und 18 (10%) einen hohen Score ( $\geq 3$  Punkte). In der Gruppe der Patienten mit einem niedrigen ABCR-Score lag das mediane Überleben bei 20,2 Monaten und bei 9,8 Monaten bei Patienten in der Gruppe mit einem hohen Score (79). Da in der ursprünglichen Publikation des ABCR Scores die größte Trennschärfe durch eine weitere Unterteilung erreicht werden konnte (72), wurden auch hier drei Untergruppen eingeteilt. Daraus ergab sich für 114 Patienten (82%) ein niedriger ABCR-Score ( $\leq 0$  Punkte), für 57 Patienten (32%) ein mittlerer (Intermediate) Score (1-3 Punkte) und für 5 Patienten (3%) ein hoher Score ( $\geq 4$  Punkte). Das jeweilige mediane Überleben der drei Gruppen lag bei 24,5 Monaten (95% CI, 20.1–27.4 M), 15,8 Monaten (95% CI, 12,5- 17,9 M), und 5,8 Monaten (95% CI, 2,1- 9,8 M).

Für den ART-Score wurde ein Harrell's C Index von 0.572 berechnet. Somit wäre der Score lediglich minimal genauer als die Ratewahrscheinlichkeit per Münzwurf. Der Integrated Brier Score für die ersten 6 Jahre ergab einen Wert von 0,135. Für den ABCR-Score konnte ein Harrell's C Index von 0,543 berechnet werden und ein IBS für die ersten 6 Jahre von 0,136. Für die Unterteilung in drei Subgruppen ergab sich ein Harrell's C Index von 0,608 und ein IBS (für die ersten 6 Jahre) von 0,128. Im Vergleich zur nicht stratifizierten Stichprobe auf der Grundlage von Kaplan-Meier-Schätzungen betrug der IBS 0,138.

Mittels Cox-Regressionsanalyse konnten drei Faktoren mit signifikantem prognostischem Wert identifiziert werden. Zum einen das BCLC-Stadium zum Zeitpunkt der Diagnose und die radiologische „Tumor response“ nach der ersten TACE. Weiter war ein Anstieg des Child Pugh Score um 1, 2 oder  $\geq 2$  Punkten nach der ersten TACE signifikant. In Bezug auf das Überleben konnte kein signifikanter Zusammenhang zu einem AFP-Wert  $\geq 200$  ng/ml zum Diagnosezeitpunkt oder einem Anstieg der Aspartat-Aminotransferase um mehr als 25% nach der ersten TACE hergestellt werden (79).

**Tabelle 6:** Ergebnisse der externen Validierung des ART und ABCR Score (79).

Score/Parameter	HR (95% CI)	P Value
<b>ART</b>		
Increase in Child–Pugh score after first chemoembolization		
1 point	1.63 (1.01–2.62)	.0455
≥ 2 points	2.55 (1.28–5.09)	.0080
Absence of radiologic tumor response after first chemoembolization	2.10 (1.47–2.99)	<.0001
AST increase by > 25% after first chemoembolization	1.39 (0.89–2.16)	.1507
<b>ABCR</b>		
BCLC stage at baseline	2.00 (1.23–3.26)	.0053
AFP ≥ 200 ng/mL at baseline	1.99 (0.75–1.63)	.6802
Increase in Child–Pugh score by ≥ 2 after first chemoembolization	3.35 (1.58–7.13)	.0016
Positive radiologic tumor response after first chemoembolization	0.58 (0.40–0.84)	.0034

ABCR =  $\alpha$ -fetoprotein, Barcelona Clinic Liver Cancer, Child–Pugh, and response [score]; ART = Assessment for Retreatment with Transarterial Chemoembolization [score]; AST = aspartate aminotransferase; BCLC = Barcelona Clinic Liver Cancer; CI = confidence interval; HR = hazard ratio; OS = overall survival.

Um eine eindeutige klinische Entscheidung zu treffen, waren beide Scores mit ihren begrenzten Vorhersagequalitäten nicht ausreichend (79).

Beim Versuch den SNACOR zu validieren wurden Daten von 268 HCC-Patienten ausgewertet. Jeder Patient wurde vor der ersten TACE mit einem Kontrastmittel verstärkten CT oder MRT untersucht. Sechs Wochen nach der ersten TACE sowie vor der zweiten TACE wurde die CT beziehungsweise MRT-Untersuchung wiederholt. Diese Untersuchungen wurden als Basis zur radiologischen Bestimmung der Tumor response genutzt. Diese wurde mittels der eindimensionalen EASL-Kriterien bestimmt (80). Der SNACOR Score wurde, nach den Vorgaben der ursprünglichen Arbeit (74) für jeden Patienten zum Zeitpunkt der Diagnosestellung sowie am Tag vor der zweiten TACE berechnet. Von den inkludierten 268 Patienten waren 94 (35,1%) in der Gruppe mit dem geringsten Risiko (0-2 Punkte). 144 Patienten (53,7%) waren in der Intermediate Risk Gruppe (3-6 Punkte) und 30 Patienten (11,2%) wurden der Gruppe mit dem höchsten Risiko zugeordnet. Die mediane Überlebenszeit betrug in der Low Risk Gruppe 31,5 Monate (95% KI 23,1- 46,0 Monate), in der Intermediate Risk Group waren es 19,9 Monate (95% KI 17,1-26,2) und in der Hochrisikogruppe waren es 9,2 Monate (95% KI 6,2-21,7) (80).



**Tabelle 7:** Vergleich des Gesamtüberlebens der Validierungsstudie im Vergleich zur ursprünglichen SNACOR Studie (80).

SNACOR, 3 subgroups	Low risk, 0–2 points	Intermediate risk, 3–6 points	High risk, 7–10 points	P-value
This study: median OS (95% CI), m	31.5 (23.1–46.0)	19.9 (17.1–26.2)	9.2 (6.2–21.7)	< 0.001
Original SNACOR study: median OS (95% CI), m	49.8 (34.3–65.3)	30.7 (25.8–35.6)	12.4 (5.9–18.9)	< 0.001

Zur statistischen Analyse wurde als Gesamtüberleben der Zeitraum zwischen dem Tag vor der ersten TACE und dem Todesdatum beziehungsweise dem Datum des letzten Follow-Ups definiert. Kaplan Meier Überlebenskurven wurden mit R 3.4.2 erstellt. (A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, [https:// www.R-project.org](https://www.R-project.org); accessed 2017). Die Überlebenszeit der einzelnen Kurven wurden mit Hilfe eines log-rank Test verglichen (80). Um den Vergleich zur ursprünglichen Arbeit (74) zu ermöglichen, wurden zu bestimmten Zeitpunkten Area under the Curve Receiver Operating Characteristics bestimmt. Um alle Parameter des SNACOR Scores zu analysieren und um unabhängige Prädiktoren für das Überleben zu identifizieren, wurden multivariate Analysen durchgeführt. Weiterhin wurden Hazard Ratios und die zugehörigen 95% Konfidenzintervalle (KI) berechnet (80).

**Tabelle 8:** Proportionales Risikomodell zur Ermittlung unabhängiger Prädiktoren für das Überleben. Vergleich der Daten von Patienten mit hepatozellulärem Karzinom aus der ursprünglichen Studie mit den Daten dieser Studie (80).

	SNACOR parameters		Hazard ratio (95% CI)	P-value
This study	Tumour size	≥5 cm vs. < 5 cm	2.51 (1.88–3.36)	< 0.001
	Tumour number	≥4 vs. < 4	1.28 (0.93–1.75)	0.127
	Baseline AFP level	≥400 ng/ml vs. < 400 ng/ml	1.76 (1.28–2.43)	< 0.001
	Child-Pugh class	A vs. B	1.56 (1.16–2.12)	0.004
	Objective radiological response	CR + PR vs. SD + PD	0.97 (0.73–1.28)	0.821
Original SNACOR study	Tumour size	≥5 cm vs. < 5 cm	1.29 (0.95–0.17)	0.100
	Tumour number	≥4 vs. < 4	1.68 (1.24–2.28)	0.001
	Baseline AFP level	≥400 ng/ml vs. < 400 ng/ml	2.09 (1.55–2.82)	< 0.001
	Child-Pugh class	A vs. B	1.44 (0.96–2.14)	0.074
	Objective radiological response	CR + PR vs. SD + PD	2.24 (1.65–3.03)	< 0.001

Die AUC für das Gesamtüberleben für ein Jahr, drei und sechs Jahre ergaben jeweils einen Wert von 0,641, 0,633 und 0,609. Der Harrell's C Index ergab einen Wert von 0,59. Der IBS für die ersten sechs Jahre lag bei 0,175. Bei Verwendung von Kaplan-Meier-Schätzungen der nicht stratifizierten Stichprobe lag der IBS bei 0,184. Bei der Cox-Regressionsanalyse zeigten nur die Tumorgröße, der Ausgangswert des AFP und die Child-Pugh Klasse einen signifikanten prognostischen Wert (HR=2.51,  $P < 0.001$ ; HR=1.76,  $P < 0.001$ ; HR=1.56,  $P = 0.004$ ). Objective radiological response (HR = 0.97,  $P = 0.821$ ), und die Anzahl der Tumorherde (HR = 1.28,  $P = 0.127$ ) waren keine zusätzlichen unabhängigen Prädiktoren für das Überleben. Auch wenn das SNACOR-System eine gewisse Aussagekraft zur Unterscheidung zwischen Patienten mit einer günstigen Prognose nach TACE und Patienten mit einer schlechten Prognose zeigt, war dieses System allein nicht ausreichend genau, um verschiedene prognostische Gruppen zuverlässig zu unterscheiden. Somit reicht der SNACOR alleine nicht aus um eine eindeutige klinische Entscheidung zu treffen (80).

### **1.11 Zielsetzung**

Die bisherigen Scoring Systeme (ABCR, ART, SNACOR) beruhen auf konventionellen statistischen Modellen. Da die prognostische Genauigkeit dieser Systeme alleine keine adäquate Hilfestellung zur Entscheidungsfindung ist (79–81), wurde ein neuer Ansatz gewählt. Um die Entscheidung über die weitere Therapie bei HCC-Patienten zu erleichtern, wurde ein künstliches, neuronales Netzwerk (KNN) erstellt, welches durch maschinelles Lernen die Überlebenszeit der Patienten vorhersagt.

In den vergangenen Jahren wurden vermehrt machine learning Algorithmen und KNNs erstellt, um Vorhersagen aller Art zu treffen. Künstliche neuronale Netzwerke zeigen vielversprechende Ergebnisse im Vergleich zu konventionellen statistischen Herangehensweisen.

Beispielsweise konnten verglichen mit dem MELD-Score, mit KNNs genauere Angaben zur Überlebenszeit gemacht werden (82).

Außerdem zeigten KNNs sowohl in der präoperativen Vorhersage des Tumor Gradings als auch der mikrovaskulären Invasion im Vergleich zu konventionellen linearen Modellen eine höhere Genauigkeit (83).

Das Ziel dieser Arbeit ist es, ein neues System zu entwickeln, welches die Therapieentscheidung bezüglich weiterer TACE bei HCC-Patienten erleichtern soll. Weiterhin war das Ziel, diesen neuen Ansatz mit den herkömmlichen Systemen zu vergleichen.

## **2 Methoden und Material**

Vor Beginn der Studie wurde bei der Ethikkommission der Rheinland-Pfälzischen Landesärztekammer die Zustimmung eingeholt (Genehmigungsnummer 2018-13619). Bei dieser Studie handelt es sich um eine retrospektive Kohortenstudie.

Alle Patientendaten sowie klinische Informationen wurden vor Studienbeginn pseudonymisiert. Einige Daten, stammen aus der Clinical Registry Software der Universitätsmedizin Mainz (84).

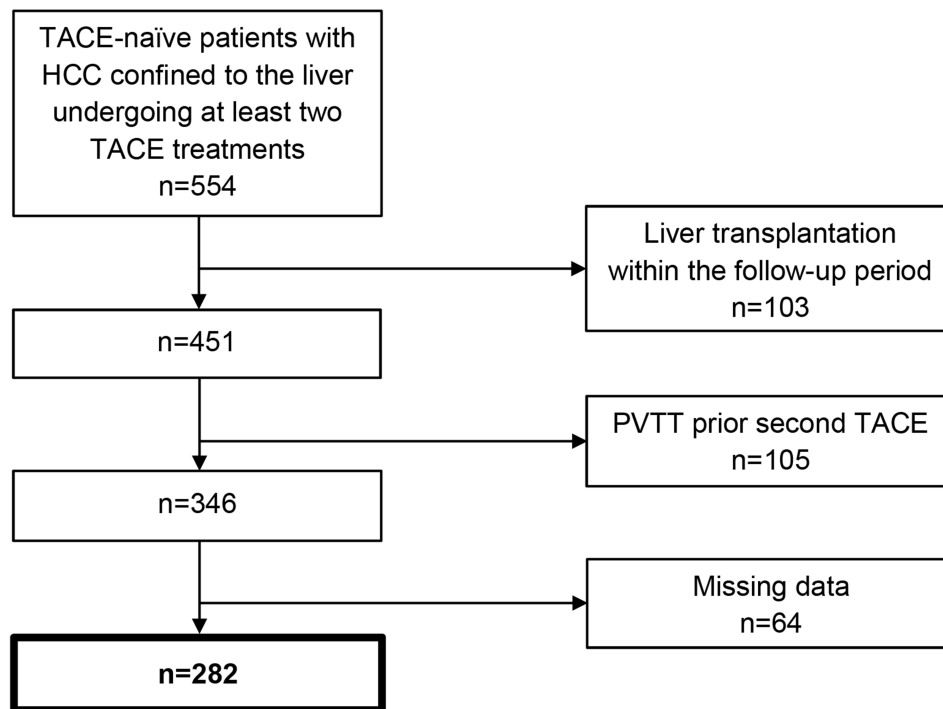
Besonderer Fokus wurde bei dieser Arbeit auf die Tansparent reporting of a multivariate prediction model for induvidual prognosis or diagnosis (TRIPOD) Richtlinien gelegt (85).

### **2.1 Patientenkollektiv**

Zu Beginn der Studie wurde zunächst eine Abfrage der Datenbank durchgeführt. Diese ergab, dass von Januar 2005 bis Dezember 2017 860 HCC - Patienten mittels TACE in der Klinik und Poliklinik für Diagnostische und Interventionelle Radiologie der Universitätsmedizin Mainz behandelt wurden. Da mindestens ein Jahr Follow-Up gegeben sein sollte, wurde der 31.12.2018 als finales Datum der Datenerfassung festgelegt.

Es wurden nur TACE-naive Patienten mit einem ausschließlich auf die Leber begrenzten HCC in die Studie aufgenommen, welche mindestens zwei TACE-erhielten. Ausgeschlossen wurden Patienten, welche eine Portalvenenthrombose vor der zweiten TACE aufwiesen oder eine Lebertransplantation innerhalb des ersten Jahres nach der zweiten TACE erhielten. Weiterhin wurden alle Patienten ausgeschlossen bei denen Parameter fehlten.

Anhand dieser Kriterien wurden 282 Patienten in die Studie aufgenommen.



**Abbildung 11:** Drop-Out Kriterien (86)

Die Sicherung der Diagnose des HCC erfolgte wenn nötig histologisch, ansonsten radiologisch anhand der Kriterien der American Association for the Study of Liver Diseases (AASLD) beziehungsweise der European Association for the Study of the Liver (9,87). Die Durchführung der TACE erfolgte wie in folgendem Artikeln beschrieben (88,89).

Um die Vergleichbarkeit zu den bekannten, konventionellen Systemen zu gewährleisten, wurde ein Intervall zwischen dem Tag vor der zweiten TACE, bis zum Tod des Patienten, respektive bis zum letzten Follow-Up festgelegt.

## 2.2 Parameter

Die Daten stammen aus der Labordatenbank der Klinik und der klinischen Registry-Software, welche speziell für die Charakterisierung von Patienten mit einem HCC entwickelt wurde (84).

Um das Gesamtüberleben möglichst genau vorherzusagen, wurden insgesamt 46 Parameter erfasst.

Dazu zählen grundlegende, demographische Angaben wie Alter und Geschlecht.

Es wurde unterschieden zwischen einer cTACE mit Mitomycin C und Lipiodol und einer TACE mittels Partikeln, welche das Chemotherapeutikum langsam abgeben (DEB-TACE) (62,65). Die jeweils vor der Therapie durchgeführte Bildgebung wurde in CT und MRT unterteilt.

Anhand der bekannten Scoring-Systeme wurden weitere Werte inkludiert, welche BCLC-Stadium, Alpha-Fetoprotein-Level, Aspartat-Aminotransferase, Tumorgröße und Tumoranzahl, Child-Pugh Stadium und das Ansprechen des Tumors auf die Therapie beinhalteten.

Bei allen Patienten erfolgte eine CT oder MRT-Untersuchung vor der ersten und zweiten TACE. Anhand dieser Untersuchungen wurde radiologisch das Ansprechen des Tumors auf die Therapie bestimmt. Dies erfolgte anhand der mRECIST Kriterien und wurde unterteilt in CR, PR, SD und PD (75).

Als weiterer klinisch relevanter Parameter wurde die Ätiologie des HCCs erfasst und in Alkoholabusus, Hepatitis B, Hepatitis C und nichtalkoholische Steatohepatitis unterteilt (90). Weiterhin wurden Komorbiditäten wie Nikotinabusus (91), Übergewicht (92), und Diabetes (90), sowie die Wachstumsform des Tumors (93) erfasst. Zur Bewertung der Leberfunktion wurden MELD-Score (94), Bilirubin (94,95), Albumin (95) und die INR (94) einbezogen. Miteinbezogen wurden weitere Parameter, die zur TACE-Therapie erhoben wurden. Zu diesen zählen die Alanin-Aminotransferase (96) und die Thrombozytenzahl (97). Natrium-Level (98) und Sarkopenie (92) wurden ebenfalls mit einbezogen. Zur Erfassung der Sarkopenie wurde der Skelettmuskel-Index berechnet (92). Dafür wurde jeweils auf Höhe des dritten Lendenwirbelkörpers die Rückenmuskulatur gemessen. Die entsprechenden Volumetrien wurden mit Hilfe des Picture Archiving and Communication Systems (Sectra R, Linköping, Schweden)

durchgeführt. Alle Werte wurden sowohl vor der ersten TACE, als auch vor der zweiten erhoben.

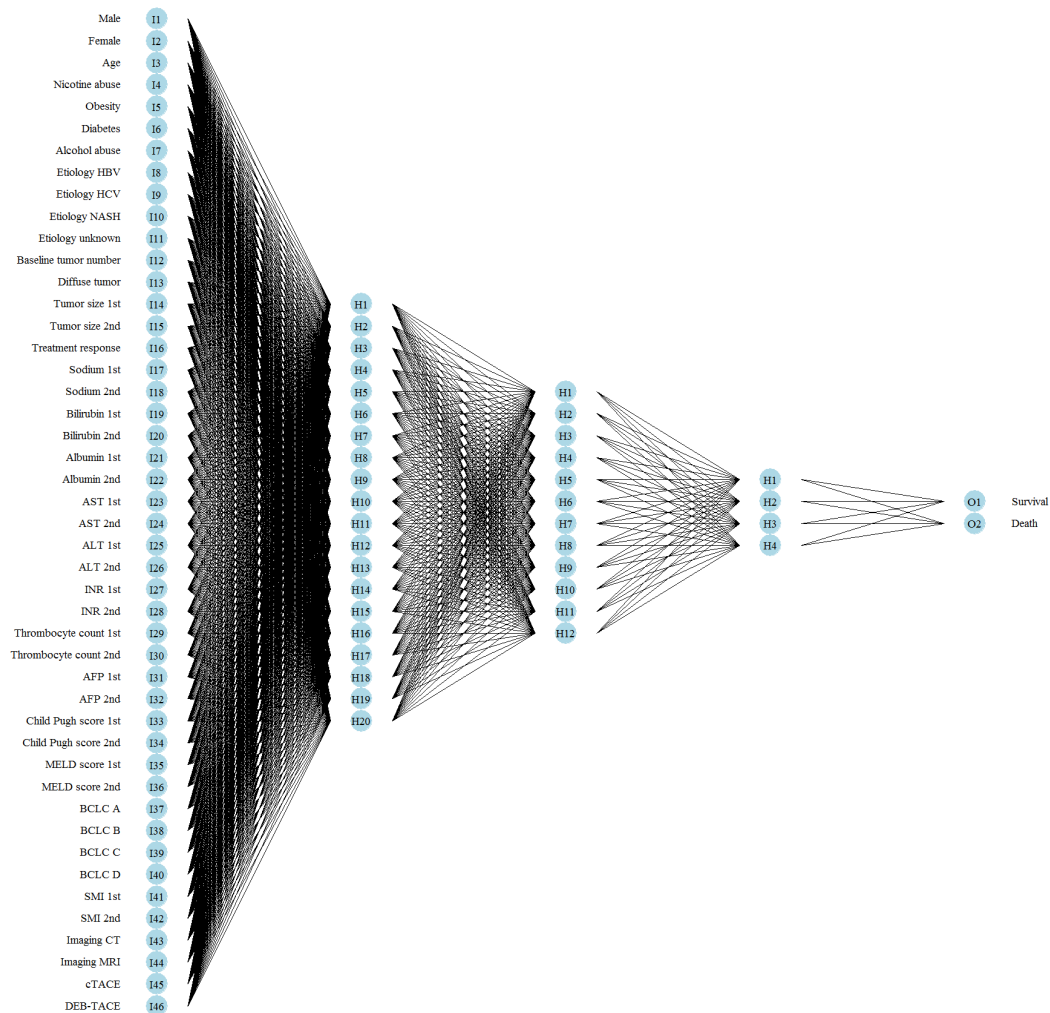
### **2.3 Aufbau des neuronalen Netzwerks**

Als Algorithmus für das maschinelle Lernen haben wir uns für ein KNN entschieden. Genauer gesagt entspricht die Struktur des KNN dem eines mehrschichtigen Perzeptrons. Die Struktur dieses mehrschichtigen Netzwerks besteht aus einer zusammenhängenden Gruppe von Knoten, welche in verschiedenen Schichten (Layern) des Netzes angeordnet sind. Das Netz besteht aus einem Input-Layer, einem oder mehreren Hidden-Layer und einem Output-Layer. Die Eingabeparameter des Netzwerkes bilden den Input-Layer. Der Output-Layer stellt alle möglichen Ergebnisse des Netzwerks dar. Die verschiedenen Hidden-Layer bestehen aus sogenannten versteckten Knoten (Hidden-Nodes). Diese verbinden die Eingabe und Ausgabeknoten und können so nichtlineare Interaktionen zwischen den Eingabevariablen ermöglichen. Diese versteckten Knoten haben kein entsprechendes klinisches Korrelat. Untereinander sind diese Knoten mittels sogenannter Links verbunden. Jedem dieser Links wird eine andere Gewichtung zugeordnet. Trainiert wurde das Netz mit einer Reihe gepaarter Ein- und Ausgabe-Daten. Das KNN lernt, indem die Ergebnisse des Netzwerks mit den erwarteten Ausgaben verglichen werden. Anhand dieses Vergleichs werden die Gewichtungen der Links angepasst. Das fertige Netz kann so die vorher festgelegten Wichtungen auf neue Eingabedaten anwenden und Vorhersagen treffen (83,99,100).

Unser KNN wurde in Python 3.7.3 mit scikit-learn (<https://scikit-learn.org/stable/>) (0.19.2) entwickelt. Das Netzwerk besteht aus 46 Input-Nodes. Beim Feintuning zeigte sich, dass die besten Ergebnisse mit einer Netzwerk Architektur bestehend aus drei vollständig verknüpften Hidden-Layern mit 20, 12 und 4 Nodes erzielt werden konnten. Für alle Hidden-Layer wurde als Aktivierungsfunktion ReLU angewandt. Für die vollständig verknüpfte Schicht wurde die Softmax-Funktion verwendet. Weiterhin wurde um eine Überanpassung (Overfitting) im Netzwerk zu vermeiden eine strikte L2-Regularisierung angewandt. Das KNN wurde genutzt, um das Gesamtüberleben nach dem ersten Jahr zu bestimmen. Als Startzeitpunkt wurde das Datum des Tages vor der zweiten TACE festgelegt. Somit repräsentieren die beiden Nodes des Output-Layers Überleben (=1) und Tod (=2).

Alle 46 erfassten Parameter bilden jeweils einen Input-Node.

Alle kontinuierlichen Eingangsparameter wurden durch Subtraktion des Mittelwerts und Division durch die Standardabweichung standardisiert.



**Abbildung 12:** Struktur des KNN bestehend aus einem Input-Layer, 3 Hidden-Layer und dem Output-Layer (86).

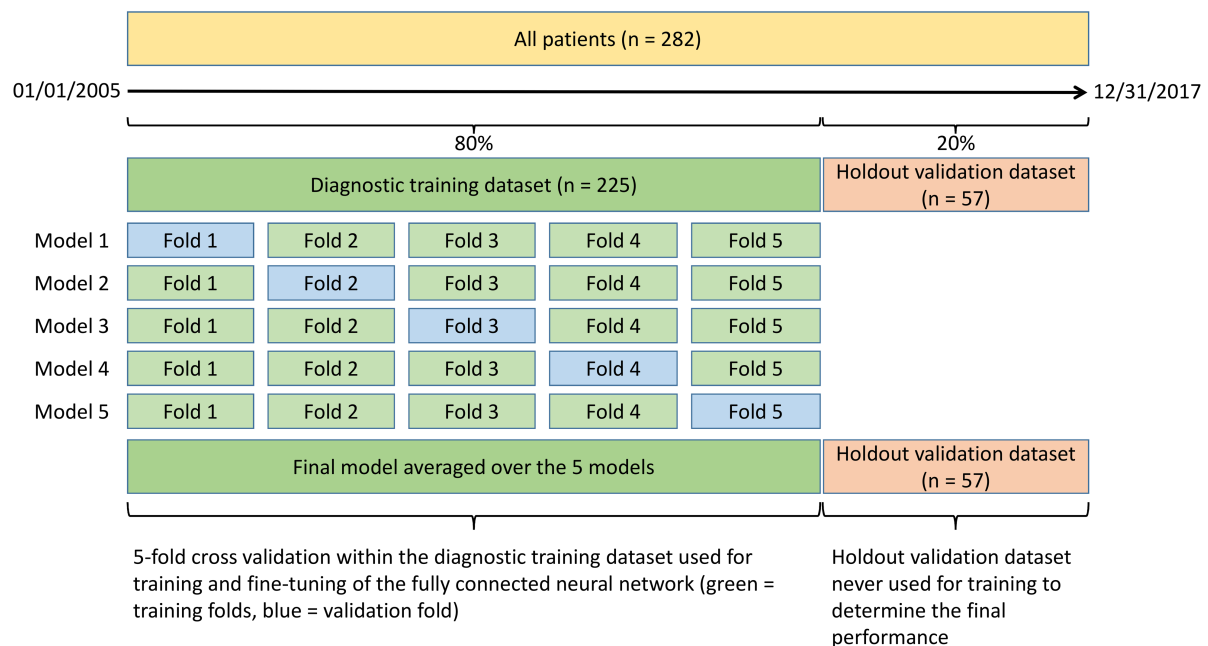
## 2.4 Training und Validierung des Neuronalen Netzwerks

Zuerst wurden die 282 Patienten in zwei Gruppen zu 225 (80%) und 57 Patienten (20%) aufgeteilt. Die Daten der ersten 225 Patienten wurden als Trainingsdatensatz und zum Feintuning des KNN genutzt.

Die übrigen 57 Patienten, welche zuletzt behandelt wurden, bildeten den unabhängigen Holdout-Validierungsdatensatz.

Um die Trainingskapazität zu maximieren, wurde bei den diagnostischen Trainingsdaten eine 5-Fache Kreuzvalidierung angewandt. Dies bedeutet, dass in Summe 5 verschiedene Modelle trainiert wurden. Hierbei wurde je eine Untergruppe von 180 Patienten zum Training und 45 Patienten zum Testen verwendet. Das vollständige Modell ergibt sich aus dem Durchschnitt der fünf Modelle.

Für die abschließende Evaluation des Netzwerks wurde der zu Beginn separierte und damit unabhängige Validierungsdatensatz der restlichen 57 Patienten genutzt. Anhand dieser Daten konnte zudem ein direkter Vergleich mit vorhandenen Scoring Systemen durchgeführt werden.



**Abbildung 13:** Modell der 5-fachen Kreuzvalidierung (86).



Beide Gruppen wurden bezüglich der Ausgangsmerkmale stratifiziert.

**Tabelle 9:** Detaillierte Aufstellung der Ausgangsparameter einschließlich entsprechender P-Werte. (86).

	Diagnostic training dataset (n = 225)	Holdout validation dataset (n = 57)	P-value
Age, years	66 (16-90)	68 (49-85)	.020
Gender			.809
Male	188 (83.6%)	49 (86.0%)	
Female	37 (16.4%)	8 (14.0%)	
BCLC stage			.139
A	37 (16.4%)	5 (8.8%)	
B	173 (76.9%)	51 (89.5%)	
C	11 (4.9%)	0 (0.0%)	
D	4 (1.8%)	1 (1.7%)	
Child-Pugh stage			.397
A	168 (74.7%)	43 (75.4%)	
B	53 (23.5%)	13 (22.8%)	
C	4 (1.8%)	1 (1.8%)	
Tumour size, mm	49 (10-270)	60 (10-180)	.293
Tumour number			.205
1	90 (40.0%)	23 (40.4%)	
2	34 (15.1%)	10 (17.5%)	
3	22 (9.8%)	11 (19.3%)	
4	24 (10.7%)	2 (3.5%)	
5	9 (4.0%)	2 (3.5%)	
6	9 (4.0%)	2 (3.5%)	
7	1 (0.4%)	2 (3.5%)	
8	2 (0.9%)	1 (1.8%)	
9	3 (1.3%)	0 (0.0%)	
≥10	31 (13.8%)	4 (7.0%)	
Diffuse tumour			.316
Yes	29 (12.9%)	4 (7.0%)	
No	196 (87.1%)	53 (93.0%)	
Alcohol <sup>a</sup>			.954
Yes	110 (48.9%)	27 (47.4%)	
No	115 (51.1%)	30 (52.6%)	
HBV <sup>a</sup>			.111
Yes	42 (18.7%)	5 (8.8%)	
No	183 (81.3%)	52 (91.2%)	
HCV <sup>a</sup>			.270
Yes	62 (27.6%)	11 (19.3%)	
No	163 (72.4%)	46 (80.7%)	
NASH <sup>a</sup>			.507
Yes	12 (5.3%)	5 (8.8%)	
No	213 (94.7%)	52 (91.2%)	
Unknown aetiology <sup>a</sup>			1.000
Yes	30 (13.3%)	9 (15.8%)	
No	195 (86.7%)	48 (84.2%)	
Nicotine abuse			.684
Yes	51 (22.7%)	15 (26.3%)	

	Diagnostic training dataset (n = 225)	Holdout validation dataset (n = 57)	P-value
No	174 (77.3%)	42 (73.7%)	
Obesity			.339
Yes	89 (39.6%)	18 (31.6%)	
No	136 (60.4%)	39 (68.4%)	
Diabetes			.307
Yes	99 (44.0%)	30 (52.6%)	
No	126 (56.0%)	27 (47.4%)	
MELD score	10 (5-24)	10 (5-17)	.790
SMI	0.215 (0.125-0.324)	0.212 (0.139-0.341)	.843
AFP, ng/ml	16 (1-164852)	14 (1-26881)	.342
Bilirubin, mg/dl	1.05 (0.3-5.9)	1.04 (0.3-4.5)	.915
Albumin, g/L	35 (20-54)	34 (22-42)	.019
INR	1.1 (0.9-3.2)	1.2 (0.9-1.9)	.364
ALT, U/L	43 (4-477)	36 (12-245)	.283
AST, U/L	56 (16-358)	58 (23-187)	.997
Thrombocyte count, per nl	135 (34-720)	130 (34-458)	.345
Sodium, mmol/L	138 (124-147)	138 (131-142)	.373
TACE-type			<.001
cTACE	144 (64.0%)	8 (14.0%)	
DEB-TACE	81 (36.0%)	49 (86.0%)	
Previous therapy before first TACE			.772
Surgical resection	44 (19.6%)	12 (21.1%)	
Local ablation therapy	4 (1.8%)	0 (0.0%)	
Sorafenib therapy	3 (1.3%)	1 (1.7%)	
none	174 (77.3%)	44 (77.2%)	
Subsequent therapy after second TACE <sup>b</sup>			.493
TACE only	175 (77.8%)	50 (87.7%)	
Surgical resection	9 (4.0%)	0 (0%)	
Local ablation therapy	10 (4.4%)	1 (1.8%)	
SIRT	5 (2.2%)	1 (1.8%)	
Systemic therapy	24 (10.7%) <sup>c</sup>	5 (8.8%) <sup>d</sup>	
SIRT + Sorafenib	2 (0.9%)	0 (0.0%)	
Type of imaging before the second TACE			<.001
CT	189 (84.0%)	17 (29.8%)	
MRI	36 (16.0%)	40 (70.2%)	

Sofern nicht anders vermerkt, sind die Daten in n (Prozent) oder im Median (Spannweite) angegeben.

<sup>a</sup> Die Summe aller Ätiologien kann 100% übersteigen, da Patienten mehr als eine Ätiologie aufweisen konnten.

<sup>b</sup> Es wurden nur Therapien berücksichtigt, die in den 12 Monaten der Beobachtungszeit durchgeführt wurden.

<sup>c</sup> Alle 24 Patienten wurden mit Sorafenib behandelt. Bei 8 der Patienten begann die Sorafenib-Therapie zwischen der ersten und der zweiten TACE.

<sup>d</sup> Zwei Patienten wurden mit Nivolumab behandelt. Alle weiteren mit Sorafenib.

Beide Gruppen waren bezüglich der Parameter vergleichbar. Nur Alter ( $p = 0,020$ ), Albumin ( $p = 0,019$ ), Art der TACE ( $p < 0,001$ ) und Art der Bildgebung vor der zweiten TACE ( $p < 0,001$ ) waren signifikant unterschiedlich.

## 2.5 Statistische Analyse

Alle kontinuierlichen Variablen wurden anhand des Medians und der Spannweite bewertet und gegebenenfalls mit einem zweiseitigen ungepaarten Student's t-Test oder Wilcoxon-Test verglichen.

Kategoriale Merkmale wurden anhand von Prozentangaben beschrieben und mittels des Chi-Quadrat-Tests, beziehungsweise des exakten Fisher-Tests verglichen.

Die Genauigkeit der Einschätzung des Gesamtüberlebens des KNNs wurde anhand von AUROCs bestimmt.

Dieselbe Methodik wurde genutzt um das KNN mit dem ART- und ABCR-Score sowie dem SNACOR zu vergleichen.

Die kumulativen beziehungsweise dynamischen ROCs wurden mit Python 3.7.3 und matplotlib 3.1.0 erstellt (<http://matplotlib.org/>).

Die Vorhersagegenauigkeit des KNN wurde zusätzlich durch Sensitivität, Spezifität und positivem- sowie negativem prädiktivem Wert beschrieben.

Als Signifikanzniveau wurde ein p Wert  $< 0,05$  festgelegt.

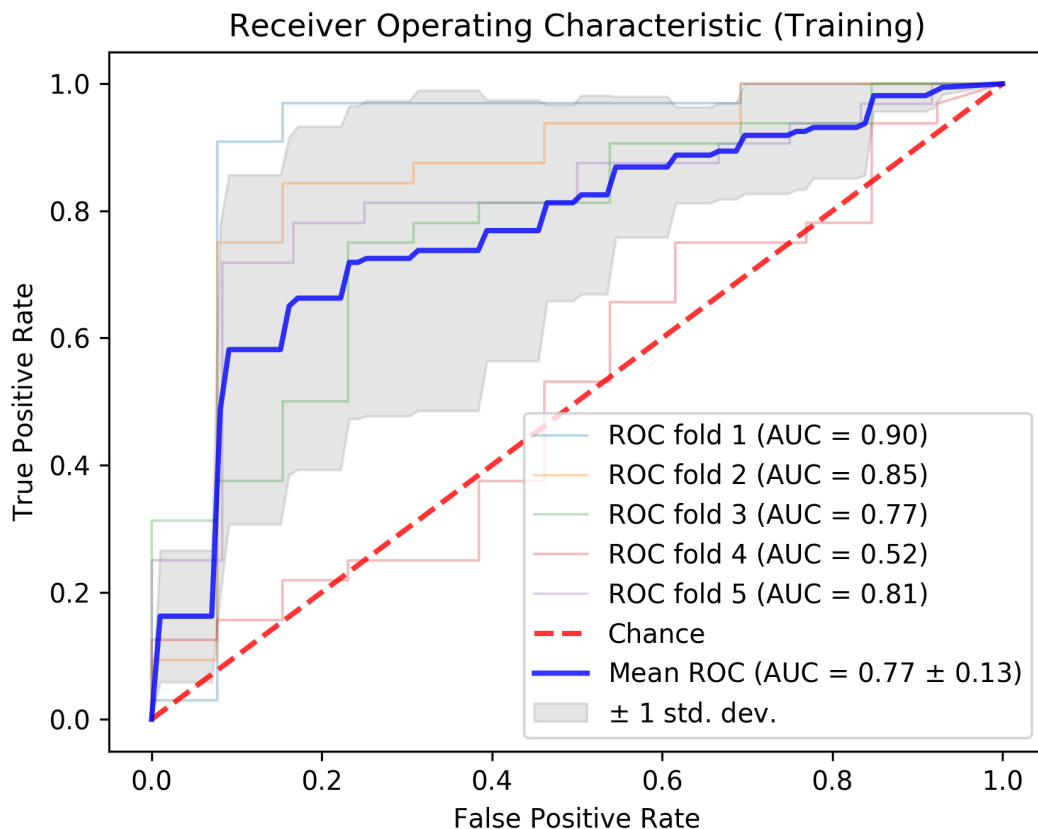
Da diese Analyse rein explorativ gedacht war, sollten die P-Werte deskriptiv interpretiert werden. R 3.5.2 und R 3.6.0 (A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, <https://www.R-project.org>; Zugriff 2018/2019) wurden für statistische Analysen verwendet.

## 3 Ergebnis

In der diagnostischen Trainingskohorte lag die Ein-Jahresüberlebensrate bei 71,5 %. Bei der Validierungsgruppe waren es 63,1 % (P= 0,283).

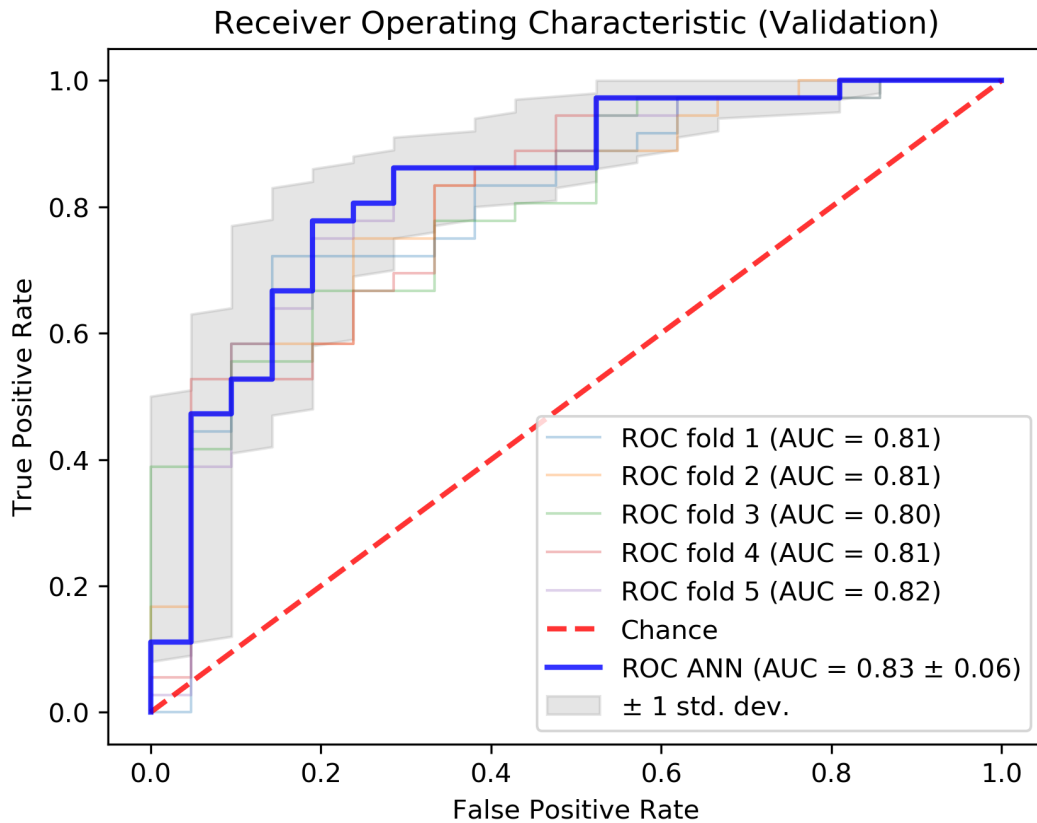
### 3.1 Genauigkeit des KNN

Für die Abschätzung des 1-Jahresüberlebens innerhalb der Trainingskohorte erreichte das KNN eine AUC von im Mittel  $0,77 \pm 0,13$ .



**Abbildung 14:** ROC der Trainingskohorte (86).

Diese Ergebnisse wurden durch die Validierungsgruppe verifiziert. Hier wurde im Mittel eine AUC von  $0,83 \pm 0,06$  erreicht. Der positiv prädiktive Wert lag bei 87,5% und der negativ prädiktive Wert bei 68%. Die Sensitivität des KNNs war 77,8% und die Spezifität 81%.

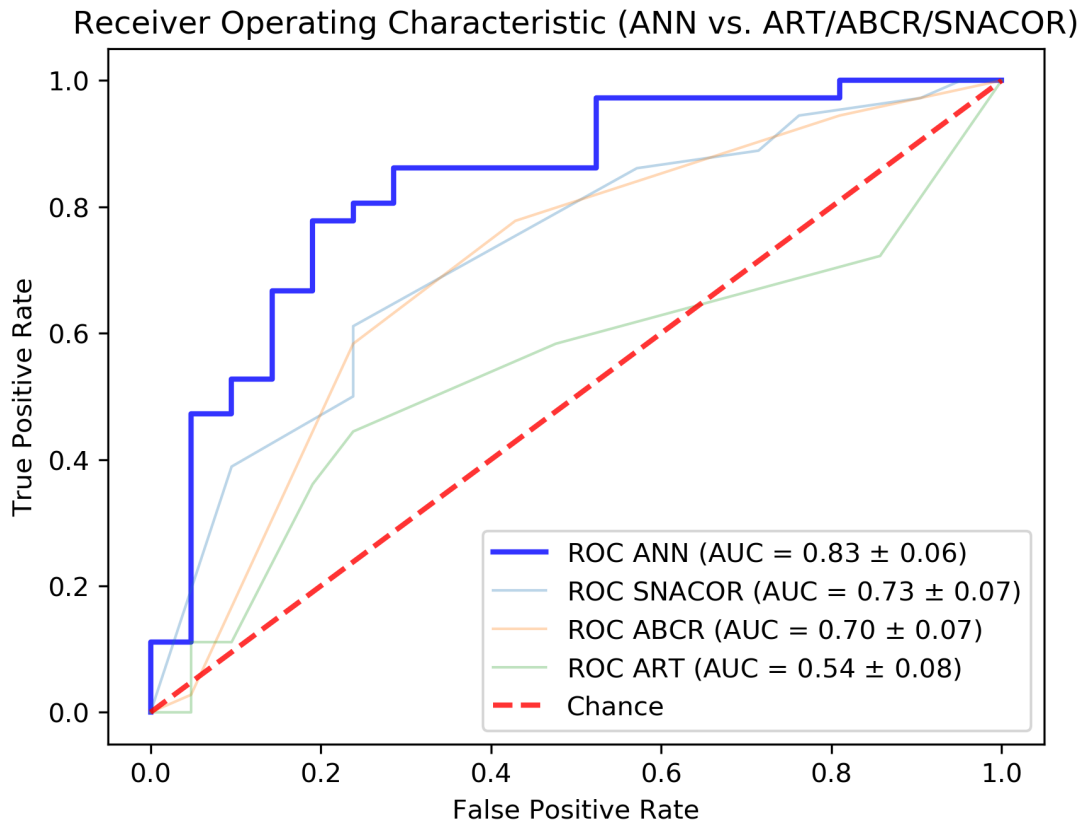


**Abbildung 15:** ROC der Validierungskohorte (86).

### 3.2 Genauigkeit des KNN im Vergleich zu bekannten Scoring-Systemen

Im letzten Schritt wurde das KNN mit den bereits existierenden Bewertungssystemen ART, ABCR und SNACOR verglichen (71,72,74). Die AUC des ART-Scores betrug 0,54 +/- 0,08, die des ABCR-Score: 0,70 +/- 0,07 und der SNACOR erreichte eine AUC von 0,73 +/- 0,07. Somit waren alle AUCs niedriger als bei unserem KNN.

Im Falle des ART-Scores war der Unterschied zum ANN signifikant ( $p < 0,001$ ), im Vergleich zu ABCR ( $p = 0,143$ ) und SNACOR ( $p = 0,201$ ) war der Unterschied nicht signifikant.



**Abbildung 16:** Vergleich der ROCs zwischen dem KNN, ART, ABRC und SNACOR (86).

Für das KNN wurde mit einem Wert von 0,83 eine „gute Vorhersage“, für ABCR 0,7 und SNACOR mit 0,73 eine „mittelmäßige Vorhersage“ und für den ART-Score mit 0,54 nur eine „sehr schlechte Vorhersage“ erreicht.

Das vollständige KNN ist online, öffentlich im Mendeley-Repository verfügbar (101). Um eine einfache Implementierung des Modells zu ermöglichen, umfasst das Repository das Python-Skript, Beispieldaten und ein ausführliches Handbuch.

## 4 Diskussion

Bei Patienten mit einem primären Leberzellkarzinom ist eine möglichst präzise Schätzung der Überlebenszeit für eine bestmögliche Therapieplanung entscheidend. Da dies jedoch auch für erfahrene Ärzte teilweise schwierig abzuschätzen ist, wurden sogenannte Scoring Systeme entwickelt. Nachdem diese in mehreren Studien unbefriedigende Ergebnisse gezeigt haben, ist in der vorliegenden Studie erstmals ein künstliches neuronales Netzwerk zum Einsatz gekommen, um das Gesamtüberleben bei HCC-Patienten unter TACE-Therapie einzuschätzen.

Dieses erzielte bei der Vorhersage des 1-Jahresüberlebens vor der zweiten TACE vielversprechende Ergebnisse. Mit einer AUC von 0,83 war es den bisher bekannten Scoring-Systemen überlegen.

Das HCC gehört weltweit zu den häufigsten zum Tode führenden Krebserkrankungen (102). Somit ist ein zuverlässiges Tool um die Überlebenszeit abzuschätzen und so auch die Entscheidung über die bestmögliche Therapie zu erleichtern ein wichtiger Schritt. Da beim HCC die möglichen Faktoren, welche die Überlebenszeit beeinflussen vielfältig sind, ist ein neuronales Netz eine vielversprechende Grundlage. Ein neuronales Netz kann über die Berücksichtigung der einzelnen Faktoren hinaus auch diese durch maschinelle Lernvorgänge mit unterschiedlichen Wichtungen versehen, um eine möglichst genaue Vorhersage zu treffen.

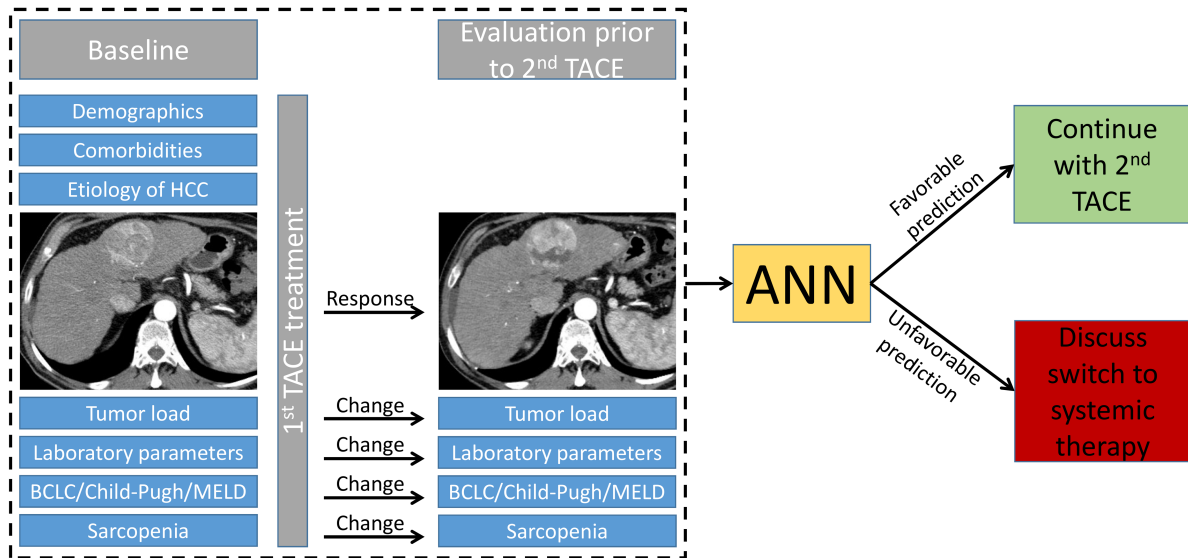
Konventionelle Scoring-Systeme wie der ART- und ABCR-Score sowie der SNACOR wurden entwickelt um diese Therapieentscheidung vor der zweiten TACE zu objektivieren (71,72,74). So soll entschieden werden, ob der Patient von weiteren Interventionen profitiert. Die genannten Systeme erreichten eine AUC/- bzw. einen Harrell's C Index von 0,6 und 0,75 für die Abschätzung des 1-Jahresüberlebens (72,74). Dies konnte allerdings von verschiedenen Forschungsgruppen in Studien zur externen Validierung nicht reproduziert werden (79–81,103). In einer Studie unserer Klinik wurde zum Beispiel ein Harrell's C Index von nur 0,54 und 0,59 erreicht (79,80). Diese Differenz ist wahrscheinlich zurückzuführen auf das sogenannte Overfitting. Zu diesem Phänomen kommt es, wenn ein datenabhängiges Diagnosemodell zu sehr an die zugrunde liegenden Daten angepasst wird. So kann aufgrund zufälliger Schwankungen der Patientencharakteristika in unterschiedlichen klinischen und demografischen Umgebungen die Vorhersagegenauigkeit nicht reproduziert werden

(104). Das Phänomen der Nicht-Reproduzierbarkeit ist vor allem in den Biowissenschaften ein bekanntes Problem. Hierbei führen verschiedene Faktoren wie Systemeigenschaften, Verzerrungen der Studienergebnisse, Probleme beim Studiendesign, der Durchführung oder der Interpretation, zu unterschiedlichen Ergebnissen zwischen der ursprünglichen Studie und dem Validierungsversuch (105).

Bisher wurde der Ansatz des machine learnings im Zusammenhang mit HCC Patienten unter TACE Therapie nur selten versucht (106,107). Diese Studien unterscheiden sich jedoch zu unserer in der genutzten Methodik, sowie in den gewünschten Ausgabeparametern. Peng et. al nutzten ein neuronales Faltungsnetzwerk basierend auf einer Mustererkennung in CT-Bildern (106). Abajian et al nutzten hingegen 5 Parameter, davon zwei basierend auf MRT-Bildern und drei klinische Parameter. Hier wurden eine logistische Regression und Random-Forest-Models angewandt. (107). Das Ziel dieser beiden Studien war es, das Ansprechen des Tumors vor der ersten TACE einzuschätzen (106,107). Im Gegensatz dazu benutzten wir ein voll verbundenes KNN basierend auf einer sehr viel größeren Zahl an klinischen Parametern (n=46) um das Gesamtüberleben vor der zweiten TACE vorherzusagen. Ähnliche Ansätze wurden bereits im Zusammenhang mit Tumorresektionen vorgestellt (108–110). Im Feld der interventionellen Onkologie haben Wu et. al versucht das tumorfreie Überleben bei HCC-Patienten nach Radiofrequenz Ablation vorherzusagen (111). Zum Zeitpunkt unserer Studie gab es diesen Ansatz im Zusammenhang mit einer TACE-Therapie noch nicht. Mit diesem neuen Ansatz erreichten wir mit einer AUC von 0,83 eine vielversprechende Vorhersageleistung (Predictive Performance).

Ist das KNN einmal trainiert kann man es einfach in den Klinikalltag integrieren, um über das weitere therapeutische Vorgehen zu entscheiden.





**Abbildung 17:** Studienablauf und Endpunkte des KNN (86)

Im direkten Vergleich unseres KNN mit dem ART, ABCR und SNACOR Score zeigte das KNN die höchste AUC, welche insgesamt einer guten Vorhersagequalität entspricht. Die AUC für den ABCR-Score und SNACOR entsprechen einer mittleren Vorhersagequalität. Den niedrigsten Wert erreichte der ART-Score mit einer „sehr schlechten“ Vorhersagequalität. Die Genauigkeit der Vorhersage unseres KNNs war jedoch nur im Vergleich zum ART-Score signifikant besser. Im Vergleich zu ABCR-Score und SNACOR war der Unterschied nicht signifikant. Da dieser Vergleich lediglich auf der Holdout-Validierungsgruppe, bestehend aus nur 57 Patienten, beruht, könnte das nicht-signifikante Ergebnis einer möglichen underpower geschuldet sein.

Einer der größten Vorteile eines solchen KNNs ist die große Anzahl an Variablen, die berücksichtigt werden können. In unserem Fall sind es insgesamt 46. Die Auswahl dieser Variablen erfolgte anhand von etablierten, evidenzbasierten, prognostischen Parametern, welche im klinischen Alltag bei Patienten mit einer Zirrhose beziehungsweise einem HCC genutzt werden. Außerdem sind KNNs bei Zunahme der Komplexität, der Anzahl der Muster und der Anzahl der Eingaben des Datensatzes leicht skalierbar. Dies kann ein Vorteil gegenüber dem klassischen machine learning wie zum Beispiel dem random forest classifier sein (112).

Das KNN hat allerdings auch einige Schwächen. Im Gegensatz zu herkömmlichen statistischen Ansätzen kann das KNN weder mit fehlenden Werten umgehen, noch ist es einfach, Parameter zu identifizieren, welche die Vorhersage besonders gut beeinflussen. Um multiple Imputationen zu vermeiden, muss der Datensatz daher so vollständig wie möglich sein. Leider sind viele medizinische Patientendaten unvollständig, oder nur schwer aus verschiedenen Datenquellen (z. B. separates Radiologie-Informationssystem, Krankenhaus-Informationssystem, Labor-Informationssystem usw.) zu akquirieren. Ein weiteres Problem sind Daten, welche noch nicht digitalisiert wurden und somit lediglich in Papierform vorhanden sind. In der Zukunft könnte dies allerdings durch neue Programme wie zum Beispiel das Einführen von strukturierten Befunden deutlich vereinfacht werden und somit die Qualität, Vollständigkeit und die Verfügbarkeit der Daten verbessert werden. Durch diese Neuerungen kann ein neuronales Netz leichter aufgebaut und trainiert werden.

Weiterhin weist unsere Analyse einige Limitationen auf. Die wohl wesentlichste Einschränkung ist die fehlende externe Validierung unserer Studie. Auch wenn die Holdout-Validierungsgruppe nicht zum Training des Netzwerks genutzt wurde, sondern nur zur Validierung, ist eine externe Validierung unumgänglich. Ohne diese kann die Reproduzierbarkeit nicht sichergestellt werden. Wir haben das KNN zum Download im Mendeley Repositorium bereitgestellt, um eine Verifizierung unserer Ergebnisse durch unabhängige Studiengruppen zu erleichtern (101).

Darüber hinaus war die Stichprobe aufgrund der retrospektiven Datenerhebung mit nur 282 Patienten nur mäßig groß. Höchstwahrscheinlich war der Datensatz zu klein, um den vollen Umfang des KNNs auszunutzen. Aus diesem Grund ist es möglich, dass unser KNN den konventionellen Ansätzen nicht überlegen ist. Es ist jedoch eher unwahrscheinlich, dass unser Ansatz schlechter ist als die klassischen Scoring-Systeme. Idealerweise sollte in der Zukunft multizentrisch mit einer ausreichend großen Patientenkohorte zunächst das KNN trainiert und anschließend validiert werden. Ein solcher Ansatz würde sowohl das Problem des Overfittings lösen als auch die Robustheit des KNNs verbessern. Weiterhin haben wir sämtliche Variablen aus allen bekannten Bewertungssystemen einbezogen, einschließlich auch einiger eigens berechneter Werte wie zum Beispiel das Child-Pugh sowie das BCLC-Stadium. Diese erfordern immer noch zusätzliche Benutzereingaben. Insgesamt handelt es sich bei den hier verwendeten Parametern jedoch um klinische Parameter zur Charakterisierung von Patienten mit Lebererkrankungen, welche in den meisten

Leberzentren verfügbar sein sollten.

Auch wenn wir eine große Bandbreite an potenziell, klinisch relevanten Parametern eingeschlossen haben, ist es möglich, dass wir weitere relevante Variablen nicht erfasst haben. Diese hätten möglicherweise das Netzwerk weiter verbessern können. Darüber hinaus haben wir manche Parameter nicht einbezogen, da sie nicht für alle Patienten verfügbar waren. Es fehlte zum Beispiel bei einigen Patienten das Tumorgrading sowie der Status der Infiltration kleiner Gefäße, da sie nicht-invasiv (mittels Bildgebung) diagnostiziert wurden. Darüber hinaus besteht die Möglichkeit die Inklusion weiterer komplexer Parameter die Vorhersage weiter verbessern könnte. Dazu zählen zum Beispiel radiologische Daten einschließlich Texturanalyse (113–115).

Ein weiterer möglicher Grund für eine Verzerrung könnte die große Zeitspanne der Datenerhebung sein, welche in Summe 13 Jahre umfasst. In dieser Zeit kam es zu vielen technischen Verbesserungen in der TACE-Therapie. Auch wenn bereits 2006 die DEB-TACE an der Universitätsmedizin Mainz eingeführt wurde und so beide Verfahren über die gesamte Beobachtungszeit angewendet wurden, kam es zu einer ungleichen Verteilung von cTACE und DEB-TACE zwischen dem Training und der Holdout-Validierungsgruppe. Diese ungleiche Verteilung kommt durch einen verstärkten Einsatz der DEB-TACE in den letzten Jahren zustande. In den beiden größten, multizentrischen randomisierten klinischen Studien waren jedoch beide Therapien gleichermaßen erfolgreich (88,116). Aus diesem Grund werden beide Therapiemöglichkeiten äquivalent von den EASL-Leitlinien empfohlen. Die Entscheidung über das verwendete Chemotherapeutikum bleibt dem Untersucher überlassen (9). Außerdem kamen neue systemische Medikamente auf den Markt, welche den Wechsel von TACE zur systemischen Therapie beeinflussen (117).

Um die Tumorlast und das Therapieansprechen des Tumors zu bestimmen, wurden CT und MRT-Untersuchungen eingesetzt. Da in den letzten Jahren vermehrt die MRT zum Einsatz kam, ist in der Validierungsgruppe die Anzahl der MRT-Bildgebungen größer als in der Trainingskohorte. Allerdings sind beide Verfahren zur Bildgebung bei HCC-Patienten und zur Bestimmung des Therapieansprechens gleichermaßen anerkannt (9,75). Darüber hinaus haben wir primär nur kaukasische Patienten eingeschlossen. Aufgrund grundlegender Unterschiede in den Patientencharakteristika, sind die Ergebnisse möglicherweise nicht auf asiatische

Patienten übertragbar, da sie sich z.B. hinsichtlich der Ätiologie der Lebererkrankung unterscheiden. Zuletzt ist noch zu nennen, dass wir uns für ein neuronales Netz mit drei hidden-layers entschieden haben. Da für ähnliche Zwecke kein allgemein anerkanntes Design solcher Netzwerke verfügbar ist, ist es durchaus möglich, dass das gewählte Design nicht das ideale ist. Ein perfekt konzipiertes Netz könnte somit eine noch bessere Vorhersage ermöglichen.

## **5 Zusammenfassung**

Bei der Vorhersage des Überlebens von HCC-Patienten nach TACE könnten Neuronale Netzwerke den bereits existierenden Bewertungssystemen, welche einen konventionellen statistischen Ansatz verfolgen, überlegen sein. Ist ein solches Netz einmal etabliert, ist die Implementierung in den klinischen Alltag problemlos möglich und kann zu einer optimalen Patienten Versorgung beitragen. Besonders weniger erfahrene Untersucher könnten von einer Entscheidungshilfe, welche auf machine learning basiert, profitieren.

Nichtsdestotrotz ist der klinische Alltag meist komplexer als es in einem Netzwerk abgebildet werden kann. Deshalb kann die langjährige Erfahrung eines Arztes auch nicht durch machine learning basierte Neuronale Netze ersetzt werden. Das KNN kann lediglich eines von mehreren Komponenten im Prozess der Entscheidungsfindung sein. Das Hinzufügen weiterer Parameter kann die Vorhersagegenauigkeit des Netzes womöglich weiter steigern. Zu diesen Parametern könnten zum Beispiel bereits als prognostisch wertvoll erkannte Parameter wie das Postembolisations-Syndrom nach der ersten TACE (118) aber auch neue Parameter wie die sogenannte Texturanalyse gehören (113–115). Eine Kombination unseres KNN mit einem konvolutionalen neuronalen Netz, welches das Prinzip der Mustererkennung nutzt, könnte womöglich die Vorhersage weiter verbessern. Um die Übertragbarkeit zu verbessern und das Problem des Overfittings zu reduzieren, sollte das Netzwerk aus einer größeren Datenbank mit klinischen Daten aus mehreren Zentren aufgebaut werden.

## 6 Literaturverzeichnis

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. November 2018;68(6):394–424.
2. Petrick JL, Florio AA, Znaor A, Ruggieri D, Laversanne M, Alvarez CS, u. a. International trends in hepatocellular carcinoma incidence, 1978–2012. *Int J Cancer*. 15. Juli 2020;147(2):317–30.
3. Global Burden of Disease Liver Cancer Collaboration, Akinyemiju T, Abera S, Ahmed M, Alam N, Alemayohu MA, u. a. The Burden of Primary Liver Cancer and Underlying Etiologies From 1990 to 2015 at the Global, Regional, and National Level: Results From the Global Burden of Disease Study 2015. *JAMA Oncol*. 1. Dezember 2017;3(12):1683–91.
4. IARC Monographs of Carcinogenic Hazards to Humans and Handbooks of Cancer Prevention [Internet]. 2018 [zitiert 18. September 2021]. Verfügbar unter: <http://monographs.iarc.fr/ENG/Publications/OrganSitePoster.pdf>
5. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body Fatness and Cancer — Viewpoint of the IARC Working Group. *N Engl J Med*. 25. August 2016;375(8):794–8.
6. El-Serag HB. Epidemiology of Viral Hepatitis and Hepatocellular Carcinoma. *Gastroenterology*. Mai 2012;142(6):1264-1273.e1.
7. White DL, Thrift AP, Kanwal F, Davila J, El-Serag HB. Incidence of Hepatocellular Carcinoma in All 50 United States, From 2000 Through 2012. *Gastroenterology*. März 2017;152(4):812-820.e5.
8. El-Serag HB, Rudolph KL. Hepatocellular Carcinoma: Epidemiology and Molecular Carcinogenesis. *Gastroenterology*. Juni 2007;132(7):2557–76.
9. Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul JL, u. a. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol*. Juli 2018;69(1):182–236.
10. Baenkler HW, Goldschmidt H, Hahn JM, Hinterseer M, Knez A, Lafrenz M, u. a.

Kurzlehrbuch Innere Medizin. 3. Thieme; 2015.

11. Simonetti RG, Camma C, Fiorello F, Politi F, D'Amico G, Pagliaro L. Hepatocellular carcinoma: A worldwide problem and the major risk factors. *Dig Dis Sci.* Juli 1991;36(7):962–72.
12. Sangiovanni A, Prati GM, Fasani P, Ronchi G, Romeo R, Manini M, u. a. The natural history of compensated cirrhosis due to hepatitis C virus: A 17-year cohort study of 214 patients. *Hepatology.* Juni 2006;43(6):1303–10.
13. Szabó E, Lotz G, Páska C, Kiss A, Schaff Z. Viral hepatitis: New data on hepatitis C infection. *Pathol Oncol Res.* Dezember 2003;9(4):215–21.
14. Halliday J, Klenerman P, Barnes E. Vaccination for hepatitis C virus: closing in on an evasive target. *Expert Rev Vaccines.* Mai 2011;10(5):659–72.
15. Raimondi S, Bruno S, Mondelli MU, Maisonneuve P. Hepatitis C virus genotype 1b as a risk factor for hepatocellular carcinoma development: A meta-analysis. *J Hepatol.* 2009;50:1142–54.
16. Kanwal F, Kramer JR, Ilyas J, Duan Z, El HB. HCV Genotype 3 is Associated with an Increased Risk of Cirrhosis and Hepatocellular Cancer in a National Sample of U.S. Veterans with HCV. 2015;60(1):98–105.
17. Karayiannis P. Hepatitis B virus: virology, molecular biology, life cycle and intrahepatic spread. *Hepatol Int.* November 2017;11(6):500–8.
18. Chen CJ, Yang HI, Su J, Jen CL, You SL, Lu SN, u. a. Risk of Hepatocellular Carcinoma Across a Biological Gradient of Serum Hepatitis B Virus DNA Level. 295(1):65–73.
19. Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: Incidence and risk factors. *Gastroenterology.* November 2004;127(5):S35–50.
20. Kanwal F, Kramer JR, Asch SM, Cao Y, Li L, El-Serag HB. Long-Term Risk of Hepatocellular Carcinoma in HCV Patients Treated With Direct Acting Antiviral Agents. *Hepatology.* Januar 2020;71(1):44–55.
21. Raffetti E, Fattovich G, Donato F. Incidence of hepatocellular carcinoma in

untreated subjects with chronic hepatitis B: a systematic review and meta-analysis. *Liver Int.* September 2016;36(9):1239–51.

22. Papatheodoridis GV, Chan HLY, Hansen BE, Janssen HLA, Lampertico P. Risk of hepatocellular carcinoma in chronic hepatitis B: Assessment and modification with current antiviral therapy. *J Hepatol.* April 2015;62(4):956–67.

23. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF): Diagnostik und Therapie des Hepatozellulären Karzinoms und biliärer Karziome, Langversion 2.0, 2021, AWMF Registernummer: 032/053OL, <https://www.leitlinienprogramm-onkologie.de/leitlinien/hcc-und-billiäre-karzinome> (Zugriff am 11.01.2022).

24. Anstee QM, Targher G, Day CP. Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis. *Nat Rev Gastroenterol Hepatol.* Juni 2013;10(6):330–44.

25. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes: *Hepatology.* Juli 2016;64(1):73–84.

26. Perlmutter DH. Pathogenesis of Chronic Liver Injury and Hepatocellular Carcinoma in Alpha-1-Antitrypsin Deficiency. *Pediatr Res.* August 2006;60(2):233–8.

27. Zhou WC, Zhang QB, Qiao L. Pathogenesis of liver cirrhosis. *World J Gastroenterol.* 2014;20(23):7312.

28. Deugnier YM, Guyader D, Crantock L, Lopez JM, Turlin B, Yaouanq J, u. a. Primary liver cancer in genetic hemochromatosis: A clinical, pathological, and pathogenetic study of 54 cases. *Gastroenterology.* Januar 1993;104(1):228–34.

29. Bruix J, Sherman M. Management of hepatocellular carcinoma: An update. *Hepatology.* März 2011;53(3):1020–2.

30. Brouwer WP, van der Meer AJP, Boonstra A, Plompen EPC, Pas SD, de Kneegt RJ, u. a. Prediction of long-term clinical outcome in a diverse chronic hepatitis B population: Role of the PAGE-B score. *J Viral Hepat.* November 2017;24(11):1023–31.

31. Papatheodoridis G, Dalekos G, Sypsa V, Yurdaydin C, Buti M, Goulis J, u. a. PAGE-B predicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitis B on 5-year antiviral therapy. *J Hepatol.* April 2016;64(4):800–6.
32. Bruix J, Sherman M. Management of hepatocellular carcinoma. *Hepatology.* November 2005;42(5):1208–36.
33. Tzartzeva K, Obi J, Rich NE, Parikh ND, Marrero JA, Yopp A, u. a. Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis. *Gastroenterology.* Mai 2018;154(6):1706-1718.e1.
34. Giorgio A, Montesarchio L, Gatti P, Amendola F, Matteucci P, Santoro B, u. a. Contrast-Enhanced Ultrasound: a Simple and Effective Tool in Defining a Rapid Diagnostic Work-up for Small Nodules Detected in Cirrhotic Patients during Surveillance. *J Gastrointestin Liver Dis.* 2016;25(2):7.
35. Ayuso C, Rimola J, Vilana R, Burrel M, Darnell A, García-Criado Á, u. a. Diagnosis and staging of hepatocellular carcinoma (HCC): current guidelines. *Eur J Radiol.* April 2018;101:72–81.
36. Lee YJ, Lee JM, Lee JS, Lee HY, Park BH, Kim YH, u. a. Hepatocellular Carcinoma: Diagnostic Performance of Multidetector CT and MR Imaging—A Systematic Review and Meta-Analysis. *Radiology.* April 2015;275(1):97–109.
37. Khalili K, Kim TK, Jang HJ, Haider MA, Khan L, Guindi M, u. a. Optimization of imaging diagnosis of 1–2cm hepatocellular carcinoma: An analysis of diagnostic performance and resource utilization. *J Hepatol.* April 2011;54(4):723–8.
38. Forner A, Vilana R, Ayuso C, Bianchi L, Solé M, Ayuso JR, u. a. Diagnosis of hepatic nodules 20 mm or smaller in cirrhosis: Prospective validation of the noninvasive diagnostic criteria for hepatocellular carcinoma. *Hepatology.* 10. Dezember 2007;47(1):97–104.
39. Sangiovanni A, Manini MA, Iavarone M, Romeo R, Forzenigo LV, Fraquelli M, u. a. The diagnostic and economic impact of contrast imaging techniques in the diagnosis of small hepatocellular carcinoma in cirrhosis. *Gut.* 1. Mai 2010;59(5):638–44.



40. Lee YH, Hsu CY, Chu CW, Liu PH, Hsia CY, Huang YH, u. a. A New Child-Turcotte-Pugh Class 0 for Patients with Hepatocellular Carcinoma: Determinants, Prognostic Impact and Ability to Improve the Current Staging Systems. PLOS ONE. 2014;9(6):11.
41. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PCJ. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. Hepatology. April 2000;31(4):864–71.
42. Eurotransplant [Internet]. Eurotransplant. [zitiert 9. August 2021]. Verfügbar unter: <https://www.eurotransplant.org/organs/liver/>
43. Kim WR, Wiesner RH, Edwards E. Hyponatremia and Mortality among Patients on the Liver-Transplant Waiting List. N Engl J Med. 2008;359(10):1018–26.
44. Oken, MM, Creech, RH, Tormey, DC, Horton, J, Davis, TE, McFadden, ET, u. a. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. Dezember 1982;5(6):649–55.
45. Llovet J, Brú C, Bruix J. Prognosis of Hepatocellular Carcinoma: The BCLC Staging Classification. Semin Liver Dis. 1999;19(03):329–38.
46. Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, u. a. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. J Hepatol. März 2022;76(3):681–93.
47. Roayaie S, Obeidat K, Sposito C, Mariani L, Bhoori S, Pellegrinelli A, u. a. Resection of hepatocellular cancer  $\leq 2$  cm: Results from two Western centers. Hepatology. April 2013;57(4):1426–35.
48. Llovet JM, Fuster J, Bruix J, Barcelona Clinic Liver Cancer (BCLC) Group. Intention-to-treat analysis of surgical treatment for early hepatocellular carcinoma: Resection versus transplantation. Hepatology. Dezember 1999;30(6):1434–40.
49. Berzigotti A, Reig M, Abraldes JG, Bosch J, Bruix J. Portal hypertension and the outcome of surgery for hepatocellular carcinoma in compensated cirrhosis: A systematic review and meta-analysis. Hepatology. Februar 2015;61(2):526–36.
50. Hasegawa K, Kokudo N, Makuuchi M, Izumi N, Ichida T, Kudo M, u. a.

Comparison of resection and ablation for hepatocellular carcinoma: A cohort study based on a Japanese nationwide survey. *J Hepatol.* April 2013;58(4):724–9.

51. N’Kontchou G, Mahamoudi A, Aout M, Ganne-Carrié N, Grando V, Coderc E, u. a. Radiofrequency ablation of hepatocellular carcinoma: Long-term results and prognostic factors in 235 Western patients with cirrhosis. *Hepatology.* November 2009;50(5):1475–83.

52. Choi D, Lim HK, Rhim H, Kim Y sun, Lee WJ, Paik SW, u. a. Percutaneous radiofrequency ablation for early-stage hepatocellular carcinoma as a first-line treatment: long-term results and prognostic factors in a large single-institution series. *Eur Radiol.* 12. Februar 2007;17(3):684–92.

53. Sala M, Llovet JM, Vilana R, Bianchi L, Solé M, Ayuso C, u. a. Initial response to percutaneous ablation predicts survival in patients with hepatocellular carcinoma. *Hepatology.* Dezember 2004;40(6):1352–60.

54. Mazzaferro V, Bhoori S, Sposito C, Bongini M, Langer M, Miceli R, u. a. Milan criteria in liver transplantation for hepatocellular carcinoma: An evidence-based analysis of 15 years of experience. *Liver Transpl.* Oktober 2011;17(S2):S44–57.

55. Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, u. a. Liver Transplantation for the Treatment of Small Hepatocellular Carcinomas in Patients with Cirrhosis. *N Engl J Med.* 14. März 1996;334(11):693–700.

56. Finn RS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim TY, u. a. Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *N Engl J Med.* 14. Mai 2020;382(20):1894–905.

57. Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, u. a. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. *Lancet Oncol.* Januar 2009;10(1):25–34.

58. Llovet JM, Hilgard P, de Oliveira AC, Forner A, Zeuzem S, Galle PR, u. a. Sorafenib in Advanced Hepatocellular Carcinoma. *N Engl J Med.* 2008;359(4):378–90.

59. Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, u. a. Lenvatinib versus

sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *The Lancet*. März 2018;391(10126):1163–73.

60. Shin SW. The Current Practice of Transarterial Chemoembolization for the Treatment of Hepatocellular Carcinoma. *Korean J Radiol*. 2009;10(5):425–34.

61. Breedis C, Young G. The blood supply of neoplasms in the liver. *Am J Pathol*. 1954;30(5):969–77.

62. Tsurusaki M, Murakami T. Surgical and Locoregional Therapy of HCC: TACE. *Liver Cancer*. 2015;4(3):165–75.

63. Ohishi, H, Uchida, H, Yoshimura, H, Ohue, S, Ueda, J, Katsuragi, M, u. a. Hepatocellular carcinoma detected by iodized oil. Use of anticancer agents. *Radiology*. Januar 1985;154(1):25–9.

64. Kan, Zuxing, McCuskey, Patricia A., Wright, Kenneth C., Wallace, Sidney. Role of Kupffer cells in iodized oil embolization. *Invest Radiol*. November 1994;29(11):990–3.

65. Varela M, Real MI, Burrel M, Forner A, Sala M, Brunet M, u. a. Chemoembolization of hepatocellular carcinoma with drug eluting beads: Efficacy and doxorubicin pharmacokinetics. *J Hepatol*. März 2007;46(3):474–81.

66. What is Harrell's C [Internet]. [statisticaloddsandends.wordpress.com](https://statisticaloddsandends.wordpress.com). 2021 [zitiert 24. August 2021]. Verfügbar unter: <https://statisticaloddsandends.wordpress.com/2019/10/26/what-is-harrells-c-index/>

67. Stephanie Glen. C-Statistic: Definition, Examples, Weighting and Significance [Internet]. [www.statisticshowto.com](http://www.statisticshowto.com). 2021 [zitiert 24. August 2021]. Verfügbar unter: <https://www.statisticshowto.com/c-statistic/>

68. Stephanie Glen. Receiver Operating Characteristic (ROC) Curve: Definition, Example [Internet]. [www.StatisticsHowTo.com](http://www.StatisticsHowTo.com). 2021 [zitiert 24. August 2021]. Verfügbar unter: <https://www.statisticshowto.com/receiver-operating-characteristic-roc-curve/>

69. Stephanie Glen. Brier Score: Definition, Examples [Internet].

StatisticsHowTo.com. 2021 [zitiert 25. August 2021]. Verfügbar unter: <https://www.statisticshowto.com/brier-score/>

70. Kronek LP, Reddy A. Logical analysis of survival data: prognostic survival models by detecting high-degree interactions in right-censored data. *Bioinformatics*. 15. August 2008;24(16):i248–53.

71. Sieghart W, Huckle F, Pinter M, Graziadei I, Vogel W, Müller C, u. a. The ART of decision making: Retreatment with transarterial chemoembolization in patients with hepatocellular carcinoma. *Hepatology*. Juni 2013;57(6):2261–73.

72. Adhoute X, Penaranda G, Naude S, Raoul JL, Perrier H, Bayle O, u. a. Retreatment with TACE: The ABCR SCORE, an aid to the decision-making process. *J Hepatol*. April 2015;62(4):855–62.

73. Gillmore R, Stuart S, Kirkwood A, Hameeduddin A, Woodward N, Burroughs AK, u. a. EASL and mRECIST responses are independent prognostic factors for survival in hepatocellular cancer patients treated with transarterial embolization. *J Hepatol*. Dezember 2011;55(6):1309–16.

74. Kim BK, Shim JH, Kim SU, Park JY, Kim DY, Ahn SH, u. a. Risk prediction for patients with hepatocellular carcinoma undergoing chemoembolization: development of a prediction model. *Liver Int*. Januar 2016;36(1):92–9.

75. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis*. Februar 2010;30(1):52–60.

76. Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol Rev*. 1958;65(6):386–408.

77. Luber S. Was ist ein Perzeptron? [Internet]. 2019 [zitiert 18. September 2021]. Verfügbar unter: <https://www.bigdata-insider.de/was-ist-ein-perzeptron-a-798367/>

78. Ranschaert E, Morozov S, Algra P. *Artificial Intelligence in Medical Imaging*. Springer Cham; 2019.

79. Kloeckner R, Pitton MB, Dueber C, Schmidtmann I, Galle PR, Koch S, u. a. Validation of Clinical Scoring Systems ART and ABCR after Transarterial Chemoembolization of Hepatocellular Carcinoma. *J Vasc Interv Radiol*. Januar

2017;28(1):94–102.

80. Mähringer-Kunz A, Weinmann A, Schmidtmann I, Koch S, Schotten S, Pinto dos Santos D, u. a. Validation of the SNACOR clinical scoring system after transarterial chemoembolisation in patients with hepatocellular carcinoma. *BMC Cancer*. Dezember 2018;18(1):489.
81. Terzi E, Terenzi L, Venerandi L, Croci L, Renzulli M, Mosconi C, u. a. The ART Score Is Not Effective to Select Patients for Transarterial Chemoembolization Retreatment in an Italian Series. *Dig Dis*. 2014;32(6):711–6.
82. Cucchetti A, Vivarelli M, Heaton ND, Phillips S, Piscaglia F, Bolondi L, u. a. Artificial neural network is superior to MELD in predicting mortality of patients with end-stage liver disease. *Gut*. 1. Februar 2007;56(2):253–8.
83. Cucchetti A, Piscaglia F, Grigioni AD, Ravaioli M, Cescon M, Zanello M, u. a. Preoperative prediction of hepatocellular carcinoma tumour grade and micro-vascular invasion by means of artificial neural network: A pilot study. *J Hepatol*. Juni 2010;52(6):880–8.
84. Weinmann A, Koch S, Niederle IM, Schulze-Bergkamen H, König J, Hoppe-Lotichius M, u. a. Trends in Epidemiology, Treatment, and Survival of Hepatocellular Carcinoma Patients Between 1998 and 2009: An Analysis of 1066 Cases of a German HCC Registry. *J Clin Gastroenterol*. März 2014;48(3):279–89.
85. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med*. 6. Januar 2015;162(1):55–63.
86. Mähringer-Kunz A, Wagner F, Hahn F, Weinmann A, Brodehl S, Schotten S, u. a. Predicting survival after transarterial chemoembolization for hepatocellular carcinoma using a neural network: A Pilot Study. *Liver Int*. März 2020;40(3):694–703.
87. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, u. a. Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases: Marrero et al. *Hepatology*. August 2018;68(2):723–50.
88. Lammer J, Malagari K, Vogl T, Pilleul F, Denys A, Watkinson A, u. a.

Prospective Randomized Study of Doxorubicin-Eluting-Bead Embolization in the Treatment of Hepatocellular Carcinoma: Results of the PRECISION V Study. *Cardiovasc Intervent Radiol*. Februar 2010;33(1):41–52.

89. Lencioni R, de Baere T, Burrel M, Caridi JG, Lammer J, Malagari K, u. a. Transcatheter Treatment of Hepatocellular Carcinoma with Doxorubicin-loaded DC Bead (DEBDOX): Technical Recommendations. *Cardiovasc Intervent Radiol*. Oktober 2012;35(5):980–5.

90. Raffetti E, Portolani N, Molfino S, Baiocchi GL, Limina RM, Caccamo G, u. a. Role of aetiology, diabetes, tobacco smoking and hypertension in hepatocellular carcinoma survival. *Dig Liver Dis*. November 2015;47(11):950–6.

91. Kolly P, Knöpfli M, Dufour JF. Effect of smoking on survival of patients with hepatocellular carcinoma. *Liver Int*. November 2017;37(11):1682–7.

92. Fujiwara N, Nakagawa H, Kudo Y, Tateishi R, Taguri M, Watadani T, u. a. Sarcopenia, intramuscular fat deposition, and visceral adiposity independently predict the outcomes of hepatocellular carcinoma. *J Hepatol*. Juli 2015;63(1):131–40.

93. Siriwardana RC, Liyanage CAH, Gunetilleke B, Niriella MA, de Silva J, Dassanayake AS, u. a. Diffuse-Type Hepatoma: A Grave Prognostic Marker. *Gastrointest Tumors*. 2017;4(1–2):20–7.

94. Kamath P. A model to predict survival in patients with end-stage liver disease. *Hepatology*. Februar 2001;33(2):464–70.

95. Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, u. a. Assessment of Liver Function in Patients With Hepatocellular Carcinoma: A New Evidence-Based Approach—The ALBI Grade. *J Clin Oncol*. 20. Februar 2015;33(6):550–8.

96. Lin ZH, Li X, Hong YF, Ma XK, Wu DH, Huang M, u. a. Alanine aminotransferase to hemoglobin ratio is an indicator for disease progression for hepatocellular carcinoma patients receiving transcatheter arterial chemoembolization. *Tumor Biol*. März 2016;37(3):2951–9.

97. Pang Q, Qu K, Zhang JY, Song SD, Liu SS, Tai MH, u. a. The Prognostic Value of Platelet Count in Patients With Hepatocellular Carcinoma: A Systematic Review and

Meta-Analysis. *Medicine (Baltimore)*. September 2015;94(37):e1431.

98. Biolato M. Hepatocellular carcinoma treated by conventional transarterial chemoembolization in field-practice: Serum sodium predicts survival. *World J Gastroenterol*. 2014;20(25):8158–65.

99. Cross, SS, Harrison, RF, Kennedy, RL. Introduction to neural networks. *Lancet*. 21. Oktober 1995;346(8982):1075–9.

100. Shi HY, Lee KT, Lee HH, Ho WH, Sun DP, Wang JJ, u. a. Comparison of Artificial Neural Network and Logistic Regression Models for Predicting In-Hospital Mortality after Primary Liver Cancer Surgery. Coleman WB, Herausgeber. *PLoS ONE*. 26. April 2012;7(4):e35781.

101. Kloeckner, Roman. Predicting Survival after Transarterial Chemoembolization for Hepatocellular Carcinoma Using a Neural Network: A Pilot Study. Mendeley Data [Internet]. 2019 [zitiert 10. September 2021];V2. Verfügbar unter: <https://data.mendeley.com/datasets/cdf8tb3pxm/2>

102. Wang H, Naghavi M, Allen C, Barber RM, Bhutta ZA, Carter A, u. a. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*. Oktober 2016;388(10053):1459–544.

103. Arizumi T, Ueshima K, Iwanishi M, Minami T, Chishina H, Kono M, u. a. Evaluation of ART Scores for Repeated Transarterial Chemoembolization in Japanese Patients with Hepatocellular Carcinoma. *Oncology*. 19. November 2015;89(2):4–10.

104. Facciorusso A, Bhoori S, Sposito C, Mazzaferro V. Repeated transarterial chemoembolization: An overfitting effort? *J Hepatol*. Juni 2015;62(6):1440–2.

105. National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. Washington, D.C.: National Academies Press; 2019.

106. Peng J, Kang S, Ning Z, Deng H, Shen J, Xu Y, u. a. Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from CT imaging. *Eur Radiol*. Januar 2020;30(1):413–24.

107. Abajian A, Murali N, Savic LJ, Laage-Gaupp FM, Nezami N, Duncan JS, u. a.

Predicting Treatment Response to Intra-arterial Therapies for Hepatocellular Carcinoma with the Use of Supervised Machine Learning—An Artificial Intelligence Concept. *J Vasc Interv Radiol.* Juni 2018;29(6):850-857.e1.

108. Qiao G, Li J, Huang A, Yan Z, Lau WY, Shen F. An artificial neural networking model for the prediction of post-hepatectomy survival of patients with early hepatocellular carcinoma: ANN predicting survival of EHCC. *J Gastroenterol Hepatol.* Dezember 2014;29(12):2014–20.

109. Chiu HC, Ho TW, Lee KT, Chen HY, Ho WH. Mortality Predicted Accuracy for Hepatocellular Carcinoma Patients with Hepatic Resection Using Artificial Neural Network. *Sci World J.* 2013;2013:1–10.

110. Ho WH, Lee KT, Chen HY, Ho TW, Chiu HC. Disease-Free Survival after Hepatic Resection in Hepatocellular Carcinoma Patients: A Prediction Approach Using Artificial Neural Network. Algül H, Herausgeber. *PLoS ONE.* 3. Januar 2012;7(1):e29179.

111. Wu CF, Wu YJ, Liang PC, Wu CH, Peng SF, Chiu HW. Disease-free survival assessment by artificial neural networks for hepatocellular carcinoma patients after radiofrequency ablation. *J Formos Med Assoc.* Oktober 2017;116(10):765–73.

112. Fernandez-Delgado M, Cernadas E, Barro S, Amorim D. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *J Mach Learn Res.* 2014;15:3133–81.

113. Xu X, Zhang HL, Liu QP, Sun SW, Zhang J, Zhu FP, u. a. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *J Hepatol.* Juni 2019;70(6):1133–44.

114. Ni M, Zhou X, Lv Q, Li Z, Gao Y, Tan Y, u. a. Radiomics models for diagnosing microvascular invasion in hepatocellular carcinoma: which model is the best model? *Cancer Imaging.* Dezember 2019;19(1):60.

115. Yuan C, Wang Z, Gu D, Tian J, Zhao P, Wei J, u. a. Prediction early recurrence of hepatocellular carcinoma eligible for curative ablation using a Radiomics nomogram. *Cancer Imaging.* Dezember 2019;19(1):21.

116. Golfieri R, Giampalma E, Renzulli M, Cioni R, Bargellini I, Bartolozzi C, u. a.



Randomised controlled trial of doxorubicin-eluting beads vs conventional chemoembolisation for hepatocellular carcinoma. *Br J Cancer*. Juli 2014;111(2):255–64.

117. Kudo M. Systemic Therapy for Hepatocellular Carcinoma: Latest Advances. *Cancers*. 30. Oktober 2018;10(11):412.

118. Mason MC, Massarweh NN, Salami A, Sultenfuss MA, Anaya DA. Post-embolization syndrome as an early predictor of overall survival after transarterial chemoembolization for hepatocellular carcinoma. *HPB*. Dezember 2015;17(12):1137–44.

## 7 Danksagung

Ich bedanke mich an dieser Stelle bei all denjenigen, die mich während der Anfertigung meiner Doktorarbeit unterstützt haben.

Der größte Dank gebührt meinem Betreuer für die großartige Begleitung dieser Doktorarbeit. Herzlichen Dank für jede augenblicklich beantwortete E-Mail, jedes Gespräch und überhaupt jede Minute Zeit, die du dir genommen hast.

Vielen Dank auch an meinen Mann, der mich in den letzten 6,5 Jahren nicht nur während des Studiums unterstützt, sondern auch zu vielen weiteren Abenteuern ermutigt hat.

Nicht zuletzt möchte ich meinen Eltern, Schwiegereltern und meinem Bruder danken, die während des Studiums und zum Anfertigen dieser Arbeit viele Stunden auf unsere kleinen Matröschen aufgepasst haben.