

Retrospektive Identifikation von Risikogruppen sowie Entwicklung
und Validierung eines Vorhersagemodells für das Gesamtüberleben
von Patienten mit Parotiskarzinomen

Inauguraldissertation
zur Erlangung des Grades eines Doktors der Zahnmedizin des
Fachbereichs Medizin
der Justus-Liebig-Universität Gießen

vorgelegt von
Krauss, Christian Alexander

aus Karlsruhe

Gießen 2020

Aus der Klinik für Hals-Nasen-Ohrenheilkunde, Kopf-Halschirurgie
und plastische Operationen unter der kommissarischen Leitung von
Dr. Gregor Wolf des Fachbereichs Medizin
der Justus-Liebig-Universität Gießen

Gutachter: Prof. Dr. Claus Wittekindt
Gutachter: Prof. Dr. Henning Schneider

Tag der Disputation: 06.10.2020

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Anatomie und Epidemiologie	1
1.2	Diagnostik	2
1.3	Prognose.....	6
1.4	Statistik in der Überlebenszeitmodellierung.....	6
1.5	Fragestellung	8
2	Material und Methoden	9
2.1	Retrospektive Datengrundlage	9
2.2	Einschlusskriterien	10
2.3	Umgang mit fehlenden Werten	11
2.4	Kohorteneinteilung und deskriptive Statistik	11
2.5	Univariate Überlebenszeitanalyse nach Kaplan-Meier/Cox und Variablenselektion....	11
2.6	Multivariate Überlebenszeitanalyse mit Cox-Regression, Nomogramm und WebApp	12
2.7	Klassifikations- und Regressionsbäume / Rekursives Partitionieren.....	12
3	Ergebnisse.....	14
3.1	Deskriptive Analyse von patientenbezogenen Merkmalen	14
3.1.1	Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich des Geschlechts	17
3.1.2	Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich der Altersgruppe.....	18
3.1.3	Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich der Histologie	19
3.2	Univariate Analyse der Einflussfaktoren auf die Überlebenszeit	21
3.2.1	Standort	22
3.2.2	Geschlecht und Alter	23
3.2.3	Histologie	24
3.2.4	TNM- und UICC-Stadien.....	26
3.2.5	Malignität.....	28
3.2.6	Therapieverfahren/radikale OP/Neck-Dissection/Rezidiv	29
3.3	Multivariate Cox-Regression und Modellbildung	31
3.4	Nomogramm	34
3.4.1	Entwicklung.....	34
3.4.2	Validierung.....	35
3.5	Rekursive Partitionsanalyse / Klassifikations- und Regressionsbaum	37
3.5.1	Entwicklung.....	37
3.5.2	Validierung.....	40
3.6	Praktische Umsetzung des Modells in Form einer Webanwendung	44

4	Diskussion	46
4.1	Prognosefaktoren	46
4.1.1	Geschlecht und Alter	46
4.1.2	Histologie	47
4.1.3	TNM/UICC-Staging	48
4.1.4	Malignität	49
4.2	Variablenselektion / Cox-Regression / Nomogramm	49
4.3	Rekursive Partitionsanalyse / Klassifikations- und Regressionsbaum	51
4.4	Überblick über bereits publizierte Modelle	52
4.5	Praktische Anwendbarkeit der Ergebnisse	54
4.6	Ausblick	54
4.7	Stärken und Grenzen der Arbeit	55
5	Zusammenfassung	58
6	Summary	59
	Abkürzungsverzeichnis	60
	Abbildungsverzeichnis	62
	Tabellenverzeichnis	64
	Literaturverzeichnis	66
	Anhang	72
	Publikationsverzeichnis	79
	Erklärung zur Dissertation	80
	Danksagung	81
	Lebenslauf	82

1 Einleitung

Maligne Speicheldrüsentumore gelten als selten und zeichnen sich durch ihre histologische Vielfalt aus (Ali et al. 2013). Für Ärzte, Patienten und das Gesundheitssystem ist es von großem Interesse, Krankheitsverläufe abhängig von der Diagnose und der gewählten Therapie vorherzusagen, um eine Prognose abzuschätzen (Royston und Altman 2013). Die histologische Vielfalt und das seltene Auftreten der Speicheldrüsenkarzinome erschweren die Abschätzung der Prognose. Mit TNM- und UICC-*Staging* haben sich einfache Prognose-Systeme etabliert, welche als „Goldstandard“ anerkannt sind. Zahlreiche klinische Variablen (wie Patientenalter oder Histologie des Tumors), welche einen nachgewiesenen signifikanten Einfluss auf das Überleben haben, bleiben beim TNM- und UICC-*Staging* jedoch unberücksichtigt (Chou et al. 2017; Ali et al. 2014). TNM und vergleichbare *Staging*-Systeme werden daher mittlerweile als unzureichend angesehen (Kreppel et al. 2013).

In den letzten Jahren wurden mehrere fortschrittlichere Prognosemodelle publiziert. Die bisher veröffentlichten Arbeiten weisen aufgrund der niedrigen Fallzahl oder der Beschränkung auf Daten von nur einem Standort jedoch Schwächen auf.

1.1 Anatomie und Epidemiologie

Der menschliche Körper verfügt über große und kleine Kopfspeicheldrüsen. Die drei großen Kopfspeicheldrüsen sind die Ohrspeicheldrüse (Glandula parotis), die Unterkieferspeicheldrüse (Glandula submandibularis) und die Unterzungenspeicheldrüse (Glandula sublingualis). Die meisten Speicheldrüsen gehören zur Gruppe der kleinen Speicheldrüsen, welche sich in der Mukosa des Mundhöhlenepithels am Gaumen, der Wange, der Lippen oder der Zunge befinden. Im Gegensatz zu den großen Speicheldrüsen münden sie meist direkt in die Mundhöhle (Reiß 2009, S. 550–555).

Typische Erkrankungen der Speicheldrüsen sind Verletzungen, entzündliche Prozesse (z.B. Parotitis), Steinbildung (Sialolithiasis), Zysten, Sialadenosen und Tumoren. Diese Arbeit beschäftigt sich mit malignen Erkrankungen der Glandula Parotis. Es handelt sich dabei um eine rein seröse Drüse, welche sich direkt vor dem äußeren Ohr befindet. Sie findet ihre anatomischen Grenzen am Vorderrand des Musculus massetericus, dem Jochbeinbogen und dem Unterkieferwinkel (Reiß 2009, S. 550–555).

Für die USA wurde festgestellt, dass 6% aller malignen Kopf-Hals-Tumoren Speicheldrüsenkarzinome sind. Diese repräsentieren 0,3% aller malignen Tumoren (Barnes et al. 2005, S. 213). Bösartige Tumoren sind u.a. gekennzeichnet durch infiltratives Wachstum, atypische Mitosen, regionären Lymphknotenbefall und einer Pleomorphie der Tumorzellen (Reiß 2009, S. 550–555).

Die Ursachen und Risikofaktoren zur Entstehung von malignen Speicheldrüsentumoren sind nicht komplett geklärt. Als mögliche Ätiologien werden Diät, Bestrahlungstherapie, Krebsvorgeschichte, genetische Faktoren, die berufliche Tätigkeit in einer Gummifabrik oder einem Kosmetikstudio, Nickelexposition, Rauchen, Alkoholkonsum, verarbeitetes Fleisch, Adipositas, Virenexposition (HIV, EBV, HPV) und diverse weitere Lebensstilfaktoren genannt (Carlson et al. 2013; Pan et al. 2017).

1.2 Diagnostik

Die Diagnostik von Speicheldrüsentumoren besteht üblicherweise aus einer körperlichen und zytologischen Untersuchung sowie bildgebenden Verfahren. Klinisch erscheinen maligne Speicheldrüsentumoren meist zunächst als schmerzlose, tastbare Schwellungen, welche sich schwer von benignen Tumoren abgrenzen lassen (Ettl et al. 2012). Die Feinnadelaspirations-Zytologie kann ultraschallunterstützt angewandt werden, um bereits präoperativ Kenntnisse über die Malignität des Tumors zu erlangen. Dieser Technik wird eine hohe Spezifität und eine niedrige Sensitivität zugerechnet. Eine neue Alternative dazu ist die Vakuumbiopsie (Ettl et al. 2012; Carlson et al. 2013). Eine definitive pathohistologische Diagnose des Tumors steht oftmals allerdings erst nach der OP fest. Dies stellt ein Problem dar, da manche Therapieentscheidungen wie Ausmaß der Resektion, adjuvante Therapien oder *Neck-Dissection* abhängig von der Histologie getroffen werden können (Ettl et al. 2012). Die Sonografie ist ein nicht-invasives bildgebendes Verfahren zur Darstellung der Drüsen und der Lymphknoten. Die Bildgebung dient der Abschätzung der Tumorgrenze sowie der OP-Planung. Mit einem Kontrastmittel-MRT kann darüber hinaus zwischen malignen und benignen Tumoren unterschieden werden. Das Kontrastmittel-CT stellt eine alternative Bildgebung dar. Das MRT ist allerdings vorzuziehen, da Spezifität und Sensitivität besser sind. Im *Staging* und der Tumornachsorge hat sich das PET-CT als ein probates Mittel erwiesen. Die dabei eingesetzte radioaktiv markierte Glukose zeigt seine Stärken in der Darstellung

von stoffwechselaktiven Zellverbänden (z.B. bei Fernmetastasen und Lokalrezidiven) (Ettl et al. 2012).

Zur Einteilung der Tumorerkrankungen nach ihrem „Schweregrad“ wurden mehrere Systeme entwickelt. Der TNM-Status stellt dabei unabhängig von der Tumorentität den weltweit akzeptierten Standard dar und wird regelmäßig überarbeitet und angepasst; siehe Tabelle 1 (Ali et al. 2014; Brierley et al. 2017).

Tabelle 1: TNM-Klassifikation für große Speicheldrüsen (8. Auflage) (Brierley et al. 2017).

T-Primärtumor	
T0	Kein Tumor feststellbar
T1	Tumor ≤ 2cm in der größten Ausdehnung und ohne extraparenchymatöse Ausbreitung
T2	Tumor >2cm aber nicht > 4cm in der größten Ausdehnung ohne extraparenchymatöse Ausbreitung
T3	Tumor > 4cm und/oder Tumor mit extraparenchymatöse Ausbreitung
T4a	Tumor infiltriert Haut, Mandibula, Gehörgang oder N. facialis
T4b	Tumor infiltriert Schädelbasis, Proc. pterygoideus oder Ummauerung von A. carotis
N-Regionärer Lymphknotenbefall	
N0	Kein regionärer Lymphknotenbefall feststellbar
N1	Eine ipsilaterale LK-Metastase, ≤ 3cm in der größten Ausdehnung
N2a	Eine ipsilaterale LK-Metastase, < 3cm in der größten Ausdehnung mit extranodaler Ausbreitung oder 3cm bis 6cm in der größten Ausdehnung ohne extranodale Ausbreitung
N2b	Multiple ipsilaterale LK-Metastasen, keine > 6cm in der größten Ausdehnung und ohne extranodale Ausbreitung
N2c	Bilaterale oder kontralaterale LK-Metastasen, keine > 6cm in der größten Ausdehnung und ohne extranodale Ausbreitung
N3a	LK-Metastase > 6cm in der größten Ausdehnung und ohne extranodale Ausbreitung
N3b	LK-Metastase > 3cm mit extranodaler Ausbreitung oder multiple LK-Metastasen sowie kontralaterale/bilaterale Metastase mit extranodaler Ausbreitung
M-Fernmetastasen	
M0	Keine Fernmetastasen
M1	Fernmetastasen

Das AJCC/UICC-*Staging* hat ebenfalls eine große Verbreitung gefunden und ist im klinischen Alltag bewährt. Es setzt sich aus den jeweiligen TNM-Stadien zusammen und stellt eine Vereinfachung dar; siehe Tabelle 2 (Brierley et al. 2017; Edge 2010; Chou et al. 2017).

Tabelle 2: UICC-*Staging* für große Speicheldrüsen (8. Auflage) (Brierley et al. 2017).

	T	N	M
0	Tis	N0	M0
I	T1	N0	M0
II	T2	N0	M0
III	T3	N0	M0
	T1, T2, T3	N1	M0
IVA	T1, T2, T3	N2	M0
	T4a	N0, N1, N2	M0
IVB	T4b	Alle N	M0
	Alle T	N3	M0
IVC	Alle T	Alle N	M1

Eine Übersichtsarbeit über *Staging*-Systeme bei Parotis-Karzinomen kommt zu dem Schluss, dass das alleinige Anwenden von tumorbezogenen Kriterien (wie des TNM) für die Prognoseabschätzung nicht mehr angemessen ist (Kreppel et al. 2013).

Die pathohistologische Diagnostik erfolgt nach einer WHO-Klassifikation, welche ebenfalls regelmäßig aktualisiert wird (siehe Tabelle 3). Die Klassifikation von 2017 enthält 20 unterschiedliche histologische Entitäten. Die Zusammensetzung ändert sich je nach Version der Klassifikation. Im Vergleich zur Klassifikation von 2005 fand eine Reduktion um drei Entitäten statt. Die bisher als eigenständig geführten Entitäten Zystadenokarzinom, muzinöses Adenokarzinom und das intestinale Adenokarzinom werden nun wieder unter Adenokarzinom *NOS* gemeinsam geführt. Mit dem mamma-analogen sekretorischen Karzinom kam eine neue Entität hinzu (Ihrler et al. 2018; El-Naggar et al. 2017).

Tabelle 3: Übersicht der unterschiedlichen WHO-Klassifikationen (Seifert et al. 1990; Barnes et al. 2005; El-Naggar et al. 2017)

WHO Klassifikation der malignen Speicheldrüsentumoren					
1991 (Seifert/Sobin) und früher	ICD-O	2005 (Barnes et. al.)	ICD-O	2017 (El-Naggar et. al., IARC)	ICD-O
1. Azinuszellkarzinom	8550/3	1. Azinuszellkarzinom	8550/3	1. Azinuszellkarzinom	8550/3
2. Mukoepidermoidkarzinom	8430/3	2. Mukoepidermoidkarzinom	8430/3	2. Mukoepidermoidkarzinom	8430/3
3. Adenoid-zystisches Karzinom	8200/3	3. Adenoid-zystisches Karzinom	8200/3	3. Adenoid-zystisches Karzinom	8200/3
4. Polymorphes Low-grade-Adenokarzinom (terminales Gangadenokarzinom)	8525/3	4. Polymorphes Low-Grade Adenokarzinom	8525/3	4. Polymorphes Adenokarzinom	8525/3
5. Epithelial-myoepitheliales Karzinom		5. Epithelial-myoepitheliales Karzinom	8562/3	5. Epithelial-myoepitheliales Karzinom	8562/3
6. Basalzelladenokarzinom	8147/3	6. Basalzelladenokarzinom	8147/3	6. Basalzelladenokarzinom	8147/3
7. Talgdrüsenkarzinom	8410/3	7. Talgdrüsenkarzinom	8410/3	7. Talgdrüsenadenokarzinom	8410/3
8. Papilläres Zystadenokarzinom	8450/3	8. Talgdrüsenlymphadenokarzinom	8410/3		
		9. Zystadenokarzinom/Low-Grade cribriformes Zystadenokarzinom	8440/3		
9. Muzinöses Adenokarzinom	8480/3	10. Muzinöses Adenokarzinom	8480/3		
10. Onkozytöses Karzinom	8290/3	11. Onkozytöses Karzinom	8290/3	8. Onkozytöses Karzinom	8290/3
11. Speichelgangkarzinom	8500/3	12. Speichelgangskarzinom	8500/3	9. Speichelgangskarzinom	8500/3
12. Adenokarzinom	8140/3	13. Adenokarzinom, NOS	8140/3	10. Intraduktales Karzinom	8500/2
13. Malignes Myoepithelium (malignes myoepitheliales Karzinom)	8982/3	14. Myoepitheliales Karzinom	8982/3	11. Adenokarzinom, NOS	8140/3
14. Karzinom im pleomorphen Adenom(bösartiger Misch tumor)	8941/3	15. Karzinom im pleomorphen Adenom	8941/3	12. Myoepitheliales Karzinom	8982/3
15. Plattenepithelkarzinom	8070/3	16. Metastasierendes pleomorphes Adenom	8940/1	13. Karzinom im pleomorphen Adenom	8941/3
16. Kleinzelliges Karzinom	8041/3	17. Plattenepithelkarzinom	8070/3	14. Plattenepithelkarzinom	8070/3
17. Undifferenziertes Karzinom	8020/3	18. Kleinzelliges Karzinom	8041/3	15. Gering differenziertes Karzinom (undifferenziertes Karzinom, kleinzelliges neuroendokrines Karzinom, großzellig neuroendokrines Karzinom)	8020/3
18. Andere Karzinome		19. Großzelliges Karzinom	8012/3		8041/3
		20. Klarzellkarzinom, NOS	8310/3		8013/3
		21. Karzinosarkom	8980/3	16. Klarzellkarzinom	8310/3
		22. Lymphoepitheliales Karzinom	8082/3	17. Karzinosarkom	8980/3
		23. Sialoblastom	8974/1	18. Lymphoepitheliales Karzinom	8082/3
				19. Sialoblastom	8974/1
				20. (Mamma-analogenes) sekretorisches Karzinom	8502/3

Grün: Seit 2017 unter Adenokarzinom NOS zusammengefasste Entitäten, Gelb: Gering differenziertes Karzinom fasst die bisherigen Entitäten des kleinzelligen Karzinoms und des großzelligen Karzinoms zusammen, Blau: Neue Entität

1.3 Therapie

Die chirurgische Resektion des Tumors in einem Schritt (Parotidektomie) unter Schonung des Nervus facialis stellt die Therapie der Wahl dar. Je nach Ausdehnung des Tumors unterscheidet man zwischen kompletter und partieller Parotidektomie (National Comprehensive Cancer Network 2019). Bei chirurgischen Eingriffen in diesem Gebiet besteht immer eine Paralysegefahr durch iatrogene Nervverletzung. Sofern der Nervus facialis bereits durch das Tumorleiden befallen wurde, wird dieser mitentfernt und rekonstruiert. Als Transplantat eignet sich dabei z.B. der Nervus auricularis (Ettl et al. 2012; Carlson et al. 2013). Es gibt mehrere Möglichkeiten den Verlauf des Nervs während einer OP darzustellen. Dies gelingt beim „*nerve mapping*“ durch perkutane Stimulation und Registrierung der Muskelaktionspotentiale (CMAP). Intraoperativ wird meist ein Facialis-Monitoring durchgeführt, welches auf der Elektromyografie basiert. Dadurch kann bei einer Parotidektomie eine Verletzung des Nervs vermieden werden (Guntinas-Lichius et al. 2018).

Ein positiver Lymphknotenbefall steht für die Ausbreitung des Tumors und führt grundsätzlich zu einem geringeren Überleben. Daher werden bei entsprechenden Fällen die regionären Lymphknoten chirurgisch entfernt. Man unterscheidet dabei zwischen therapeutischer und prophylaktischer Lymphknotenausräumung (*Neck-Dissection*). Die therapeutische *Neck-Dissection* wird immer angewandt, sobald ein klinisches oder bildgebendes Anzeichen für mögliche regionäre Metastasierung vorliegt. Wenn kein definitiver Lymphknotenbefall feststeht, ist eine prophylaktische *Neck-Dissection* möglich, deren Indikationsbereich hingegen ärztliche Auslegungssache ist (Ettl et al. 2012). Unter adjuvanter Therapie versteht man die zusätzliche Radio- oder Chemotherapie als Ergänzung zur chirurgischen Resektion. Eine Bestrahlung wird immer dann durchgeführt, wenn das Tumorstadium schon fortgeschritten ist und begonnen hat zu metastasieren (regionärer Lymphknotenstatus, infiltratives Wachstum). Studien belegen, dass die adjuvante Radiotherapie lokale Tumorrezidive verringert. Sie wird bei Metastasen und Rezidiven angewandt, welche weder chirurgisch noch radioonkologisch therapierbar sind. Als rein palliative Therapie wird die alleinige Chemotherapie angesehen (Ettl et al. 2012). In den letzten Jahren ist die gezielte Therapie (*targeted therapy*) bedeutsamer geworden. Dabei handelt es sich um zielgerichtete, medikamentöse Eingriffe in die Signalkaskaden, welche das Tumorstadium steuern (z.B. Therapie der chronischen myeloischen Leukämie: Imatinib führt zur Blockade des KIT-Rezeptor (Hantschel et al. 2008)). Je nach

molekularen Wechselwirkungen entsteht dabei ein individueller Therapieansatz (Ettl et al. 2012; Keller et al. 2017).

1.4 Prognose

Die Seltenheit und Heterogenität der malignen Speicheldrüsentumoren erschweren Aussagen über die Prognose der Patienten. In der Vergangenheit wurden zahlreiche statistische Auswertungen durchgeführt, um den Einfluss von Faktoren auf das Überleben festzustellen. Die Prognose hängt unter anderem von der Tumorgroße (T-Stadium), Lymphknotenbefall (N-Stadium), Fernmetastasierung (M-Stadium), Alter, Geschlecht, Perineuralinvasion, Infiltration angrenzenden Gewebes, Nervbeteiligung, Gefäßinvasion, Malignität (*Grading*), Resektionsstatus, Therapie, Komorbidität, molekulare Marker wie ki67 etc. ab und differiert stark innerhalb der Literatur (siehe Tabelle 4).

Tabelle 4: Übersicht über prognostische Faktoren für die Überlebenszeit von Parotiskarzinomen in der Literatur

Autor	n	Faktoren
Spiro (1989)	380	Stadium, Histologie, LN, Facialisfehlfunktion, Alter
Kane (1991)	194	Stadium, Malignität, Symptome, Geschlecht, Alter
Calearo (1998)	167	Stadium, Malignität, Facialisfehlfunktion
Vander Poorten (1999)	151	T-Stage, N-Stage, Alter, Symptome, Ausbreitung ins Gesicht, Facialisfehlfunktion
Harbo (2002)	152	TNM, lokale Ausbreitung, Malignität
Terhaard (2004)	565	T-Stage, N-Stage, Tumorlokalisation, Gesichts- und Knocheninvasion, Alter, Geschlecht
Bhattacharyya (2004)	903	Tumorgroße, Malignität, extraglanduläre Ausdehnung, LN, Radio-Therapie, Alter
Pohar (2005)	163	Stadium, Alter
Lima (2005)	126	T-Stage, N-Stage, Malignität, Alter
Cederblad (2009)	144	N-Stage, Alter, Facialisfehlfunktion
Björndal (2012)	871	Stadium 3+ 4, Malignität, Gefäßinvasion, Tumorlatenz, Resektionsränder, Alter
Ali (2014)	301	T4-Stage, Tumorgroße, Malignität, Perineuralinvasion, Alter
Erovc (2015)	215	N-Stage, Perineural- und Lymphgefäßinvasion, Alter
Li (2017)	4218	TNM, Tumorlokalisation, Malignität, Therapie (OP/Radiatio), Alter, Geschlecht
Paderno (2018)	104	N-Stage, Malignität, Perineural- und Lymphgefäßinvasion, Gesichtsinvasion
Chakrabarti (2018)	224	T-Stage, N-Stage, Resektionsränder, Malignität, Perineural- und Lymphgefäßinvasion, Alter

1.5 Statistik in der Überlebenszeitmodellierung

Vorhersagemodelle finden in zahlreichen Gebieten der Medizin Anwendung. Als klassische Verfahren haben sich die univariate Kaplan-Meier-Überlebenszeitanalyse und das multivariate Cox-Proportional-Hazard-Regressions-Modell etabliert. Bei der Cox-Regression handelt es sich um ein semi-parametrisches Verfahren. Es enthält parametrische (über den Prädiktoreneffekt) und nicht-parametrische Annahmen (*Baseline-Hazard-Funktion*) (Steyerberg 2009, S. 77). Dabei wird eine *Hazard-Funktion*

gebildet, welche das relative Sterberisiko angibt. Die Grundannahme des *Cox-Proportional-Hazard-Modells* ist die Konstanz des Sterberisikos (*Hazard Ratio*) über die Zeit (Harrell 2015, S. 475). Das Ergebnis einer Cox-Regression kann numerisch (durch die *Hazard Ratios*) als auch grafisch (durch ein Nomogramm) dargestellt werden. Durch ein Nomogramm wird die Berechnung einer individualisierten Prognose ermöglicht; Beispiel siehe Ali et al. 2013. Nach der Modellentwicklung kann die Modellgüte bewertet werden. Man spricht dabei von Validierung. Es gibt mehrere Möglichkeiten, welche unterschiedlich hohe Qualität besitzen. Die Cox-Regression wird durch das *Bootstrapping* intern validiert. Das bedeutet, dass ein Teil des Trainingsdatensatzes auf das Modell angewandt wird, um die Vorhersagen zu vergleichen. Der Goldstandard ist jedoch das Anwenden von Daten, welche nicht für die Entwicklung des Modells verwendet wurden, also unabhängig vom Modell sind (Hay et al. 2018). Die Methoden zum Beschreiben der Vorhersagekraft nennen sich Kalibrierung und Diskriminierung (Royston und Altman 2013). Diskriminierung beschreibt die Vorhersagekraft eines Modells insofern, dass das Modell die Vorhersagen aus einem Spektrum mehrerer Prognosen auswählt und diese dann real zutreffen. Kalibrierung bedeutet, dass ein Modell für einen bestimmten Anteil einer Fallgruppe ein bestimmtes Ereignis/Wahrscheinlichkeit (z.B. Todesfall) vorhersagen kann. Dies wird mit der tatsächlich eingetroffenen Ereignis/Wahrscheinlichkeit verglichen. Dabei entsteht ein sog. Kalibrierungsdiagramm (Harrell 2015, S. 506; Royston und Altman 2013; Ellwanger 2001, S. 26).

Alternativ zu den klassischen Methoden haben sich die „modernen“ *Machine-Learning*-Methoden, wie neuronale Netze und Klassifikations- und Regressionsbäume (***Classification and Regression Trees***, CART), entwickelt. *Machine-Learning* ist eng verknüpft mit den Begriffen *Data-Mining*, *Knowledge Discovery in Databases* und Künstlicher Intelligenz. Dabei wird aus Erfahrung Wissen generiert, d.h. das Modell lernt in den zu Verfügung gestellten Trainingsdaten Muster zu erkennen und anschließend zu verallgemeinern. *Machine-Learning* findet sich schon in vielen Bereichen unseres Alltags, z.B. bei der Erkennung von Kreditkartenbetrug, Text-/Spracherkennung und zunehmend auch in der Medizin (Kourou et al. 2015). CART sind eine Methode, komplexe Interaktionen darzustellen und anwendbar zu machen – ein Beispiel für einen Klassifikations- und Regressionsbaum findet sich in Wagner et al. 2017. Bei den CART wird eine sog. Zielvariable definiert (z.B. das Gesamtüberleben). Die Gesamtgruppe an Fällen wird dann Schritt-für-Schritt anhand der Variablen in zwei Gruppen unterteilt, welche die beiden Gruppen (hinsichtlich der Zielvariablen) am ehesten trennt. In den zwei entstandenen Gruppen finden sich also jeweils die Fälle, welche sich ähnlich ihrer

Überlebenszeiten verhalten. Die wichtigste Variable steht dabei ganz oben im „Baum“. Dieser Prozess wird solange wiederholt, bis ein Stopp-Kriterium erreicht ist (z.B. Mindestgröße eines Endknoten) oder keine Modellverbesserung mehr erreicht werden kann. Währenddessen findet eine Validierung statt, bei der Daten angewendet werden, um das Modell zu bestätigen (Steyerberg 2009, S. 67; Ellwanger 2001, S. 26). Um ein CART-Modell zu interpretieren, folgt z.B. der behandelnde Arzt dem „Zweig des Baums“, welcher auf den Patienten am besten zutrifft. Je nachdem in welcher Subgruppe/Endknoten der Patient landet, kann die Prognose abgeschätzt werden. Daraus können z.B. Risikogruppen erstellt werden (Steyerberg 2009, S. 329).

1.6 Fragestellung

Es ergeben sich für die vorliegende Arbeit folgende Fragestellungen:

Welchen prognostischen Einfluss haben die im Krankheitsverlauf dokumentierten Variablen auf das Gesamtüberleben?

Wie sinnvoll sind bestehende Prognosesysteme wie TNM/UICC und wie können diese ggf. erweitert werden?

Wie können Patienten anhand ihrer Charakteristika in Risikogruppen hinsichtlich des Gesamtüberlebens eingeteilt werden?

Welche Chancen und Grenzen bieten neue *Machine-Learning*-Methoden (z.B. CART) im Vergleich zu traditionellen Methoden (z.B. Cox-Regression) bei der Überlebenszeitmodellierung?

Welche Möglichkeiten bestehen, komplexe statistische Prognosemodelle in eine klinisch anwendbare Form zu bringen?

2 Material und Methoden

2.1 Retrospektive Datengrundlage

Es standen zwei Datensätze zur Verfügung. Der erste enthielt 325 Speicheldrüsenkarzinompatienten-Fälle (unabhängig von der Lokalisation), die im Universitätsklinikum Gießen behandelt wurden. Der zweite enthielt 913 Parotiskarzinompatienten-Fälle aus 9 deutschen Klinikstandorten [Gießen (GI), Köln (K), Jena (J), Erfurt (EF), Suhl (SHL), Gera (G), Nordhausen (NDH), Bonn (BN) und Ulm (UL)] der Jahre 1953 bis 2014. Aus dem Gießener Tumordokumentationssystem (GTDS) wurde der Datensatz um 36 Fälle aus den Jahren 2014 bis 2018 ergänzt. Zusätzlich wurden noch 65 weitere Fälle des Thüringer Registers von 2014 bis 2018 ergänzt. Doppelte, fehlerhaft oder unvollständig codierte Fälle (z.B. Fehlen elementarer Bestandteile wie Alter oder Überlebenszeit) wurden entfernt. Das GTDS und die anderen Tumorregister erfassen mehr als 200 Variablen je Fall. Folgende Variablen wurden ausgewertet:

Diagnostische Variablen:

Geschlecht, Diagnosealter, *T-Stage*, *N-Stage*, *M-Stage*, *UICC-Stage*, Histologie, Malignität, Rezidiv

Therapeutische Variablen:

Neck-Dissection, Radikale Operation, Therapie vereinfacht, Facialisserhalt vereinfacht

Diese ergeben sich wie folgt:

Die histologischen Entitäten wurden aufgrund des dokumentierten ICD-O-Codes definiert. Bei Patientenfällen, welche an einem Plattenepithelkarzinom der Speicheldrüse erkrankt sind und in deren Vorgeschichte ein Hauttumor dokumentiert wurde, kann von einer möglichen Metastase ausgegangen werden. Diese Fälle wurden aus der Auswertung ausgeschlossen. In den Daten kamen ebenfalls Tumorentitäten vor, welche in keiner der vorliegenden Klassifikationen geführt oder lediglich als „Karzinom“ diagnostiziert wurden. Diese wurden unter „andere Histologien“ zusammengeführt. Abhängig davon, welche WHO-Klassifikation zum Zeitpunkt der pathologischen Befundung galt, wurden die Fälle in „1991 und älter“, „2005“ sowie „2017“ gruppiert. Die Codierung des TNM richtet sich nach dem pathologischen TNM (pTNM). Bei Fällen ohne dokumentierten pTNM wurde der klinische TNM (cTNM) als Grundlage genommen. In der Auswertung wurden *Staging* (UICC), Tumorgroße (T-Stadium), Lymphknotenstatus

(N-Stadium), Fernmetastasierung (M-Stadium) und histologischer Differenzierungsgrad (Malignität) analog zur vorherrschenden Literatur dichotomisiert codiert (*early Stage I/II* vs. *advanced Stage III/IV*, T1/T2 vs. T3/T4, N0 vs. $\geq$ N1, M0 vs. M1, *low-grade* vs. *high-grade*). Diese Dichotomisierung findet statt, da bei einem fortgeschrittenem *Staging*, einer Tumorgroße mit extraparenchymatöser Beteiligung, regionärem Lymphknotenbefall, Fernmetastasierung sowie *high-grade* Tumoren von einer deutlich schlechteren Prognose ausgegangen werden kann (siehe Tabelle 4).

Das Therapieverfahren wurde vereinfacht in „primäre OP“, „adjuvante Therapie“ und „keine OP“ codiert. Als adjuvante Therapie gilt eine zum operativen Eingriff zusätzlich durchgeführte Chemotherapie und/oder Bestrahlung. Bezüglich des chirurgischen Verfahrens wurden zusätzlich die Variablen „radikale OP“ und „*Neck-Dissection*“ geschaffen. Die Entscheidung ob eine OP als „radikal“ gilt oder eine *Neck-Dissection* stattgefunden hat, wurde auf Grund der vergebenen OP-Schlüssel und Freitexte zur Behandlung getroffen. Als „radikal“ wurde eine OP betrachtet, bei der eine totale Parotidektomie, eine *Neck-Dissection* oder ein Eingriff mit Larynx- und Pharynxresektion durchgeführt wurde. Als „Facialisershalt“ wurde eine operative Schonung des Nervus facialis oder dessen Rekonstruktion gewertet. Stattgefundenes *Facialismonitoring* ohne weitere Angaben wurde als „unbekannt“ codiert.

Hinsichtlich des *Outcomes* wird die Gesamtüberlebenszeit betrachtet. Diese ist definiert als der Zeitraum zwischen Diagnosedatum und Sterbedatum bzw. letzter bekannten Info über den vitalen Status des Patienten.

Alle Auswertungen wurden mit SPSS Version 24.0 (*IBM Corporation, Armonk, NY, USA*) und R Version 3.5.0 (*R-Project, R Foundation for Statistical Computing, Vienna, Austria*) durchgeführt. Die Modellbildung und Programmierung der Web-App mit R erfolgten durch die Packages „*survival*“ (Therneau und Lumley 2019), „*rms*“ (Harrell Jr. 2019), „*rpart*“ (Therneau et al.), „*shiny*“ (Chang et al. 2019), „*party*“ (Hothorn et al. 2019a) und „*partykit*“ (Hothorn et al. 2019b).

2.2 Einschlusskriterien

Eingeschlossen wurden alle Fälle mit Diagnosedatum zwischen 1972-2018, bei denen die diagnostischen Variablen T-Stadium, N-Stadium, M-Stadium, Alter und Histologie vollständig dokumentiert wurden. Für die Analysen konnte daher ein Datensatz mit n=688 herangezogen werden.

2.3 Umgang mit fehlenden Werten

Zahlreiche Angaben der Datensätze waren nicht ausreichend dokumentiert. Fehlende Angaben führen dazu, dass der Fall von den statistischen Methoden dieser Arbeit ausgeschlossen und daher nicht in die Modellbildung einbezogen wird. Sofern diese fehlenden Angaben rein zufällig fehlen, führt der Ausschluss der Fälle zu keiner Verzerrung der Ergebnisse. Von der Möglichkeit, fehlende Werte durch statistische Verfahren (*Imputation*) zu rekonstruieren, wurde daher abgesehen. Der Fallausschluss führt naturgemäß zu einer niedrigeren Fallzahl und damit zu einer geringeren Modellgüte, muss jedoch zu Gunsten der Datenauthenzizität akzeptiert werden (Wells et al. 2013). Dies stellt eine anerkannte Vorgehensweise im prädiktiven *Modeling* dar (Masconi et al. 2015).

2.4 Kohorteneinteilung und deskriptive Statistik

Der Datensatz wird durch eine Zufallsstichprobe in eine Trainings- und eine Validierungskohorte eingeteilt. Man spricht dabei von einem *Datasplit*. Die Trainingskohorte besteht aus 70% der Fälle, die Validierungskohorte aus 30%. Diese Aufteilung ist aufgrund der Fallzahl naheliegend. Sie ermöglicht die Zusammenstellung von zahlreichen Zufallsstichproben, aus denen eine Vielzahl von Modellen trainiert und validiert werden kann (Ellwanger 2001, S. 26). Zusätzlich dazu nutzen wir die Möglichkeit des 70%/30% *Datasplit*, um aus 1000 Zufallsstichproben Modelle zu trainieren und damit die Modellstabilität zu beschreiben.

Die Daten wurden deskriptiv analysiert. Die Signifikanz von Unterschieden in der Häufigkeitsverteilung von Variablen wurde mit dem Chi2-Test festgestellt. Zur Überprüfung von Hypothesen wurde ein Signifikanzniveau $\alpha = 5\%$ festgelegt.

2.5 Univariate Überlebenszeitanalyse nach Kaplan-Meier/Cox und Variablenselektion

Es folgte eine univariate Analyse möglicher Einflussfaktoren durch die Kaplan-Meier-Überlebenszeitanalyse/Cox-Regression und den Test der Unterschiede auf statistische Signifikanz mit dem Log-Rank-Test. Die Kaplan-Meier-Überlebenszeitanalyse ist eine nicht-parametrische Methode zur Analyse der Überlebenszeit und kann auch zensierte Daten miteinbeziehen. Zensierung bedeutet, dass das Ereignis innerhalb der

Beobachtungszeit nicht eingetreten ist (z.B. das Ende einer Studie oder Patient, dessen Verbleib unbekannt ist etc.) (Steyerberg 2009, S. 79). Die Variablenselektion für die Modellentwicklung erfolgte anhand der univariaten Überlebenszeitanalyse (Kaplan-Meier-Überlebenszeitanalyse/Cox-Regression) und klinischem Erfahrungswissen.

2.6 Multivariate Überlebenszeitanalyse mit Cox-Regression, Nomogramm und WebApp

Die multivariate Überlebenszeitanalyse erfolgt durch die Cox-Regression. Das Modell wird sowohl intern validiert (Methode: *Bootstrapping*) als auch mit Daten, welche nicht bereits in die Modellentwicklung eingeflossen sind. Dafür werden die 30% der Fälle des *Datasplits* verwendet. Dieser Vorgang wird mehrfach wiederholt und ein Durchschnitt/Median berechnet, welcher zur Bewertung herangezogen wird. Die Diskriminierung wird mit Frank Harrells C-Index quantifiziert. Der C-Index kann dabei zwischen 0 und 1 variieren. Ein Wert von 1 bedeutet dabei perfekte Vorhersagen, ein Wert von 0,5 bedeutet, dass das Modell genau so gut, wie der Zufall vorhersagt. Das Modell hat dann keine Vorhersagekraft. Ein Wert von 0 bedeutet, dass genau das Gegenteil vorhergesagt wird (Hay et al. 2018; Ellwanger 2001, S. 35; Harrell 2015, S. 506). Die Kalibrierung wird mit einem Kalibrierungsdiagramm aufgezeigt. Die Darstellung in eine klinisch anwendbare Form erfolgt als Nomogramm mit der R-Anwendung „*rms*“ (Harrell Jr. 2019) und als *WebApp* mit der R-Anwendung „*shiny*“ (Chang et al. 2019).

2.7 Klassifikations- und Regressionsbäume / Rekursives Partitionieren

Es wird ein Klassifikations- und Regressionsbaum (CART) entwickelt. In dieser Arbeit wird als Zielvariable die Gesamtüberlebenszeit festgelegt. Es existieren mehrere Erweiterungen in R, mit der CART entwickelt werden können. Obwohl sie grundsätzlich nach derselben Methode vorgehen, unterscheiden sich die Ergebnisse je nach verwendetem *Package* erheblich. Das liegt daran, dass unterschiedliche Algorithmen angewandt werden. In dieser Arbeit wird das CART-Modell mit der Funktion „*ctree*“ (Hothorn et al. 2006) aus dem *Package* „*partykit*“ erstellt. „*Ctree*“ verwendet einen Signifikanztest, welcher Verzerrungen vermeidet. Dahingegen neigt der Algorithmus des ebenfalls populären *Packages* „*rpart*“ (Breiman 1984) dazu, Variablen auszuwählen, welche zu vielen *Splits* führen oder viele fehlende Werte aufweisen.

Die Validierung eines CART-Baums wurde gemäß Steyerberg et al. auf zwei Wegen durchgeführt:

- 1:** Mit zwei unterschiedlichen Datensätzen werden zwei CART-Bäume erstellt und die Ähnlichkeit miteinander verglichen. Je ähnlicher sich die beiden „Bäume“ sind, desto eher kann angenommen werden, dass das Modell allgemein angewendet werden kann.
- 2:** Der CART-Baum (entwickelt aus dem Trainingsdatensatz) wird verwendet, um die Fälle des Validierungsdatensatzes in Teilgruppen (gemäß der *terminal nodes*) einzuteilen und das Gesamtüberleben (Kaplan-Meier-Kurven) miteinander zu vergleichen (Steyerberg 2009, S. 376).

3 Ergebnisse

3.1 Deskriptive Analyse von patientenbezogenen Merkmalen

Die Fälle (n=688) haben zahlreiche unterschiedliche Histologien, ihre Häufigkeiten sind in Tabelle 5 dargestellt. Trotz der relativ hohen Fallanzahl kommen manche Entitäten der WHO-Klassifikation nur vereinzelt bzw. gar nicht vor (z.B. Onkozytäres Karzinom, Klarzellkarzinom, Lymphoepitheliales Karzinom, Sialoblastom, Mamma-analoges-sekretorisches Karzinom). Da die weiteren Analysen nur aussagekräftig sind, wenn eine hinreichende Anzahl an Fällen und Ereignissen vorhanden sind, wurden nur die fünf häufigsten Histologien (n=487) einbezogen. Die am häufigsten vertretene Histologie ist das Adenokarzinom (AdenoCa), gefolgt vom Mukoepidermoidkarzinom (MEC), Azinuszellkarzinom (AcinicCa) und dem adenoid-zystischen Karzinom (ACC); den kleinsten Anteil bildet das Plattenepithelkarzinom (PEC).

Tabelle 5: Häufigkeiten der histologischen Entitäten in n=688

WHO-Klassifikation	1991 (Seifert/Sobin) und frühere*		2005 (Barnes et al.)		2017 (El-Naggar et al.)	
	n	(%)	n	(%)	n	(%)
Entitäten	342	100	256	100	90	100
Adenokarzinom	60	17,5	46	18	11	12,2
Mukoepidermoidkarzinom	56	16,4	36	14,1	10	11,1
Azinuszellkarzinom	50	14,6	39	15,2	8	8,9
Adenoid-zystisches Karzinom	53	15,5	27	10,5	12	13,3
Plattenepithelkarzinom	43	12,6	25	9,8	11	12,2
Invasiv-duktales Karzinom	8	2,3	17	6,6	6	6,7
Epithelial-myoepitheliales Karzinom	7	2	13	5,1	4	4,4
Karzinom ex pleomorphes Adenom	12	3,5	8	3,1	2	2,2
Basalzelladenokarzinom	7	2	9	3,5	3	3,3
Myoepitheliales Karzinom	4	1,2	6	2,3	3	3,3
Adenosquamöses Karzinom	0	0	3	1,2	3	3,3
Verhornendes Plattenepithelkarzinom	0	0	0	0	5	5,6
Bösartiger Misch tumor	5	1,5	0	0	0	0
Lymphoepitheliales Karzinom	2	0,6	2	0,8	0	0
Muzinöses papilläres Zystadenokarzinom	2	0,6	0	0	1	1,1
Karzinom sarkom	0	0	0	0	3	3,3
Muzinöses Adenokarzinom	1	0,3	0	0	2	2,2
Neuroendokrines Karzinom, NOS	0	0	2	0,8	0	0
Endometrioides Karzinom	0	0	0	0	1	1,1
Basosquamöses Karzinom	1	0,3	0	0	0	0
Basaloides Plattenepithelkarzinom	0	0	1	0,4	0	0
Neurogenes Sarkom	1	0,3	0	0	0	0
Riesenzellsarkom	1	0,3	0	0	0	0
Malignes fibröses Histiozytom	0	0	1	0,4	0	0
Großzellig-nichtverhornendes Karzinom	0	0	1	0,4	0	0
Hämangiosarkom	1	0,3	0	0	0	0
Undifferenziertes Karzinom	16	4,7	9	3,5	1	1,1
Karzinom	10	2,9	11	4,3	4	4,4
Unbekannt	2	0,6	0	0	0	0

*diese Spalte fasst alle Entitäten ein, welche nach der WHO-Klassifikation von 1991 und allen vorherigen diagnostiziert wurden

Die Fälle mit den fünf häufigsten Histologien (n=487) wurden im Folgenden deskriptiv analysiert. Die Daten entstammen aus neun deutschen Kliniken (Abbildung 1). Den größten Anteil haben Patienten des Universitätsklinikum Köln (Diagnosedatum zwischen 1977-2015), gefolgt vom Thüringer Krebsregister (Diagnosedatum zwischen 1996-2016; Erfurt, Jena, Suhl, Gera, Nordhausen) und dem Universitätsklinikum Gießen (Diagnosedatum zwischen 1972-2018). Den kleinsten Anteil bilden die Patienten aus Bonn (Diagnosedatum 2014) und Ulm (Diagnosedatum 1995-2004), welche unter „Andere“ zusammengefasst wurden.

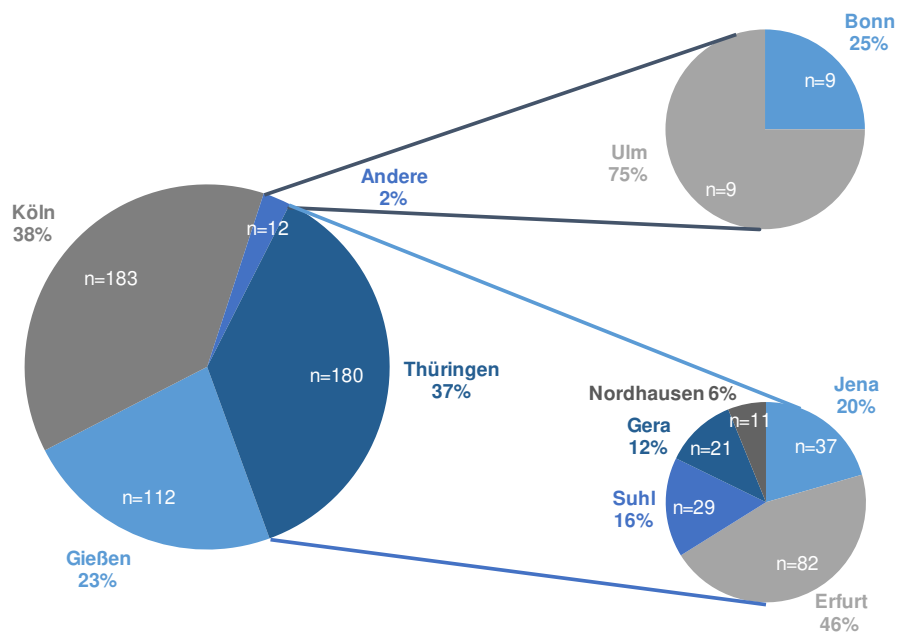


Abbildung 1: Anteile der Klinikstandorte n=487.

Die Histologie des Parotiskarzinoms ist sehr divers. Gemäß aktueller WHO-Klassifikation existieren 20 unterschiedliche maligne Entitäten. In der Modellentwicklung dieser Arbeit wurden lediglich die fünf häufigsten Histologien (n=487) betrachtet, da die übrigen 15 Histologien nicht häufig genug (Fallzahl: min. 0 und max. 24, siehe Tabelle 5) vorkommen, um eine ausreichende Aussagekraft gewinnen zu können. Abbildung 2 stellt die Zusammensetzung der fünf häufigsten Histologien dar.

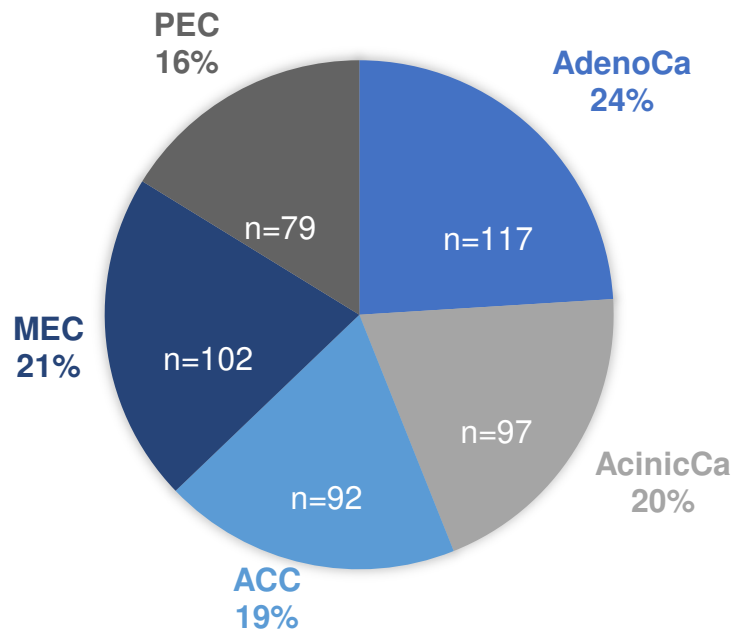


Abbildung 2: Häufigkeitsverteilung der fünf häufigsten Histologien n=487 Adenokarzinom (AdenoCa), Mukoepidermoidkarzinom (MEC), Azinuszellkarzinom (AcinicCa), adenoidzystisches Karzinom (ACC) und Plattenepithelkarzinom (PEC).

Tabelle 6 zeigt die Zusammensetzung des Patientenkollektivs. Es besteht aus 238 Frauen (48,9%) und 249 Männern (51,1%). Das mediane Alter liegt bei 62,8 Jahren (Jüngstes: 7 Jahre, Ältestes: 93 Jahre). Die mediane Nachbeobachtungszeit beträgt 32,9 Monate. Es sind 149 Todesfälle feststellbar (31%).

Tabelle 6: Übersicht über die Gesamtkohorte (n=487).

Gesamtkohorte (n=487)							
		n	(%)				
Gesamt		487	100	Gesamt			
		n	(%)				
Gesamt		487	100	Gesamt			
Standort	Gießen	112	23,0	UICC	I	108	22,2
	Köln	183	37,6		II	97	19,9
	Jena	37	7,6		III	104	21,4
	Erfurt	82	16,8		IV	176	36,1
	Suhl	29	6,0		unbekannt	2	0,4
	Gera	21	4,3	Malignität	Low Grade	174	35,7
	Nordhausen	11	2,3		High Grade	142	29,2
	Bonn	3	0,6		unbekannt	171	35,1
	Ulm	9	1,8	Therapie	adjuvante	251	51,5
Geschlecht	Weiblich	238	48,9		keine OP	18	3,7
	Männlich	249	51,1		primäre OP	191	39,2
Alter (Median:62,8)	<Median	243	49,9		unbekannt	27	5,5
	>Median	244	50,1	Radikale OP	Nein	75	15,4
Histologie	AdenoCa	117	24,0		Ja	338	69,4
	AcinicCa	97	19,9		unbekannt	74	15,2
	ACC	92	18,9	Neck- Dissection	Nein	107	22,0
	MEC	102	20,9		Ja	324	66,5
	PEC	79	16,2		unbekannt	56	11,5
T	T1	130	26,7	Facialiserhalt	Nein	20	4,1
	T2	145	29,8		Ja	124	25,5
	T3	93	19,1		unbekannt	343	70,4
	T4	119	24,4	Rezidiv	Nein	351	72,1
N	N0	306	62,8		Ja	115	23,6
	≥N1	181	37,2		unbekannt	21	4,3
M	M0	442	90,8				
	M1	45	9,2				

3.1.1 Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich des Geschlechts

Die Häufigkeitsverteilung der demographischen Merkmale (von Geschlecht und Alter; Kapitel 3.1.1 und 3.1.2) sowie der Histologie (Kapitel 3.1.3) wurden analysiert. Die Häufigkeitsverteilung für alle anderen Variablen sind im Anhang zu finden. Mit der Kreuztabelle (Tabelle 7) wird der Zusammenhang von Tumorcharakteristika und dem Geschlecht dargestellt. Dies liefert Hinweise darauf, ob männliches bzw. weibliches Geschlecht als unabhängiger Risikofaktor angesehen werden kann oder ob sich die Unterschiede in den Überlebenszeiten der männlichen bzw. weiblichen Ausprägung durch die Assoziation mit anderen Ausprägungen erklären lassen.

Tabelle 7: Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich des Geschlechts

Geschlecht		W		M		p-Wert
		n	(%)	n	(%)	
Alter (Median:62,8)	<Median	131	55,0	112	45,0	0,026
	>Median	107	45,0	137	55,0	
T	T1/2	150	63,0	125	50,2	0,004
	T3/4	88	37,0	124	49,8	
N	N0	170	71,4	136	54,6	<0,001
	≥N1	68	28,6	113	45,4	
M	M0	223	93,7	219	88,0	0,04
	M1	15	6,3	30	12,0	
UICC	II	119	50,2	86	34,7	0,001
	III/IV	118	49,8	162	65,3	
Malignität	Low Grade	88	60,3	86	50,6	0,067
	High Grade	58	39,7	84	49,4	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: statistisch signifikante Werte ($p \leq 0,05$)

In unserer Kohorte sind weibliche Patienten im Vergleich zu männlichen jünger (55% unterhalb des Medians vs. 45%). Sie haben ein niedrigeres TNM- (T1/2: 63% vs. 50,2%, N0: 71,4% vs. 54,6%, M0: 93,7% vs. 88,0%) und UICC-*Staging* (50,2% vs. 34,7%). Die Häufigkeitsverteilung hinsichtlich des Geschlechts ist für die Variablen Alter, TNM und UICC-*Staging* statistisch signifikant unterschiedlich.

3.1.2 Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich der Altersgruppe

Die Kreuztabelle (Tabelle 8) stellt den Zusammenhang von Tumorcharakteristika und dem Alter dar. Als Grenze zwischen jüngerer und älterer Gruppe wurde das mediane Alter zum Zeitpunkt der Diagnose festgelegt. Diese Auswertung gibt uns Hinweise darauf, ob das Alter als eigener Risikofaktor angesehen werden kann oder ob sich die Unterschiede in den Überlebenszeiten durch die Assoziation des Alters mit anderen Ausprägungen erklären lassen.

Tabelle 8: Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich der Altersgruppe

Alter		<Median		>Median		p-Wert
		n	(%)	n	(%)	
Geschlecht	W	131	53,9	107	43,9	0,026
	M	112	46,1	137	56,1	
T-Stage	T1/2	147	60,5	128	52,5	0,074
	T3/4	96	39,5	116	47,5	
N-Stage	N0	171	70,4	135	55,3	0,001
	≥N1	72	29,6	109	44,7	
M-Stage	M0	223	91,8	219	89,8	0,443
	M1	20	8,2	25	10,2	
UICC	II	118	48,6	87	36,0	0,005
	III/IV	125	51,4	155	64,0	
Malignität	Low Grade	92	60,1	82	50,3	0,079
	High Grade	61	39,9	81	49,7	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: statistisch signifikante Werte ($p \leq 0,05$), medianes Alter 62,8 Jahre

Die Patienten der jüngeren Gruppe sind im Gegensatz zur älteren Gruppe eher weiblich (53,9% vs. 43,9%). Sie haben seltener regionäre Lymphknotenmetastasen (70,4% vs. 55,3%) und ein niedrigeres UICC-*Staging* (48,6% vs. 36%). Statistisch signifikante Unterschiede in der Häufigkeitsverteilung ergaben sich für die Variablen Geschlecht, N-Stadium und UICC-*Staging*.

3.1.3 Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich der Histologie

In der Literatur ist bekannt, dass es Unterschiede in der Überlebenszeit abhängig von der Histologie gibt. Es ist daher von besonderem Interesse, die Verteilung der patienten- und tumorbezogenen Merkmale auf die einzelnen histologischen Typen des Tumors auszuwerten (siehe Tabelle 9).

Tabelle 9: Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich der Histologie

		Alle				AdenoCa				AcinicCa				ACC				MEC				PEC			
				Nein	Ja	p-Wert			Nein	Ja	p-Wert			Nein	Ja	p-Wert			Nein	Ja	p-Wert			Nein	Ja
		%		%			%		%			%		%			%		%			%		%	
Geschlecht	W	48,9	53,8	33,3	<0,001	44,1	68,0	<0,001	0,063	46,8	57,6	0,251	0,001	52,2	31,6	<0,001	0,001	48,9	53,8	33,3	<0,001	0,001	52,2	31,6	<0,001
	M	51,1	46,2	66,7		55,9	32,0			53,2	42,4			47,8	68,4										
Alter	<Median	49,9	53,5	38,5	0,005	48,7	54,6	0,297	0,01	47,1	62,0	<0,001	0,001	55,9	19	<0,001	0,001	49,9	53,5	38,5	0,005	0,001	55,9	19	<0,001
	>Median	50,1	46,5	61,5		51,3	45,4			52,9	38,0			44,1	81										
T-Stage	T1/2	56,5	60	45,3	0,005	53,8	67,0	0,019	0,02	59	45,7	<0,001	0,001	57,8	49,4	0,164	0,001	56,5	60	45,3	0,005	0,001	57,8	49,4	0,164
	T3/4	43,5	40	54,7		46,2	33,0			41	54,3			42,2	50,6										
N-Stage	N0	62,8	67,8	47	<0,001	60,3	73,2	0,018	0,002	59,5	77,2	0,04	0,001	66,2	45,6	<0,001	0,001	62,8	67,8	47	<0,001	0,001	66,2	45,6	<0,001
	≥N1	37,2	32,2	53		39,7	26,8			40,5	22,8			33,8	54,4										
M-Stage	M0	90,8	92,2	86,3	0,057	89,5	95,9	0,052	0,841	90,6	91,3	0,545	0,001	90,4	92,4	0,581	0,001	90,8	92,2	86,3	0,057	0,001	90,4	92,4	0,581
	M1	9,2	7,8	13,7		10,5	4,1			9,4	8,7			9,6	7,6										
UICC	I/II	42,3	45,8	31	0,005	38,9	55,7	0,003	0,362	43,3	38,0	<0,001	0,001	45,7	24,4	<0,001	0,001	42,3	45,8	31	0,005	0,001	45,7	24,4	<0,001
	III/IV	57,7	54,2	69		61,1	44,3			56,7	62,0			54,3	75,6										
Malignität	Low Grade	55,1	59	44,9	0,024	52,9	66,0	0,078	0,280	56,3	47,9	0,002	0,002	57,0	45,1	0,118	0,002	55,1	59	44,9	0,024	0,002	57,0	45,1	0,118
	High Grade	44,9	41	55,1		47,1	34,0			43,7	52,1			43,0	54,9										

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: statistisch signifikante Werte ($p \leq 0,05$); rot/grüne Rahmen: Kennzeichnung der Tendenz der Häufigkeitsverteilung bezogen auf die jeweilige Histologie im Vergleich zur Gesamtheit aller anderen Histologien.

Aufgrund der farblichen Kennzeichnung der signifikanten Unterschiede in rot/grün in Tabelle 9 erkennt man, dass das Adeno- und Plattenepithelkarzinom ausschließlich signifikante Unterschiede in der Häufigkeitsverteilung zugunsten von Ausprägungen haben, denen generell eine schlechtere Prognose zugerechnet wird (siehe rote Kästen). Azinuszellkarzinom und Mukoepidermoidkarzinom hingegen sind statistisch signifikant mehrheitlich eher mit günstigeren Ausprägungen assoziiert (siehe grüne Kästen).

Männer sind statistisch signifikant eher unter den Adenokarzinompatienten (66,7%) und bei Plattenepithelkarzinompatienten (68,4%) zu finden; Frauen hingegen eher unter den Patienten mit Azinuszellkarzinom (68%).

Angehörige der älteren Gruppe sind eher unter den Patienten des Adenokarzinoms (61,5%) und des Plattenepithelkarzinoms (81%) zu finden; dahingegen sind Angehörige der jüngeren Gruppe eher unter den Patienten anzutreffen, welche am Azinuszellkarzinom, dem adenoidzystischen Karzinom oder dem Mukoepidermoidzellkarzinom erkrankt sind. Diese Unterschiede in der Häufigkeitsverteilung sind für alle Histologien statistisch signifikant. Kleine Tumoren sind eher unter den Patienten des Azinuszellkarzinoms (67%) und des Mukoepidermoidzellkarzinoms (74,5%) zu finden. Große Tumoren sind eher bei den Patienten des Adenokarzinoms (54,7%), des adenoidzystischen Karzinoms (54,3%) und des Plattenepithelkarzinoms (50,6%) vertreten. Die Unterschiede sind für alle Histologien außer dem Plattenepithelkarzinom statistisch signifikant.

Lymphknotenbefall ist eher unter den Patienten des Adenokarzinoms (53%) und des Plattenepithelkarzinoms (54,4%) vertreten. Ein negativer Lymphknotenstatus hindessen

ist eher bei den Patienten des Azinuszellkarzinoms (73,2%), des adenoidzystischen Karzinoms (77,2%) und des Mukoepidermoidkarzinoms (71,6%) vertreten. Die Unterschiede sind für alle Histologien statistisch signifikant. Die Patienten unserer Kohorte (n=487) weisen unabhängig von der Histologie relativ selten eine Fernmetastasierung auf. Die Unterschiede in der Häufigkeitsverteilung sind für keine der fünf häufigsten Histologien statistisch signifikant. Patienten mit vorangeschrittenem UICC-Stadium (III/IV) sind eher unter den Patienten des Adenokarzinoms (69%), des adenoidzystischen Karzinoms (62%) und des Plattenepithelkarzinoms (75,6%) vertreten. Patienten mit einem niedrigeren UICC-Stadium (I/II) sind eher unter den Patienten des Azinuszellkarzinoms (55,7%) und des Mukoepidermoidkarzinoms (59,8%) anzutreffen. Die Unterschiede sind für alle Histologien außer dem adenoidzystischen Karzinom statistisch signifikant. Da sich das UICC-Stadium direkt aus dem TNM-Status ergibt, ist diese Häufigkeitsverteilung nur konsequent und stellt keine neue Erkenntnis dar. Patienten mit einem *low-grade* Tumor sind eher bei den Patienten zu finden, welche an einem Azinuszellkarzinom (66%) oder einem Mukoepidermoidkarzinom (70,7%) erkrankt sind. *High-grade* Tumoren sind eher unter den Fällen mit einem Adenokarzinom (55,1%), einem adenoidzystischen Karzinom (52,1%) oder einem Plattenepithelkarzinom (54,9%) anzutreffen. Eine statistische Signifikanz ergibt sich lediglich für das Adenokarzinom, das Mukoepidermoidkarzinom und das Plattenepithelkarzinom.

3.2 Univariate Analyse der Einflussfaktoren auf die Überlebenszeit

Die patientenspezifischen (Standort, Geschlecht, Alter), diagnostischen (TNM, Histo, Malignität, Rezidiv) und therapeutischen Variablen (Therapie, Radikale OP, *Neck-Dissection*, Facialiserhalt) wurden auf ihren Einfluss auf das Gesamtüberleben analysiert, um mögliche Prädiktoren zu identifizieren.

In Abbildung 3 wurde das Gesamtüberleben der Kohorte analysiert.

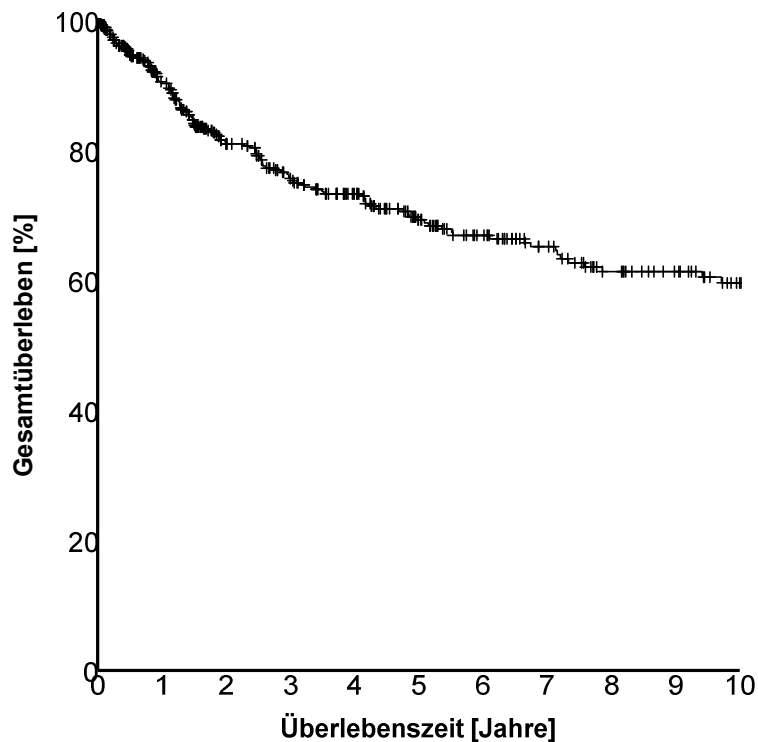


Abbildung 3: Kaplan-Meier-Überlebenszeitanalyse für alle Fälle (n=487). Zensierte Fälle durch senkrechte Striche markiert.

Das 5-Jahresüberleben in der Gesamtkohorte beträgt 69,3%. Es sind 149 Ereignisse(=Todesfälle) eingetreten. Dementsprechend sind 338 Fälle zensiert.

Das mediane Überleben war 13,7 Jahre (95% CI: 11,4 – 15,6 Jahre). Die mediane Nachbeobachtungszeit beträgt 2,7 Jahre, die mittlere Nachbeobachtungszeit 4,6 Jahre.

3.2.1 Standort

In der univariaten Kaplan-Meier-Überlebenszeitanalyse zeigen sich statistisch signifikante Unterschiede abhängig vom Klinikstandort (siehe Abbildung 4, Tabelle 10).

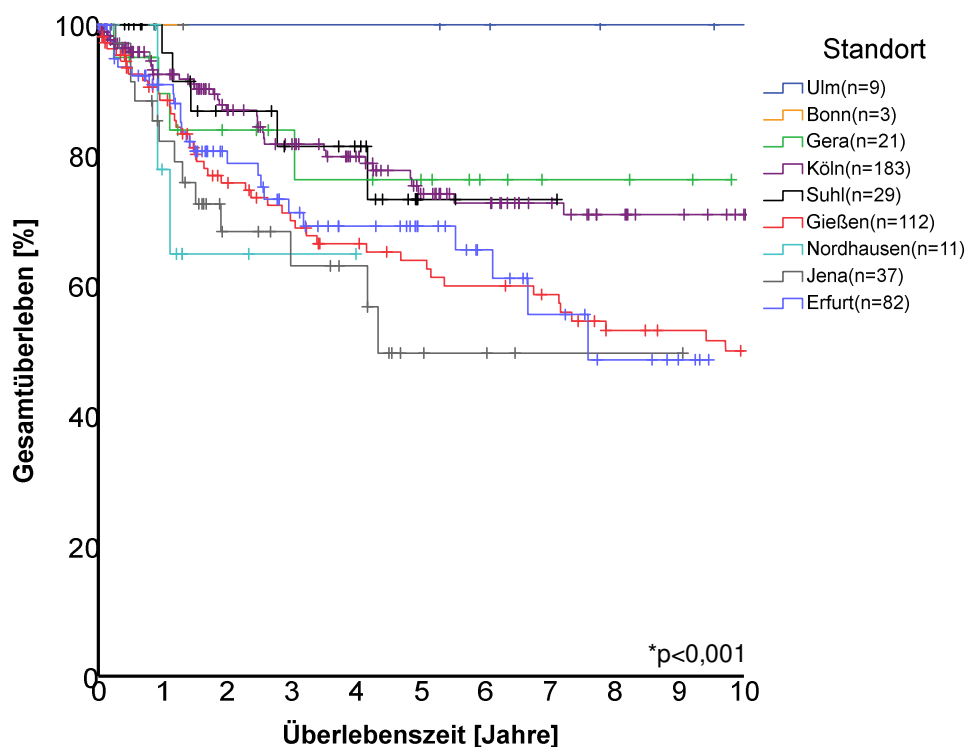


Abbildung 4: Kaplan-Meier-Überlebenszeitanalyse differenziert nach Standort, n=487 (*p: Log-rank-Test über alle Schichten; paarweise Vergleiche in Tabelle 10), zensierte Fälle durch senkrechte Striche markiert.

Tabelle 10: P-Werte (Log-Rank-Test; paarweise Vergleiche der Kaplan-Meier-Kurven in Abbildung 4) differenziert nach Standort, Grün: Statistisch signifikant(p<0,05), *: Ohne p-Wert, da alle Fälle zensiert

	n	BN	EF	G	GI	J	K	NDH	SHL	UL
BN	3		0,542	0,603	0,536	0,443	0,620	0,526	0,766	*
EF	82	0,542		0,338	0,859	0,276	0,051	0,322	0,366	0,019
G	21	0,603	0,338		0,245	0,160	0,931	0,370	0,908	0,126
GI	112	0,536	0,859	0,245		0,383	0,003	0,426	0,233	0,007
J	37	0,443	0,276	0,160	0,383		0,008	0,748	0,091	0,015
K	183	0,620	0,051	0,931	0,003	0,008		0,077	0,963	0,052
NDH	11	0,526	0,322	0,370	0,426	0,748	0,077		0,091	0,058
SHL	29	0,766	0,366	0,908	0,233	0,091	0,963	0,091		0,102
UL	9	*	0,019	0,126	0,007	0,015	0,052	0,058	0,102	

Für den Standort Ulm(n=9) lassen sich die statistisch signifikanten Unterschiede hinsichtlich des Überlebens aufgrund der fehlenden Ereignisse (kein einziger Todesfall vorhanden) erklären. Es lassen sich auch bei den Standorten mit hoher Fallzahl statistisch signifikante Unterschiede (Köln im Vergleich zu Gießen: p=0,003; Köln im Vergleich zu Jena: p=0,008) feststellen; aufgrund dessen ist eine Berücksichtigung des Standorts bei der Einteilung in Trainings- und Validierungskohorte sinnvoll.

3.2.2 Geschlecht und Alter

Es wurde der Einfluss der patientenbezogenen Variablen wie Geschlecht und Alter auf das Gesamtüberleben analysiert (siehe Abbildung 5).

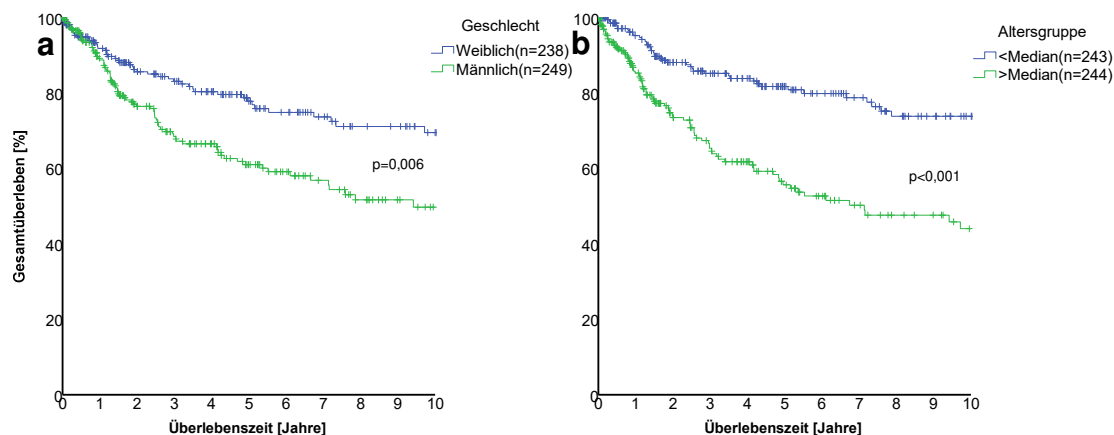


Abbildung 5: Kaplan-Meier-Überlebenszeitanalyse (p: Log-rank-Test; zensierte Fälle durch senkrechte Striche markiert): a) Geschlecht (n=487), b) Altersgruppe (n=487)

Mit 77,9% haben Patientinnen ein signifikant ($p=0,006$) besseres 5-Jahres-Gesamtüberleben als Patienten (61,1%); siehe Abbildung 5a. Das mediane Überleben für Frauen beträgt 13,9 Jahre, für Männer 9,4 Jahre. Das mediane Alter zum Diagnosezeitpunkt beträgt 62,8 Jahre. Der jüngste Patient war 7 Jahre alt, der älteste 95 Jahre. Es zeigt sich, dass die Zugehörigkeit zur älteren Gruppe mit einem geringeren Überleben assoziiert ist (5-Jahres-Überlebensrate 55,3%) als die Zugehörigkeit zur jüngeren Gruppe (5-Jahres-Überlebensrate 80,5%). Die Unterschiede in den Überlebenszeiten sind mit $p<0,001$ statistisch signifikant (siehe Abbildung 5b).

3.2.3 Histologie

Von den bereits genannten 688 Fällen entfallen 487 auf die fünf häufigsten Histologien, 201 Fälle sind von einer anderen Histologie betroffen. Dies sind oftmals Histologien, welche sehr selten vorkommen und deren Überlebensverhalten daher nur unzureichend

untersucht sind. Der Log-Rank-Test zeigt einen statistisch signifikanten Unterschied in den Überlebenszeiten über alle fünf Histologien (Abbildung 6, Tabelle 11).

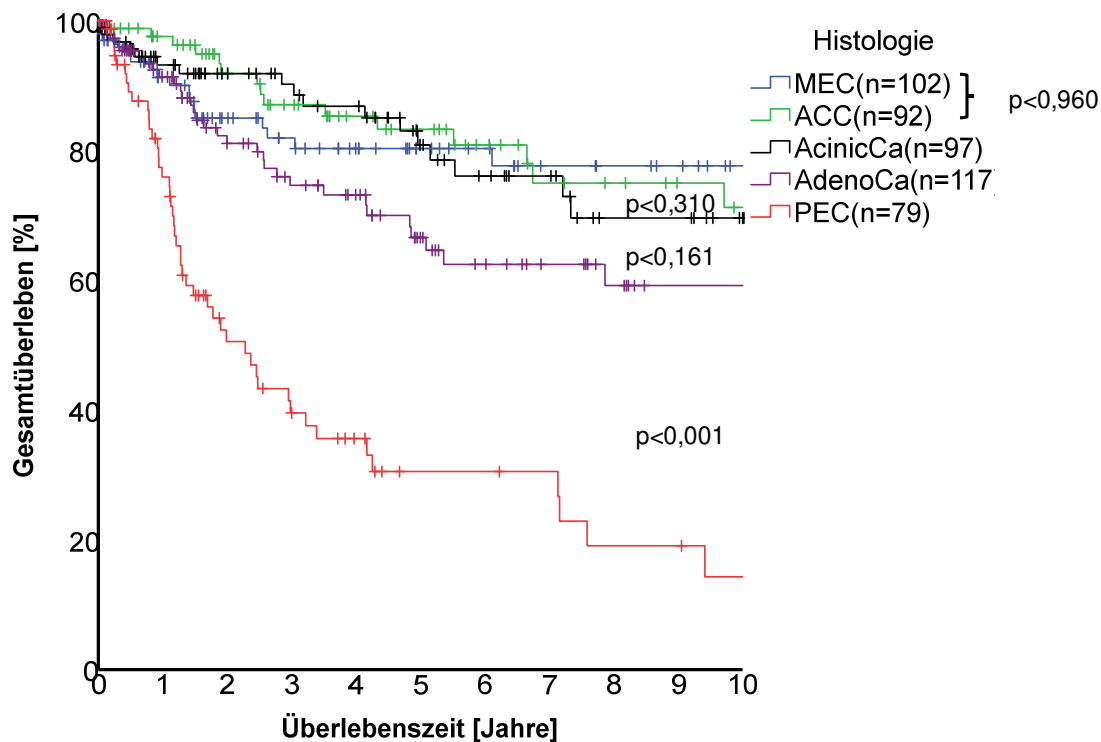


Abbildung 6: Kaplan-Meier-Überlebenszeitanalyse differenziert nach Histologie n=487 (p: Log-rank-Test jeweils zwischen benachbarten Gruppen wie dargestellt), zensierte Fälle durch senkrechte Striche markiert.

Das Plattenepithelkarzinom ist in unserer Kohorte mit dem geringsten Überleben assoziiert (5-Jahres-Überlebensrate 30,5%). Darauf folgen das Adenokarzinom (5-Jahres-Überlebensrate 66,6%), das Mukoepidermoidkarzinom (5-Jahres-Überlebensrate 80,3%), das Azinuszellkarzinom (5-Jahres-Überlebensrate 80,9%) und das Adenoid-zystische Karzinom (5-Jahres-Überlebensrate 83,3%).

Tabelle 11: P-Werte (Log-Rank-Test; paarweise Vergleiche der Kaplan-Meier-Kurven in Abbildung 6) differenziert nach häufigsten Histologien, Grün: Statistisch signifikant ($p < 0,05$)

	AdenoCa	AcinicCa	ACC	MEC	PEC
AdenoCa		0,161	0,012	0,020	<0,001
AcinicCa	0,161		0,310	0,322	<0,001
ACC	0,012	0,310		0,960	<0,001
MEC	0,020	0,322	0,960		<0,001
PEC	<0,001	<0,001	<0,001	<0,001	

Das Gesamtüberleben aller anderen Histologien mit $n > 10$ (Basallzelladenokarzinom (BCAC): 19, epithelial-myoepitheliales Karzinom (EMC): 24, invasiv-duktales Karzinom (IDC): 31, Karzinom ex pleomorphes Adenom (CaexPA): 22, myoepitheliales Karzinom

(MC): 13, wurde ebenfalls mittels der Kaplan-Meier-Überlebenszeitanalyse untersucht und ist in Abbildung 7 dargestellt.

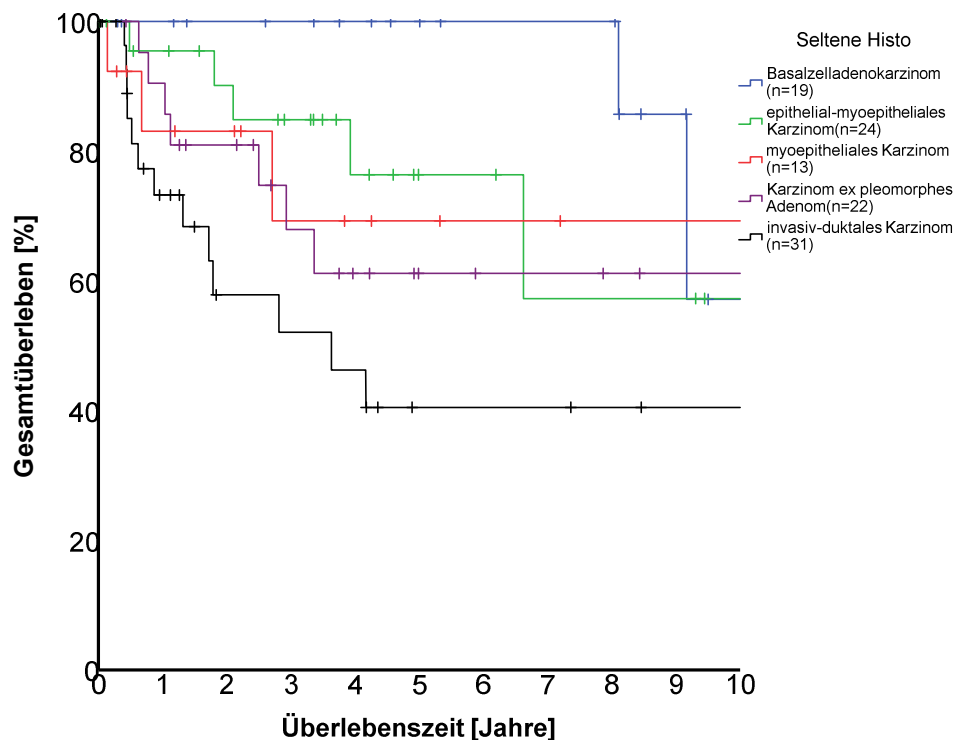


Abbildung 7: Kaplan-Meier-Überlebenszeitanalyse differenziert nach seltener Histologie n=109

Die 5-Jahres-Überlebensrate unterscheidet sich je nach seltener Histologie: (BCAC: 100%, EMC: 76,4%, IDC: 40,5%, CaexPA: 61,1%, MC: 69,2%). Statistisch signifikante Unterschiede in den Überlebenszeiten der seltenen Histologien untereinander weisen die Entitäten BCAC, EMC, IDC und CaexPa auf (Tabelle 12).

Tabelle 12: P-Werte (Log-Rank-Test; paarweise Vergleiche der Kaplan-Meier-Kurven in Abbildung 7 differenziert nach seltener Histologie, Grün: Statistisch signifikant ($p < 0,05$))

	BCAC	EMC	IDC	CaexPa	MC
BCAC		0,171	0,003	0,035	0,118
EMC	0,171		0,030	0,422	0,724
IDC	0,003	0,030		0,159	0,219
CaexPA	0,035	0,422	0,159		0,836
MC	0,118	0,724	0,219	0,836	

3.2.4 TNM- und UICC-Stadien

Informationen über Größe und Ausbreitungstendenz des Tumorgeschehens werden international mittels der TNM-Klassifikation diagnostiziert. Das UICC-*Staging* stellt eine Vereinfachung des TNM dar und Einstufung des Patienten in Risikogruppen (siehe 1.2). In der Literatur ist bekannt, dass ein steigendes TNM-/UICC-*Staging* mit einem

geringeren Überleben assoziiert ist. Mit den folgenden Kaplan-Meier-Überlebenszeitanalysen wird diese These überprüft (siehe Abbildung 8).

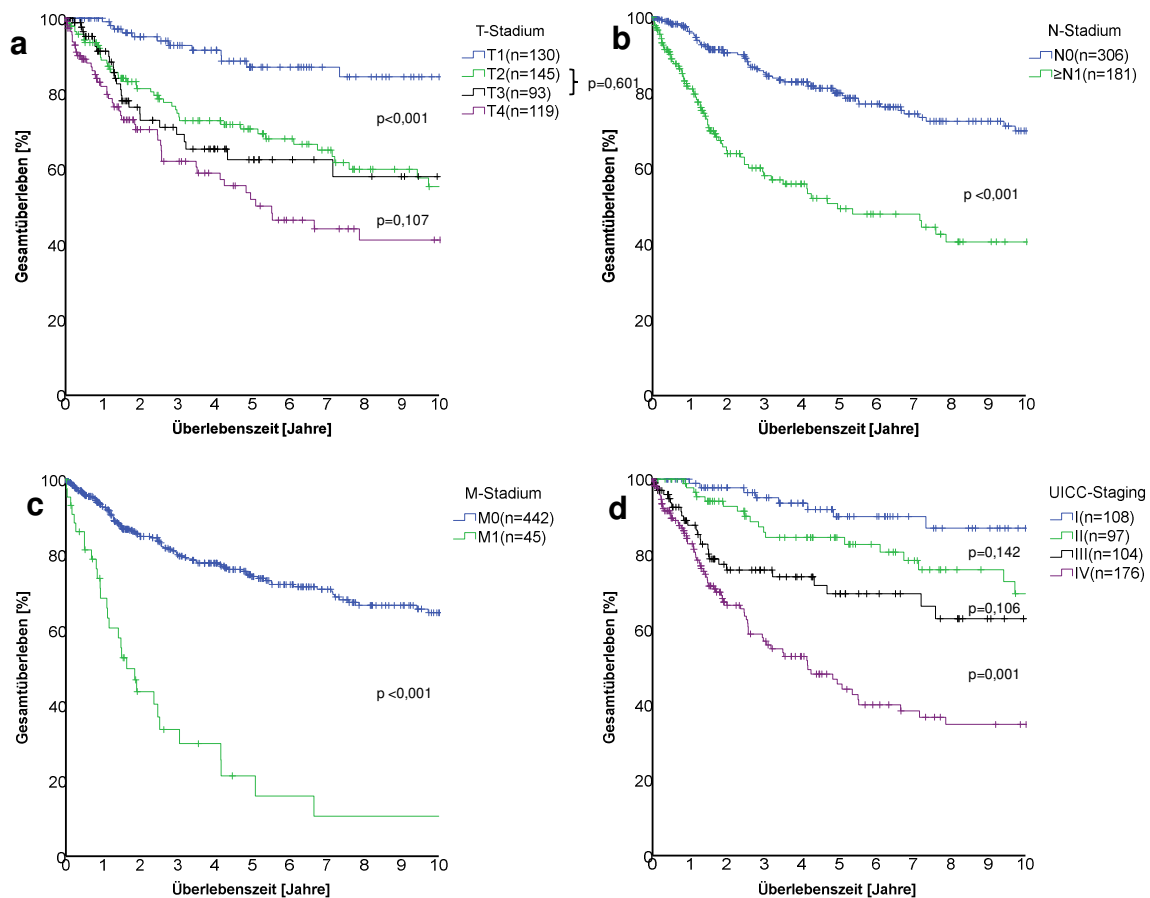


Abbildung 8: Kaplan-Meier-Überlebenszeitanalyse (p: Log-rank-Test; zensierte Fälle durch senkrechte Striche markiert): a) T-Stadium (n=487), b) N-Stadium (n=487), c) M-Stadium (n=487), d) UICC-Staging (n=485)

Ein höheres T-Stadium ist mit einem geringeren Überleben assoziiert (siehe Abbildung 8a). Das T1-Stadium hat eine 5-Jahres-Überlebensrate von 87 %, das T2-Stadium (5-Jahres-Überlebensrate 70,6%), das T3-Stadium (5-Jahres-Überlebensrate 62,5%), das T4-Stadium (5-Jahres-Überlebensrate 52%).

Betrachtet man den regionären Lymphknotenstatus (Abbildung 8b), so lässt sich ein signifikant geringeres Überleben bei positivem Lymphknotenstatus ($\geq N1$, 5-Jahres-Überlebensrate 49,4 %) im Vergleich zu einem negativen Lymphknotenstatus (N0, 5-Jahres-Überlebensrate 74,4 %) feststellen. Der Unterschied ist statistisch signifikant ($p<0,001$).

Patienten ohne Fernmetastasen haben ein signifikant besseres Überleben ($p < 0,001$, 5-Jahres-Überlebensrate 74,3 %), als Patienten mit Fernmetastasen (5-Jahres-Überlebensrate 21,4 %); siehe Abbildung 8c.

Patienten mit dokumentierten UICC-Stadien weisen je nach Stadium unterschiedliche Überlebenszeiten auf (siehe Abbildung 8d). Das UICC-Stadium I hat eine 5-Jahres-Überlebensrate von 90,1%. Für UICC-Stadium II beträgt sie 84,5 %. Für UICC-Stadium III und IV 69,6% bzw. 45,6 %. Die Unterschiede sind über alle Schichten statistisch signifikant ($p < 0,001$).

3.2.5 Malignität

Malignität beschreibt den Differenzierungsgrad der untersuchten Zellen und wird durch pathohistologisch/zytologische Untersuchungen bestimmt. Welchen Einfluss dieser Differenzierungsgrad (*low-grade* = gut differenziert, hohe Übereinstimmung mit dem Ursprungsgewebe, *high-grade* = schlecht differenziert, geringe Übereinstimmung mit dem Ursprungsgewebe) auf das Gesamtüberleben hat, wird in Abbildung 9 dargestellt. Daten zur Malignität sind oftmals unvollständig dokumentiert gewesen (171 fehlende Angaben).

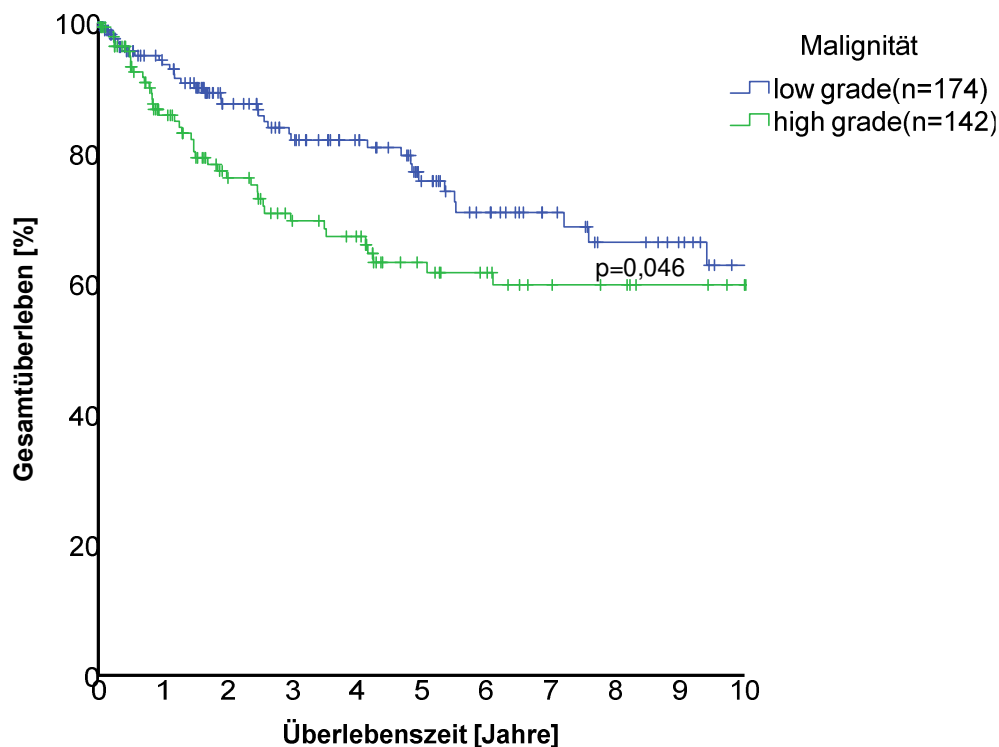


Abbildung 9: Kaplan-Meier-Überlebenszeitanalyse differenziert nach Malignität $n=316$ (p : Log-rank-Test), zensierte Fälle durch senkrechte Striche markiert.

Patienten mit einem *low-grade* Tumor haben ein signifikant besseres Überleben (5-Jahres-Überlebensrate 75,7 %), als Patienten mit einem *high-grade* Tumor (5-Jahres-Überlebensrate 63,3 %).

3.2.6 Therapieverfahren/radikale OP/Neck-Dissection/Rezidiv

Patienten in unserer Kohorte wurden unterschiedlich therapiert (siehe Abbildung 10a). 191 Patienten erhielten eine primär chirurgische Therapie (5-Jahres-Überlebensrate 87,9%), 251 Patienten eine adjuvante Therapie (5-Jahres-Überlebensrate 58,7%), 18 Patienten erhielten keine chirurgische Therapie, 27 erhielten eine andere/unbekannte Therapie (unzureichend dokumentiert; 5-Jahres-Überlebensrate 25,3%). Patienten mit einer primär chirurgischen Therapie haben eine günstigere 5-Jahres-Überlebensrate als Patienten mit einer anderen Therapie.

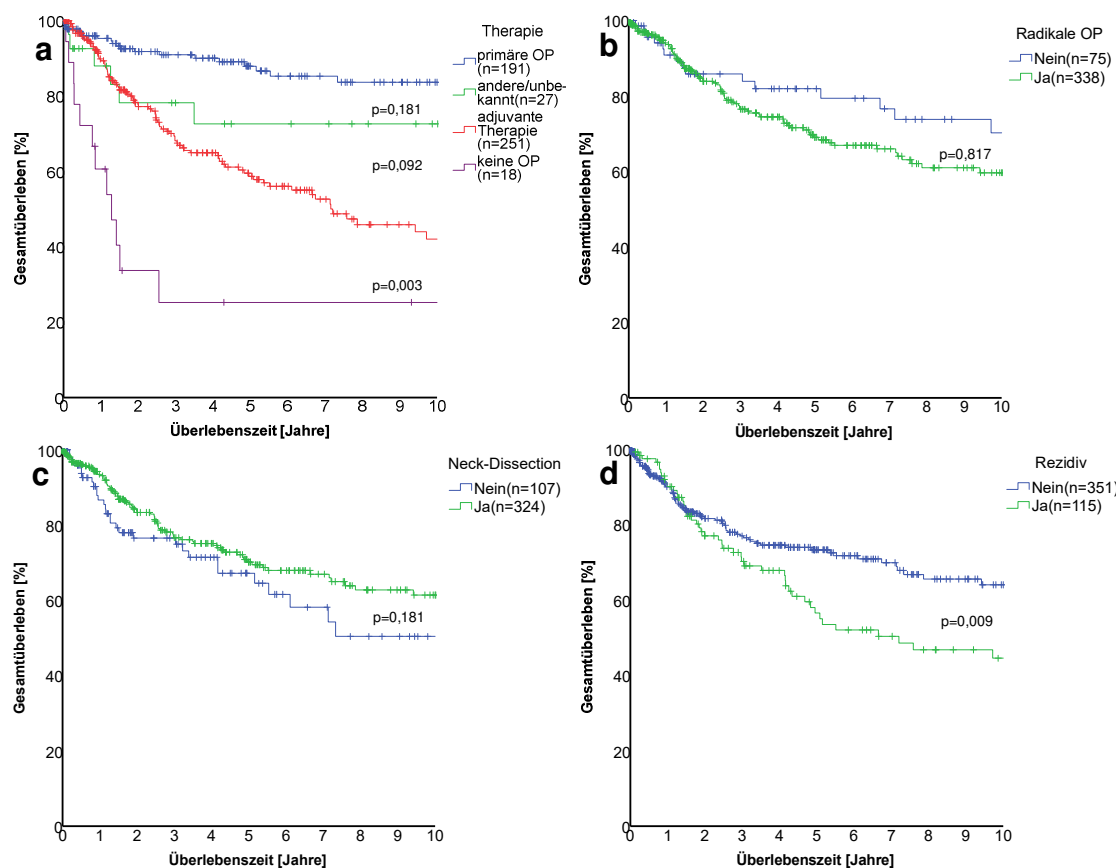


Abbildung 10: Kaplan-Meier-Überlebenszeitanalyse (p: Log-rank-Test, zensierte Fälle durch senkrechte Striche markiert): a) Therapie (n=487), b) radikales OP-Vorgehen (n=413), c) Neck-Dissection (n=431), d) Rezidiv (n=466)

75 Patienten wurden nicht radikal operiert, 338 Patienten erhielten eine radikale Operation. Die Überlebenszeiten der beiden Gruppen sind mit $p = 0,817$ nicht statistisch signifikant unterschiedlich (Abbildung 10b).

Eine *Neck-Dissection* wurde bei 324 Patienten durchgeführt. 107 Patienten erhielten keine *Neck-Dissection*. Die univariate Kaplan-Meier-Überlebenszeitanalyse legt nahe, dass die Überlebenszeiten von Patienten mit *Neck-Dissection* besser sind als ohne. Die Unterschiede sind jedoch nicht statistisch signifikant (Abbildung 10c).

Unter einem Rezidiv versteht man das erneute Auftreten einer Tumorerkrankung nach deren (mutmaßlich) erfolgreicher Therapie. Welchen Einfluss das Wiederkehren der malignen Erkrankung auf das Gesamtüberleben hat, wird in den Kaplan-Meier-Kurven von Abbildung 10d gezeigt.

Bei 115 Patienten (23,6%) konnte während des Beobachtungszeitraums ein Rezidiv festgestellt werden. Das Vorhandensein eines Rezidivs ist mit einem statistisch signifikant geringeren Überleben assoziiert ($p = 0,009$). Die 5-Jahres-Überlebensrate von Patienten mit Rezidiv beträgt 56,5% (ohne Rezidiv: 73,2%).

Auf Grundlage der Kaplan-Meier-Überlebenszeitanalyse und der univariaten Cox-Regression wurde eine Risikofaktortabelle (siehe Tabelle 13) erstellt, welche eine Übersicht über den Einfluss der Prädiktoren darstellt.

Tabelle 13: Univariate Analyse von Einflussfaktoren auf das Gesamtüberleben von Patienten der fünf häufigsten Parotiskarzinome (n=487)

Univariate Analyse (Kaplan-Meier) von Risikofaktoren (n=487)						Univariate Cox-Regression		
			Medianes OS		5-Jahres-OS (%)	P(log-rank)	HR	95% CI (unteres/oberes)
		n	Jahre	S.E.				
Gesamt		487	13,651	1,126	69,3			
Geschlecht	Weiblich	238	13,941	0,941	77,9	0,006	1	
	Männlich	249	9,410	2,837	61,1		1,579	1,138-2,192
Alter (Median:62,8)	<Median	243	20,693	2,344	81,8	<0,001	1	
	>Median	244	5,153	1,866	55,7		3,145	2,316-4,659
Histologie	AdenoCa	117	11,409	2,717	66,6	<0,001	1	
	AcinicCa	97	13,755	2,988	80,9		0,689	0,413-1,149
	ACC	92	20,693	4,689	83,3		0,502	0,289-0,871
	MEC	102	/	/	80,3		0,502	0,286-0,880
	PEC	79	2,281	0,389	30,5		3,308	2,149-5,094
T	T1	130	16,706	2,061	87	<0,001*	1	
	T2	145	11,409	2,143	70,6		2,364	1,421-3,933
	T3	93	10,223	2,010	62,5		2,670	1,504-4,741
	T4	119	5,514	1,247	52		3,951	2,367-6,594
N	N0	306	15,956	1,274	80	<0,001	1	
	≥N1	181	4,956	1,556	49,4		2,756	1,991-3,814
M	M0	442	13,941	1,320	74,3	<0,001	1	
	M1	45	1,634	0,292	21,4		4,528	3,024-6,778
UICC-Stage	I	108	16,706	2,870	90,1	<0,001*	1	
	II	97	16,000	1,496	84,5		1,574	0,836-2,965
	III	104	11,231		69,6		2,401	1,292-4,462
	IV	176	4,170	0,812	45,6		5,043	2,942-8,643
Malignität	Low Grade	174	13,941	1,843	75,7	0,046	1	
	High Grade	142	12,873	1,611	63,3		1,547	1,003-2,384

OS=Gesamtüberleben, S.E. = Standardfehler, HR=Hazard Ratio, CI= Konfidenzintervall, *P=Log-Rank-Test über alle Schichten, fett: p<0,05

3.3 Multivariate Cox-Regression und Modellbildung

Trainings- und Validierungskohorten wurden jeweils durch eine Zufallsstichprobe (70%/30%) aus der Gesamtkohorte n=487 gebildet. Man spricht dabei von einem *Datasplit*. In 3.2.1 wurde nachgewiesen, dass es statistisch signifikante Unterschiede in den Überlebenszeiten zwischen den Standorten gibt. Daher wurde die Zufallsstichprobe so gesteuert, dass in Trainings- und Validierungskohorte die Standorte zu gleichen Teilen repräsentiert sind. Die Trainingskohorte (n=340) wird verwendet, um das Modell zu entwickeln; die Validierungskohorte (n=147) um die Validität des Modells zu prüfen. Die Zusammensetzung der beiden Kohorten ist in Tabelle 14 dargestellt; der Chi2-Test

zeigt, dass es keine statistisch signifikanten Unterschiede in der Häufigkeitsverteilung gibt.

Tabelle 14: Übersicht Trainings- und Validierungskohorte (n=487)

Übersicht Trainings- und Validierungskohorte (n=487)						
		Training		Validierung		P-Wert
		n	(%)	n	(%)	
Gesamt		340	100	147	100	
Standort	Gießen	78	22,9	34	23,1	1
	Köln	128	37,6	55	37,4	
	Jena	26	7,6	11	7,5	
	Erfurt	57	16,8	25	17,0	
	Suhl	20	5,9	9	6,1	
	Gera	15	4,4	6	4,1	
	Nordhausen	8	2,4	3	2,0	
	Bonn	2	0,6	1	0,7	
	Ulm	6	1,8	3	2,0	
Geschlecht	Weiblich	160	47,1	78	53,1	0,224
	Männlich	180	52,9	69	46,9	
Alter (Median:62,8)	<Median	165	48,5	78	53,1	0,359
	>Median	175	51,5	69	50,1	
Histologie	AdenoCa	83	24,4	34	23,1	0,809
	AcinicCa	63	18,5	34	23,1	
	ACC	67	19,7	25	17,0	
	MEC	71	20,9	31	21,1	
	PEC	56	16,5	23	15,6	
T	T1	81	23,8	49	33,3	0,136
	T2	109	32,1	36	24,5	
	T3	66	19,4	27	18,4	
	T4	84	24,7	35	23,8	
N	N0	210	61,8	96	65,3	0,458
	≥N1	130	38,2	51	34,7	
M	M0	307	90,3	135	91,8	0,589
	M1	33	9,7	12	8,2	
UICC-Stage	I	71	20,9	37	25,2	0,648
	II	71	20,9	26	17,7	
	III	71	20,9	33	22,4	
	IV	125	36,8	51	34,7	
	unbekannt	2	0,6	0	0	
Malignität	Low grade	121	35,6	53	36,1	0,941
	High grade	98	28,8	44	29,9	
	unbekannt	121	35,6	50	34,0	
Therapie	adjuvant	179	52,6	72	49,0	0,639
	keine OP	14	4,1	4	2,7	
	primäre OP	130	38,2	61	41,5	
	unbekannt	17	5,0	10	6,8	
Radikale OP	Nein	48	14,1	27	18,4	0,381
	Ja	237	69,7	101	68,7	
	unbekannt	55	16,2	19	12,9	
Neck- Dissection	Nein	74	21,8	33	22,4	0,918
	Ja	228	67,1	96	65,3	
	unbekannt	38	11,2	18	12,2	
Facialiserhalt	Nein	13	3,8	7	4,8	0,860
	Ja	88	25,9	36	24,5	
	unbekannt	239	70,3	104	70,7	
Rezidiv	Nein	244	71,8	107	72,8	0,809
	Ja	80	23,5	35	23,8	
	unbekannt	16	4,7	5	3,4	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: statistisch signifikante Werte (p≤0,05);

In ein Cox-Modell müssen Variablen einbezogen werden, welche als prognostisch für die Zielgröße (Gesamtüberleben) angesehen werden. Daher findet für die Erstellung eines Cox-Modells eine Variablenselektion statt. Diese wurden aufgrund von klinischem Erfahrungswissen und statistisch signifikanten Ergebnissen der vorherigen Kapitel ausgewählt. Es wurden nur Variablen in die Modellbildung einbezogen, welche zum Diagnosezeitpunkt zur Verfügung stehen und welche in der univariaten Analyse (siehe Kapitel 3.2) einen statistisch signifikanten Unterschied in den Überlebenszeiten aufwiesen. Auf die Malignität wurde verzichtet, da es hinsichtlich dieser Variablen zu viele fehlende Angaben gibt und somit die Zahl der Fälle mit vollständigen Daten für die Modellbildung stark reduzieren würde. Auf die Einbeziehung des *UICC-Staging* wurde verzichtet, da sich das UICC aus dem TNM direkt ergibt. Das Berücksichtigen von beiden *Staging*-Systemen würde zu einer Ungenauigkeit führen, welche in der Literatur als Multikollinearität bezeichnet wird (Harrell 2015, S. 78). Zunächst wurde ein Cox-Modell inklusive der Variable Geschlecht erstellt (siehe Tabelle 15).

Tabelle 15: Multivariate Cox-Regression mit allen univariat-signifikanten Faktoren (n=340)

Variable		HR	95 % CI (unterer/oberer)	P-Wert
Geschlecht	Weiblich	1		
	Männlich	1,129	0,741-1,721	0,571
Alter		1,040	1,029-1,059	<0,001
Histologie	AdenoCa	1		<0,001
	AcinicCa	0,873	0,292-1,196	0,695
	ACC	0,591	0,556-2,190	0,143
	MEC	1,104	1,401-2,190	0,777
	PEC	2,391	0,741-4,081	0,001
T	T1	1		0,002
	T2	2,254	1,149-4,425	0,018
	T3	2,338	1,040-5,259	0,04
	T4	3,791	1,873-7,670	<0,001
N	N0	1		0,100
	≥N1	1,453	0,931-2,267	
M	M0	1		<0,001
	M1	4,101	2,408-6,984	

Es zeigte sich in der multivariaten Analyse, dass das Geschlecht kein statistisch unabhängiger Prognosefaktor ist ($p=0,571$). Die Variable Geschlecht wurde daher in eine weitere Modellierung nicht integriert. Für das N-Stadium zeigt sich eine Signifikanz von 0,10. Man kann dabei von einem statistischen Trend sprechen, weshalb die Variable dennoch eingeschlossen wurde. Das finale Modell (Tabelle 16) besteht aus folgenden Variablen: Alter, Histologie, T-Stage, N-Stage und M-Stage.

Tabelle 16: Finale Cox-Regression (n=340)

Variable		HR	95 % CI (unterer/oberer)	P-Wert
Alter		1,040	1,022-1,058	<0,001
Histologie	AdenoCa	1		<0,001
	AcinicCa	0,84	0,43-1,623	0,595
	ACC	0,57	0,285-1,151	0,118
	MEC	1,087	0,547-2,157	0,812
	PEC	2,362	1,386-4,026	0,002
T	T1	1		0,002
	T2	2,294	1,172-4,487	0,015
	T3	2,441	1,100-5,420	0,028
	T4	3,901	1,942-7,836	<0,001
N	N0	1		
	≥N1	1,428	0,919-2,219	0,113
M	M0	1		
	M1	4,182	2,464-7,099	<0,001

Die finale Cox-Regression zeigt den Einfluss der einzelnen Variablen auf das Sterberisiko (*Hazard Ratio*). Mit jedem Jahr älter, steigt dieses Risiko um 4% (HR: 1,040). Hinsichtlich der Histologie hat das Plattenepithelkarzinom ein 2,36-fach höheres Sterberisiko im Vergleich zum Adenokarzinom (Referenz), dahingegen halbiert sich das Sterberisiko des adenoidyzstischen Karzinom nahezu (HR:0,57). Ebenfalls zeigt sich der große Einfluss des TNM im Hinblick auf die Prognose. Patienten mit einem T4-Tumor haben ein nahezu vierfach höheres Sterberisiko (HR: 3,901) als mit einem T1-Tumor (Referenz). Patienten mit regionärem Lymphknotenbefall haben ein 1,43-faches Sterberisiko im Vergleich zu keinem Lymphknotenbefall (Referenz); dieses Ergebnis erweist sich als nicht statistisch signifikant (p=0,113). Erkrankte mit Fernmetastasierung haben ein nahezu vierfach höheres Sterberisiko (HR: 4,182) als welche ohne Fernmetastasen.

3.4 Nomogramm

3.4.1 Entwicklung

Aus dem Cox-Modell wurde mit der Statistiksoftware R und der dazugehörigen Erweiterung „rms“ ein Nomogramm gebildet (siehe Abbildung 11). Das Nomogramm ist eine Darstellungsform der Cox-Regression, mit der grafisch für jeden Patienten anhand von Risikofaktoren eine individuelle Prognose abgelesen werden kann.

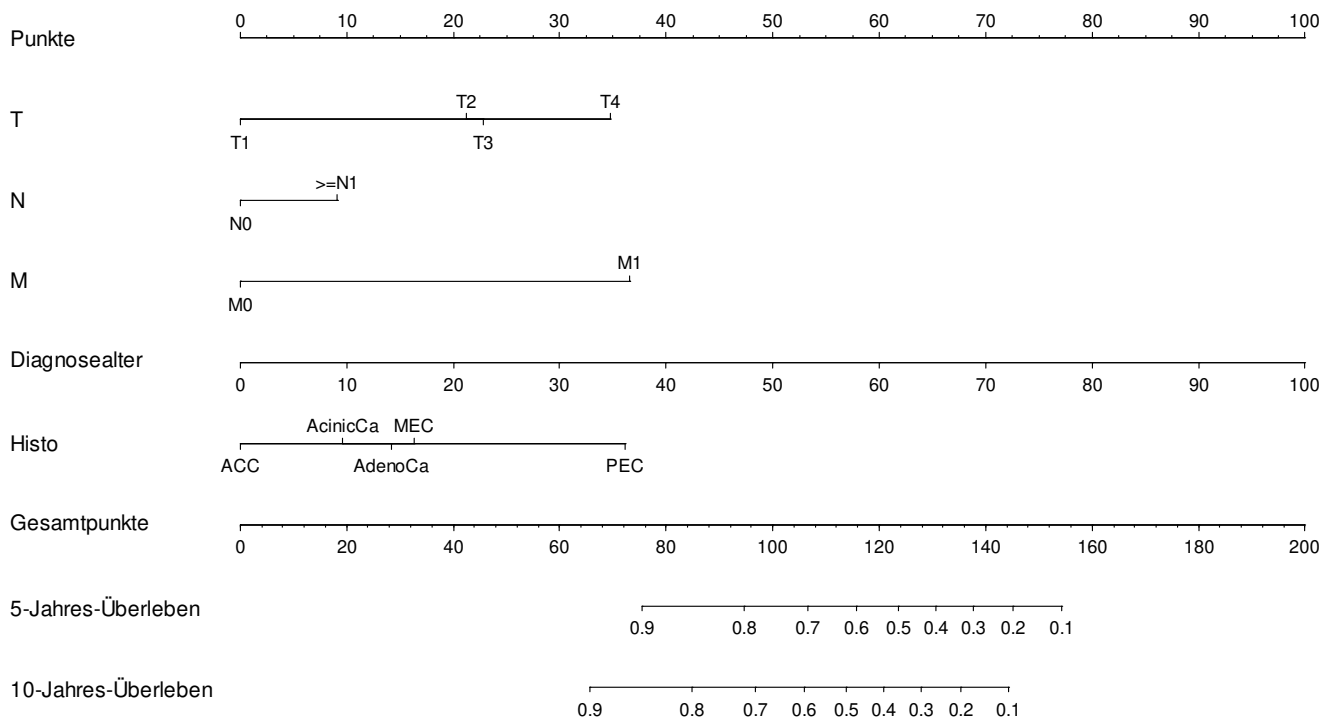


Abbildung 11: Nomogramm zur Vorhersage des 5-/10-Jahres-Gesamtüberlebens bei Parotiskarzinomen. Jedem Prädiktor wird dabei eine Punktzahl zugeordnet, die auf der oberen Skala abgelesen wird. Anhand der Gesamtpunktzahl auf der unteren Skala kann das 5- und 10-Jahres-Gesamtüberleben entsprechend abgelesen werden.

3.4.2 Validierung

Interne Validierung (Bootstrapping)

Es wurde eine interne Validierung durch *Bootstrapping* durchgeführt. Dabei werden Teile des Trainingsdatensatz ($n=340$) auf das Modell angewandt. Als Ergebnis erhält man den Somers Dxy-Wert. Dieser gibt die Korrelation zwischen der beobachteten und den vorhergesagten Wahrscheinlichkeiten an. Da dabei die Daten angewandt werden, welche auch das Modell bilden, sind die Vorhersagen in der Regel besser, als wenn man unabhängige Daten anwenden würde (Fachbegriff: Optimismus). Der Original-Dxy-Wert bei 200 Wiederholungen beträgt hier 0,681. Die *Bootstrapping*-Funktion kann diese Verzerrung berücksichtigen (Steyerberg 2009, S. 92–96). Der Modelloptimismus wird für dieses Nomogramm mit 0,0334 quantifiziert. Der Optimismus-korrigierte Dxy-Wert beträgt 0,574. Dies entspricht einem C-Index von 0,787 (Umrechnung: $Dxy = 2 \cdot (C - 0,5)$).

Multizentrische Validierung mit Validierungskohorte

Mit 30% der Gesamtkohorte ($n=147$) wurde die Validierung der Modellierung durchgeführt. Dabei wird das Nomogramm auf diese Kohorte angewandt, und das beobachtete Überleben mit dem vorhergesagten Überleben verglichen,

dabei entsteht ein Kalibrierungsdiagramm (Methode: Cox-Snell-Residuen für kumulierte Wahrscheinlichkeiten); siehe Abbildung 12. Im Diagramm ist eine gestrichelte Gerade (welche für die perfekte Übereinstimmung von beobachtetem und vorhergesagtem Überleben steht) und eine durchgezogene Kurve (welche für die Übereinstimmung von Beobachtungen und Vorhersagen des Nomogramms steht) zu sehen. Je näher sich die Kurven sind, desto genauer stimmen die beobachteten und die vorhergesagten Überlebenswahrscheinlichkeiten überein.

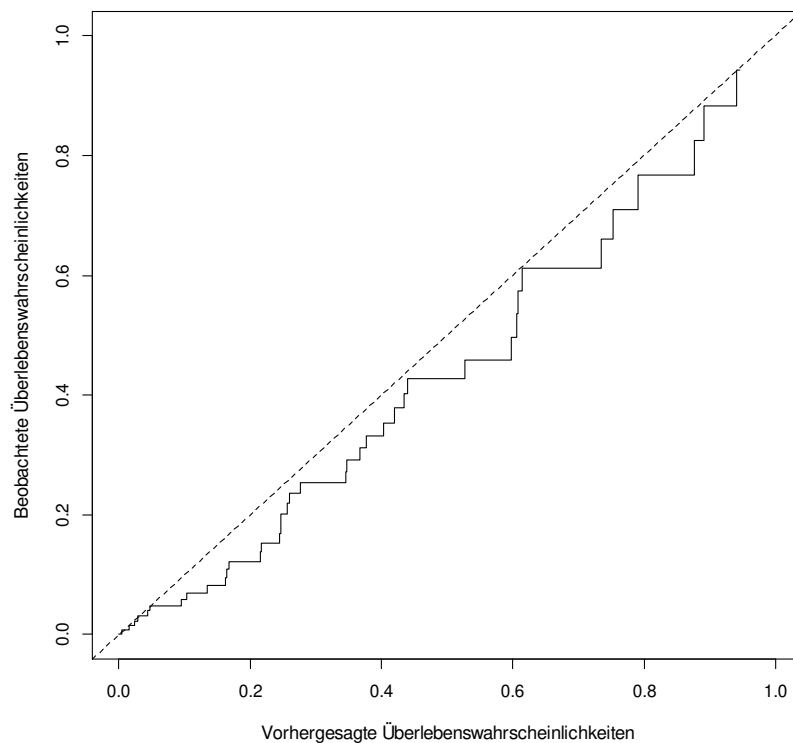


Abbildung 12: Kalibrierungsdiagramm: Die gestrichelte Linie steht für die perfekte Kalibrierung von beobachtetem und vorhergesagtem Überleben; die durchgezogene Linie steht für die Kalibrierung von beobachtetem und vorhergesagtem Überleben des Nomogramms für die Validierungsdaten

Der C-Index für das hier exemplarisch vorgestellte Nomogramm beträgt 0,861 (95% CI: 0,808-0,913). Im Vergleich zu anderen bereits publizierten Nomogrammen ist dies ein guter Wert (siehe 4.4). Um die Stabilität des Modells festzustellen, wurden 1000 Zufallsstichproben erstellt. Die daraus entwickelten 1000 Modelle wurden wie beschrieben validiert und die C-Indizes erfasst. Der Mittelwert beträgt 0,80, der Median ebenfalls 0,80. Die Standardabweichung beträgt 0,28. Der niedrigste erfasste C-Index ist 0,71, der höchste 0,89. Dies lässt den Schluss zu, dass es sich hierbei um ein stabiles Modell handelt.

3.5 Rekursive Partitionsanalyse / Klassifikations- und Regressionsbaum

3.5.1 Entwicklung

Die Trainings- und Validierungskohorten, welche im weiteren Verlauf der Arbeit erfolgreich validiert werden konnten, wurden analog zu Kapitel 3.3 durch eine Zufallsstichprobe (70%/30%) gebildet (siehe Tabelle 17). Die Standorte fanden ebenso wieder Berücksichtigung bei der Zusammensetzung der Kohorten. Die Trainingskohorte (n=340) wurde verwendet, um mit dem Package „*ctree*“ einen Klassifikations- und Regressionsbaum (CART) zu trainieren (siehe Abbildung 13 und Abbildung 18). Die Validierungskohorte (n=147) wurde verwendet, um das Modell zu validieren. Die Zusammensetzung der beiden Kohorten ist in Tabelle 17 dargestellt; der Chi2-Test zeigt, dass es statistisch signifikante Unterschiede in der Häufigkeitsverteilung gibt.

Tabelle 17: Übersicht über Trainings- und Validierungskohorte für die CART-Analyse

		Training		Validierung		P-Wert
		n	(%)	n	(%)	
Gesamt		340	100	147	100	
Standort	Gießen	78	22,9	34	23,1	1
	Köln	128	37,6	55	37,4	
	Jena	26	7,6	11	7,5	
	Erfurt	57	16,8	25	17,0	
	Suhl	20	5,9	9	6,1	
	Gera	15	4,4	6	4,1	
	Nordhausen	8	2,4	3	2,0	
	Bonn	2	0,6	1	0,7	
	Ulm	6	1,8	3	2,0	
Alter	≤69	229	67,4	88	59,9	0,111
	>69	111	32,6	59	40,1	
Histologie	AdenoCa	77	22,6	40	27,2	0,202
	AcinicCa	62	18,2	35	23,8	
	ACC	72	21,2	20	13,6	
	MEC	74	21,8	28	19,0	
	PEC	55	16,2	24	16,3	
PEC	Nein	285	83,8	123	83,7	0,967
	Ja	55	16,2	24	16,3	
T	T1	95	27,9	35	23,8	0,627
	T2	96	28,2	49	33,3	
	T3	64	18,8	29	19,7	
	T4	85	25,0	34	23,1	
T-Dichotom	T1/T2	189	55,6	84	57,1	0,751
	T3/T4	151	44,4	63	42,9	
N	N0	225	66,2	81	55,1	0,020
	≥N1	115	33,8	66	44,9	
M	M0	315	92,6	127	86,4	0,021
	M1	25	7,4	20	13,6	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: statistisch signifikante Werte ($p \leq 0,05$);

Klassifikations- und Regressions-Baum (CART)

Die Funktion „*ctree*“ entwirft einen „Baum“ mit „Zweigen“ an dessen Enden die Endknoten stehen (siehe 2.7). Dabei wird jedem dieser Endknoten anhand der zugehörigen Ausprägungen ein Teil der Patientenkohorte zugeordnet (siehe Abbildung 13). Bei der Bildung des „Baums“ werden die Endknoten nach dem Gesamtüberleben (Kaplan-Meier-Kurve) absteigend von links nach rechts sortiert. Links sind die Endknoten mit den Fällen mit dem besten Gesamtüberleben und rechts die Endknoten mit den Fällen mit dem schlechtesten Gesamtüberleben.

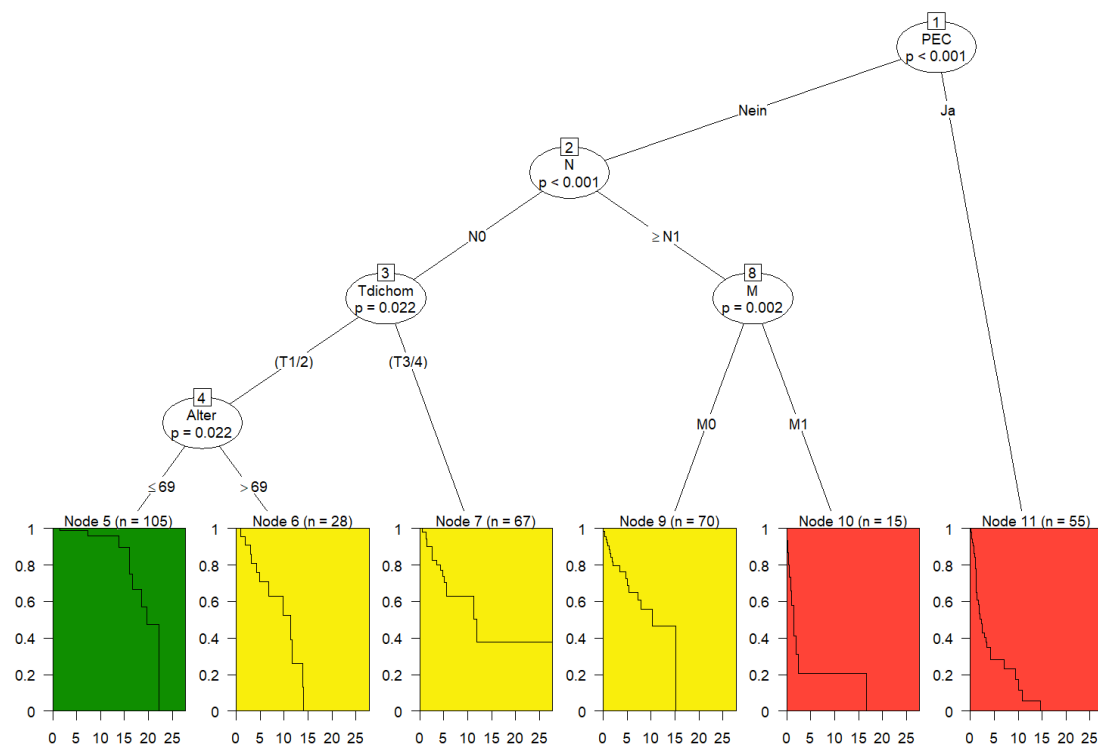


Abbildung 13: Klassifikations- und Regressions-Baum (CART) (n=340) mit Signifikanzniveau $p=0,05$, die Farben zeigen die Zugehörigkeit zu den hier definierten Risikogruppen an (rot= Hoch-Risiko-Gruppe, gelb= mittlere Risiko-Gruppe, grün= Niedrig-Risiko-Gruppe)

Anhand der Kaplan-Meier Kurven (siehe Abbildung 14) kann eine Einteilung in Risikogruppen vorgenommen werden.

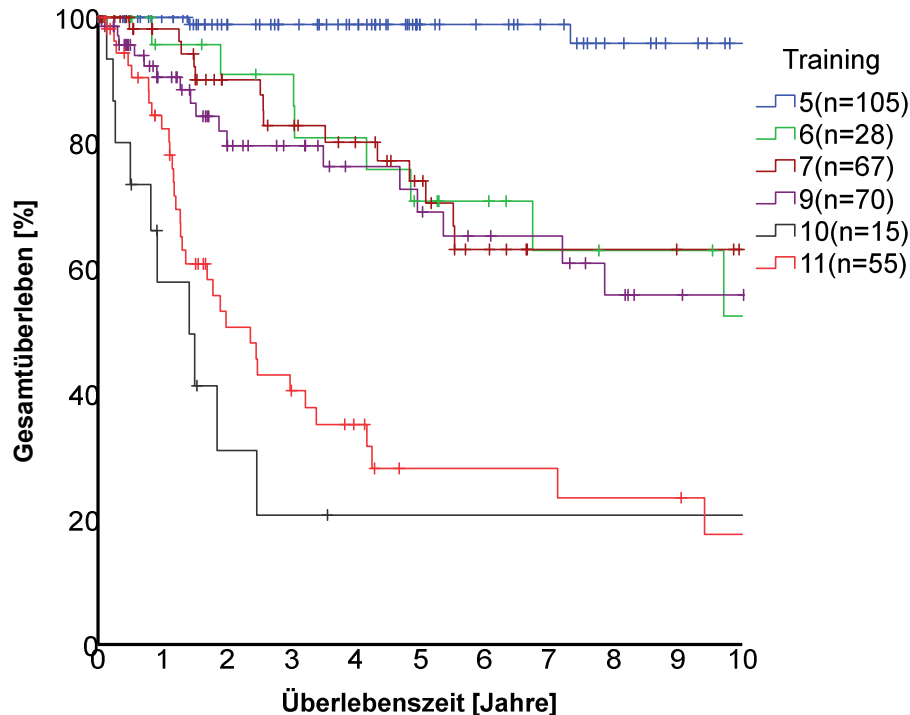


Abbildung 14: Kaplan-Meier-Überlebenszeitanalyse differenziert nach Subgruppen/Endknoten von Training (n=340)

Ein Vergleich der Ergebnisse des paarweisen Log-Rank Test (Tabelle 18) erlaubt das schrittweise Zusammenfassen von Kaplan-Meier-Kurven/Endknoten, welche sich hinsichtlich des Gesamtüberlebens statistisch nicht signifikant voneinander unterscheiden.

Tabelle 18: P-Werte (Log-Rank-Test; paarweise Vergleiche der Kaplan-Meier-Kurven der Subgruppen/Endknoten von Training in Abbildung 22), Grün: $p > 0,05$

Knoten	5	6	7	9	10	11
5		<0,001	<0,001	<0,001	<0,001	<0,001
6	<0,001		0,487	0,824	0,008	0,001
7	<0,001	0,487		0,401	<0,001	<0,001
9	<0,001	0,824	0,401		0,001	<0,001
10	<0,001	0,008	<0,001	0,001		0,465
11	<0,001	0,001	<0,001	<0,001	0,465	

Zunächst wurden die Endknoten 6 und 9 zusammengefasst und ein erneuter log-rank-Test durchgeführt. Im nächsten Schritt wurden Endknoten 7 und der „neue“ Endknoten 6+9 zusammengefasst. Im letzten Schritt wurden Endknoten 10 und 11 zusammengefasst. Es ergeben sich drei Gruppen, welche sich statistisch signifikant ($p < 0,05$) hinsichtlich ihres Gesamtüberlebens unterscheiden (siehe Abbildung 15). Diese stellen die drei Hauptrisikogruppen dar. Die Niedrig-Risiko-Gruppe (Endknoten 5)

hat eine 5-Jahres-Überlebensrate von 98,8%. Die mittlere-Risiko-Gruppe (Endknoten 9+6+7) hat eine 5-Jahres-Überlebensrate von 70,7%. Die Hoch-Risiko-Gruppe (Endknoten 10+11) hat eine 5-Jahres-Überlebensrate von 26,2%.

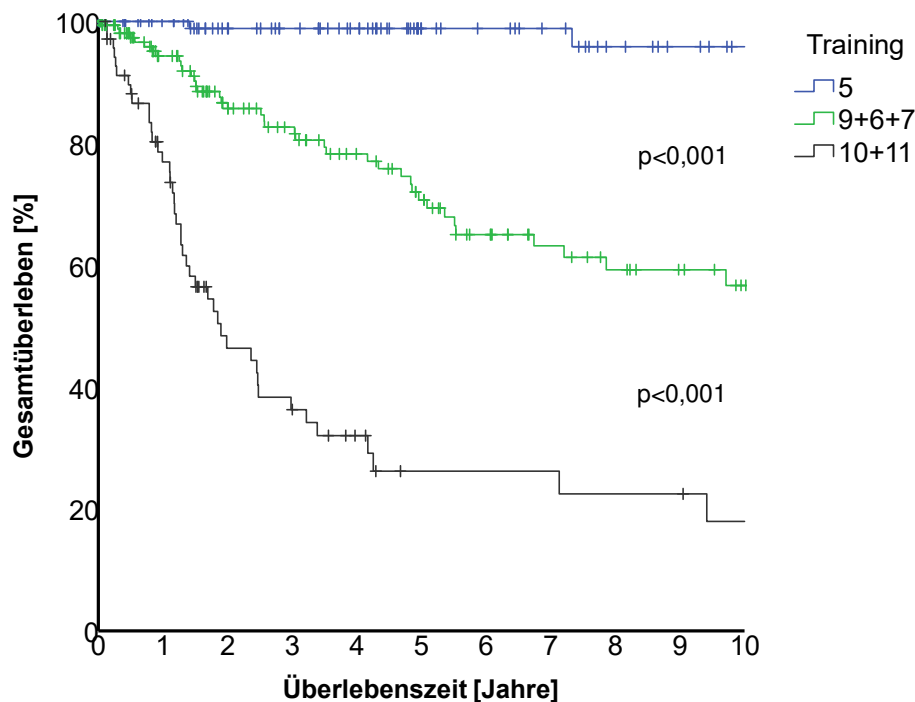


Abbildung 15: Kaplan-Meier-Überlebenszeitanalyse differenziert nach erstellten Hauptrisikogruppen aus den Subgruppen von Training (n=340) und log-rank Test

3.5.2 Validierung

Je nach verwendeter Zufallskohorte verändert sich der entstandene Baum; dies spricht grundsätzlich für ein instabiles Modell. Eine Validierung ist daher notwendig.

Validierung Methode 1: Beschreiben der Modellstabilität

Mit 1000 Zufallsstichproben wurde jeweils ein CART-Baum gebildet und der grundsätzliche Aufbau der Bäume beschrieben (siehe Tabelle 19). Am häufigsten wurde die Variable PEC als erste (wichtigste) Variable ausgewählt (in 51,8% der Modelle). Am zweithäufigsten wurde die Variable Alter ausgewählt (42,1% der Modelle). Weit seltener wurde der M-Status (4,8%) und N-Status (0,3%) ausgewählt. Somit lässt sich die Stabilität des Modells beschreiben, die näherungsweise zur Hälfte zwischen PEC und Alter als wichtigste Variable schwankt.

Tabelle 19: Häufigkeit der Variablen die in den 1000 trainierten CART-Modellen an erster Stelle aufgerufen wurden.

Übersicht: 1000 CART-Modelle		
1.Variable	Anzahl	%
PEC	518	51,8
Alter	431	43,1
M	48	4,8
N	3	0,3

Validierung Methode 2: Vergleichen des Gesamtüberlebens

Exemplarisch wurde der CART-Baum der Zufallskohorte 15 („seed 15“, n=340) entwickelt. Schrittweises Zusammenfassen der Subgruppen/Endknoten anhand des Log-Rank-Test (mit $p=0,05$ als Signifikanzgrenze) ergab drei Hauptrisikogruppen (siehe Tabelle 22). Die Validierung wurde mit n=147 durchgeführt. Es wird betrachtet, in welchen Endknoten welcher Fall des Validierungs-Datensatzes zugeordnet werden würde (siehe Tabelle 20). Anhand dessen wurden neue Kaplan-Meier-Überlebenszeitanalysen durchgeführt (siehe Abbildung 16, Tabelle 21). Diese können analog der Trainingskohorte ebenfalls durch schrittweises Vorgehen in drei Hauptrisikogruppen zusammengefasst werden.

Tabelle 20: Fallzusammensetzung: Verteilung der Fälle/Ereignisse auf die einzelnen Subgruppen/Endknoten

Validierungskohorte(n=147) und Zuordnung auf Endknoten		
Endknoten	Anzahl	Ereignisse
5	38	3
6	17	4
7	15	5
9	39	11
10	14	9
11	24	14

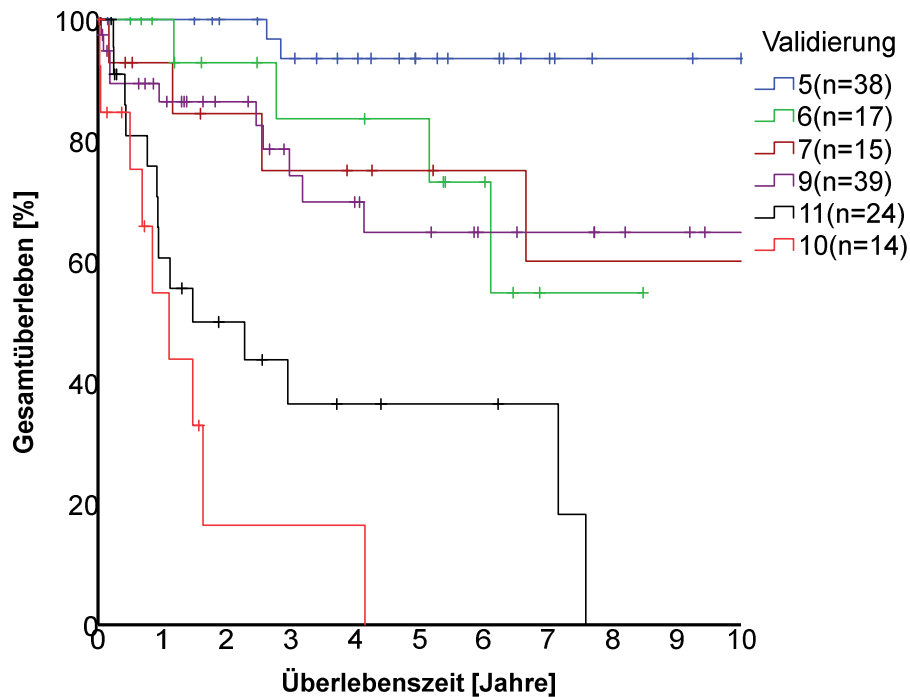


Abbildung 16: Kaplan-Meier-Überlebenszeitanalyse für die einzelnen Subgruppen/Endknoten (5-7, 9-11) von Validierung n=147

Tabelle 21: P-Werte (Log-Rank-Test; paarweise Vergleiche der Kaplan-Meier-Kurven von Subgruppen/Endknoten von Validierung in Abbildung 24), Grün: $p > 0,05$.

Knoten	5	6	7	9	10	11
5		0,015	0,012	0,006	<0,001	<0,001
6	0,015		0,957	0,707	<0,001	0,009
7	0,012	0,957		0,862	<0,001	0,015
9	0,006	0,707	0,862		<0,001	0,002
10	<0,001	<0,001	0,001	<0,001		0,117
11	<0,001	0,009	0,015	0,002	0,117	

Das Vorgehen des Zusammenfassens von Endknoten findet analog zu dem der Trainingskohorte statt. Es entstehen durch schrittweises Zusammenfassen ebenfalls drei Gruppen, welche sich statistisch signifikant hinsichtlich des Gesamtüberlebens unterscheiden. Es entstehen dieselben Hauptrisikogruppen wie bei der Trainingskohorte (vgl. Abbildung 15 und Abbildung 17).

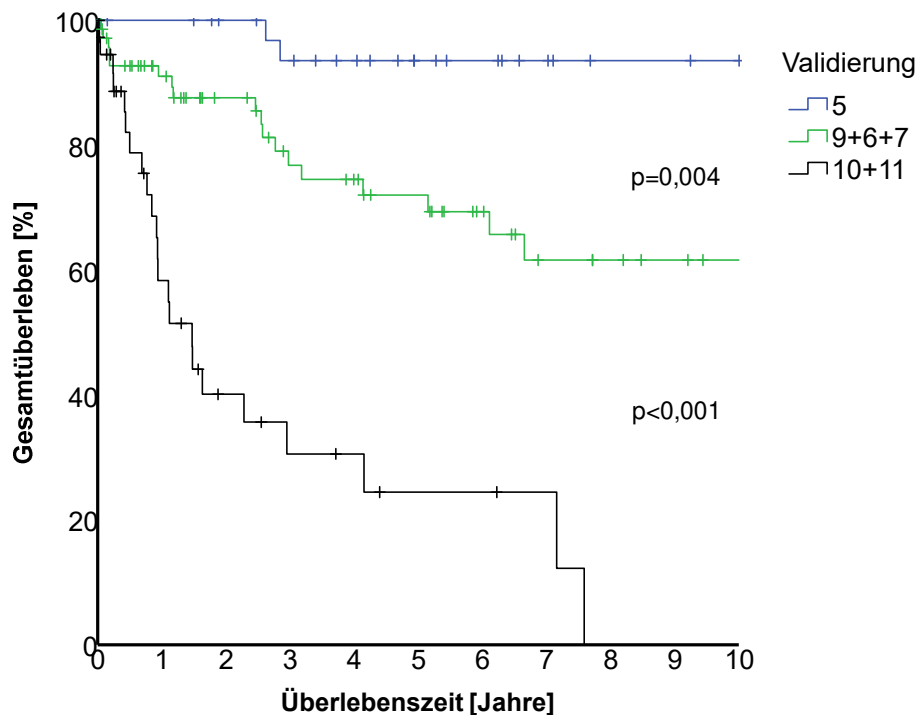


Abbildung 17: Kaplan-Meier-Überlebenszeitanalyse differenziert nach erstellten Hauptrisikogruppen aus den Subgruppen von Validierung $n=147$ und log-rank-Test(p -Wert)

Dies zeigt, das trainierte Modell würde neue Fälle (Validierungsdatensatz) genauso anhand ihres prognostizierten Gesamtüberlebens sortieren, dass ebenfalls drei Hauptrisikogruppen entstehen. Unsere Anforderungen (Bildung von drei Risikogruppen in beiden Kohorten) an die Validierung sind damit erfüllt.

Abbildung 21 zeigt die Zuteilung der Fälle den entsprechenden Risikogruppen analog zu Abbildung 18 in einem Flussdiagramm (englische Version).

entwickelt, um die 5-Jahres-Überlebenswahrscheinlichkeit zu berechnen. Es handelt sich dabei um eine digitale Form des Nomogramms.

Abbildung 19 zeigt einen Screenshot der fertigen Webanwendung mit einem Fallbeispiel.

Parotid Carcinoma Survival Probability Tool, Department of Otorhinolaryngology, University of Giessen

Enter your patient characteristics and hit Submit

Choose the T-Stage:

T2

Choose the N-Stage (N0 vs. N1 and higher):

>=N1

Choose the M-Stage (M0 vs. M1 and higher):

M0

Histological type;

PEC

Age at diagnosis:

60

Submit

Predicted Survival Probability

The Probability of 5-Year-Overall-Survival is:

43.54 %

Disclaimer:

This App is a multicenter developed and validated prognostic model. It provides the individualized 5-Year-Survival Probability for Overall Survival. This model should only be used by physicians who are familiar with the complexity of treatment decisions in cancer treatment and the limits of statistic survival modeling. This model should not be used directly by patients.

Contact:

Prof. Dr. C. Wittekindt, Claus.Wittekindt@hno.med.uni-giessen.de
ZA Christian A. Krauss, christian.a.krauss@dentist.med.uni-giessen.de

Scientific reference supporting this model:

Author1, Author2, Author3, Author4. Multicenter developed and validated nomogram for predicting overall survival in parotid gland carcinoma patients[Journal of XY]

Abbildung 19: Screenshot der Webanwendung zeigt Fallbeispiel: 60-jähriger Patient mit einem T2, N1, M0, Plattenepithelkarzinom

4 Diskussion

4.1 Prognosefaktoren

4.1.1 Geschlecht und Alter

Hinsichtlich der Geschlechtsverteilung der analysierten Kohorte im Vergleich zu anderen Studien dominiert das männliche Geschlecht mit 51,1%, vgl. mit der Literatur: 48% - 60,3% (Chakrabarti et al. 2018; Li et al. 2017; Erovic et al. 2015; Ali et al. 2014; Bjørndal et al. 2012; Terhaard et al. 2004). In der univariaten Kaplan-Meier-Überlebenszeitanalyse zeigt sich, dass das männliche Geschlecht im Vergleich zum weiblichen mit einem geringeren Gesamtüberleben assoziiert ist (siehe Abbildung 5a). In unserer Kohorte haben Männer ein älteres Diagnosealter, sind häufiger von einer aggressiveren Tumorphistologie betroffen (z.B. PEC), haben häufiger ein T3/4- bzw. III/IV-Stadium, häufiger regionären Lymphknotenbefall, Fernmetastasen und *high-grade* Tumoren (siehe Tabelle 7). Diese Prädiktoren sind mit einem geringeren Überleben assoziiert. Dies deckt sich mit der gängigen Literatur (siehe Tabelle 4). Gemäß Literaturangaben konsumieren Frauen auch weniger Alkohol und rauchen weniger (Cheung et al. 2011). Der Verdacht auf Kollinearität des Geschlechts mit anderen diagnostischen Variablen liegt daher nahe. Auf die Einbeziehung des Geschlechts in das finale Modell wurde daher verzichtet.

Mit einem medianen Alter von 62,8 Jahren sind Patienten dieser Kohorte im Vergleich zu anderen Studien ähnlich alt (medianes Alter von 56 – 63 Jahren (Li et al. 2017; Erovic et al. 2015; Ali et al. 2014; Bjørndal et al. 2012; Terhaard et al. 2004). Die Arbeit von Chakrabarti et al. sticht mit einem medianen Alter von 46 Jahren heraus - die Autoren sprechen dabei selbstkritisch von einer Stichprobenverzerrung, da zahlreiche Patienten mit geringer Nachbeobachtungszeit ausgeschlossen wurden. Genauere Angaben werden von den Autoren allerdings nicht gemacht (Chakrabarti et al. 2018). Ein hohes Patientenalter ist mit einem geringeren Gesamtüberleben assoziiert (siehe Abbildung 5b). Es ist naheliegend, dass ältere Patienten eine niedrigere Überlebensrate aufweisen als jüngere Patienten, da die allgemeine Sterblichkeit mit steigendem Lebensalter ansteigt. Die Literatur zeigt, dass dies nicht nur dem chronologischen Alter geschuldet ist, sondern auch den im Alter häufiger vorliegenden Komorbiditäten, den *high-grade* Tumoren und den höheren Tumorstadien (Becker et al. 2017; Bjørndal et al. 2016). Komorbiditäten konnten in dieser Arbeit nicht ausgewertet werden, da diese nicht

registerübergreifend erfasst wurden. Die Literaturaussagen werden dennoch von der Häufigkeitsverteilung in unserer Kohorte unterstützt. Ältere Patienten sind demnach eher männlich, öfter von „aggressiveren“ Histologien betroffen (z.B. PEC) und haben höhere Tumorstadien (siehe Tabelle 8).

4.1.2 Histologie

Die Überlebenszeiten unterscheiden sich je nach histologischer Entität statistisch signifikant voneinander. Auch in der Literatur ist bekannt, dass sich die Prognose je nach Tumorphistologie unterscheidet (siehe Tabelle 4) (Israel et al. 2019). In 3.1.3 wird dargelegt, dass manche Histologien (Azinuszellkarzinom und Mukoepidermoidkarzinom) eher mit günstigeren Ausprägungen der Risikofaktoren (wie z.B. jüngeres Patientenalter, *low-grade* etc.) assoziiert sind, während Adeno- und Plattenepithelkarzinom eher mit schlechteren Risikofaktorausprägungen assoziiert sind. Die Literatur bestätigt diese Zuordnung von Hochrisiko- und Niedrig-Risiko-Histologien (Seethala 2009). Auch in unseren Modellen zeigt sich z.B. für das Plattenepithelkarzinom (PEC) eine große Bedeutung (höchstes Sterberisiko in der Cox-Regression, siehe Abbildung 11, Tabelle 16 und wichtigster Faktor in der CART-Analyse, Abbildung 13). Da die Prognose zu großen Teilen von der diagnostizierten Histologie abhängt, ist es sinnvoll Therapieentscheidungen daran zu orientieren (siehe Kapitel 1.2 Diagnostik).

Auch unter den seltenen Histologien zeigen sich deutliche Unterschiede in den Überlebenszeiten. Von den fünf untersuchten, seltenen Histologien mit $n > 10$ wurden die 5-Jahresüberlebenszeiten berechnet (Basalzelladenokarzinom(BCAC): 100% Epithelial-Myoepitheliales Karzinom(EMC): 76,4% Invasiv-duktales Karzinom(IDC): 40,5% Karzinom-ex-pleomorphes Adenom(CaexPA): 61,1% Myoepitheliales Karzinom(MC): 69,2%). Verglichen mit der Literatur wird für das IDC ein 5-Jahresüberleben von 43% angegeben (Johnston et al. 2016), für das BCAC von 79% (Zhan und Lentsch 2016), für das EMC von 64% (Xiao et al. 2016), für das CaexPA von 50,3% (Zhao et al. 2013) und für das MC von 71% (Su et al. 2015). Obwohl für die Auswertung eine relativ große Patientenkohorte zur Verfügung stand, zeigt sich, dass die in den WHO-Klassifikationen aufgelisteten Entitäten teilweise sehr selten bzw. gar nicht vorkommen (siehe Tabelle 5). Zu beachten ist, dass sich während des Beobachtungszeitraums die WHO-Klassifikation mehrfach geändert hat (siehe Tabelle 3). Obwohl die Histologie ein wichtiger prognostischer Faktor ist, wird sie bei den gängigen *Staging*-Systemen wie TNM/UICC nicht verwendet (siehe Tabelle 1, Tabelle 2). Auch bei den alternativen *Staging*-

Systemen (z.B. TANIS-8, TANIS-3, Snyderman, Hart, Berg) bleiben histologische oder patientenspezifische Aspekte (z.B. Alter, Komorbiditäten) unberücksichtigt. Kreppel et al. hat diese Systeme miteinander verglichen und kommt zu dem Schluss, dass sie um weitere Patienten-, Tumor- und Therapie-assoziierte Faktoren erweitert werden müssen. Als mögliche Darstellungsform wird ein Nomogramm vorgeschlagen (Kreppel et al. 2013). Auch in Anbetracht der Modellentwicklung und Auswertung dieser Arbeit kann die Schlussfolgerung gestellt werden, dass das *Staging* um weitere Faktoren wie die Histologie ergänzt werden sollte. Eine definitive pathohistologische Diagnosestellung ist dafür notwendig.

4.1.3 TNM/UICC-Staging

Das Einteilen von Patienten in vereinfachte Risikogruppen ist wichtig, um Prognose und Therapieentscheidung sinnvoll treffen zu können (Kreppel et al. 2013). Das TNM-*Staging* beinhaltet die Information über Beschaffenheit und Ausbreitung des Tumors und erlaubt die Abschätzung einer Prognose. Obwohl das TNM-*Staging* der Vereinfachung der Prognosestellung dienen soll, sind über 40 unterschiedliche Kombinationen von TNM-Ausprägungen möglich. Um die Prognoseabschätzung weiter zu vereinfachen, wurden zusammenfassende Stagingssysteme entwickelt (z.B. *UICC-Staging*). Ein steigendes T-Stadium bedeutet prinzipiell, dass der Tumor größer geworden ist oder bestimmte Grenzen erreicht hat (z.B. extraparenchymatöse Ausdehnung, Haut, Schädelbasis), an die man einen „Schweregrad“ koppelt. Je größer ein Tumor geworden ist, desto größer ist auch die Wahrscheinlichkeit, dass angrenzendes Gewebe infiltriert wurde bzw. schon regionäre Lymphknoten befallen sind (Ettl et al. 2014; Kreppel et al. 2013). Das N-Stadium beschreibt den Befall der regionären Lymphknoten; dieser steht für die Ausbreitungstendenz des Tumors, und die Wahrscheinlichkeit einer Fernmetastasierung steigt. Fernmetastasen sind in der Onkologie generell mit einer schlechten Prognose verbunden, da eine alleinige chirurgische Therapie i.d.R. nicht mehr möglich ist (Ettl et al. 2012; Ali et al. 2015). In unserer Auswertung (siehe Abbildung 8) zeigt sich, dass mit steigendem TNM/UICC-Stadium die Überlebenswahrscheinlichkeit sinkt. Damit werden die bekannten Literaturaussagen bestätigt (siehe Tabelle 4). Unsere Modellierungen (Cox-Regression/CART) zeigen aber auch, dass es andere Prädiktoren wie Alter und Histologie gibt, welche einen statistisch signifikanten Einfluss auf das Gesamtüberleben haben. Eine Prognosestellung, welche allein auf dem TNM-Status basiert, ist daher als unzureichend anzusehen.

Eine generelle Schwäche von statistischen Auswertungen, die sich auf lange Beobachtungszeiträume beziehen, ist, dass *Staging*-Systeme währenddessen überarbeitet werden können und u.U. Datensätze nicht entsprechend umgewandelt werden können. So galt zu Beginn der Beobachtungszeit (1972) die 3. Auflage des TNM, während die neuesten Fälle anhand der 8. Auflage des TNM (siehe Tabelle 1) diagnostiziert wurden. Unabhängig von den Änderungen der TNM-Auflage zueinander, bleibt die prinzipielle Systematik der Klassifikation jedoch beständig.

4.1.4 Malignität

Ein geringer Differenzierungsgrad der Tumorzellen (*high-grade*) wird mit einem geringeren Überleben, mehr regionärem Lymphknotenbefall und Fernmetastasen assoziiert (Rayess et al. 2017; Ali et al. 2015). Dies deckt sich mit unserer Literaturübersicht, in der die Malignität als wichtig für die Prognose erachtet wird (siehe Tabelle 4), insbesondere in Kombination mit dem TNM-*Staging* (Harbo et al. 2002). Tumorentitäten, die als *high-grade* klassifiziert wurden, gelten als schnell wachsend bzw. als „aggressiv“ (Seethala et al. 2019). Im Gegensatz dazu gelten *low-grade* Tumoren als langsam wachsend. In der Literatur ist ebenfalls beschrieben, dass manche histologischen Entitäten eher mit einem hohen Differenzierungsgrad assoziiert sind und andere mit einem niedrigen (Freling et al. 2016). In den für diese Arbeit zur Verfügung stehenden Daten ist der Differenzierungsgrad nur unzureichend erfasst und wurde deshalb mit Rücksichtnahme auf die Fallzahl nicht in die multivariaten Analysen aufgenommen. In der univariaten Kaplan-Meier-Überlebenszeitanalyse zeigt sich die Malignität als signifikanter Faktor für das Gesamtüberleben (siehe Abbildung 9). Die Malignität sollte daher präoperativ diagnostiziert und dokumentiert werden z.B. durch Feinnadelaspirationszytologie und in die Therapieentscheidung einfließen (Jang et al. 2018).

4.2 Variablenselektion / Cox-Regression / Nomogramm

Die Cox-Regression ist eine der am häufigsten angewendeten Methoden für Überlebenszeitanalysen in der Medizin (Steyerberg 2009, S. 77). Für die Variablenselektion bei klinischen Modellen gibt es viele intensiv diskutierte Herangehensweisen. Es gibt keinen eindeutigen Konsens darüber, welches das richtige Vorgehen zur Auswahl der Variablen ist. Für ein möglichst gutes Modell zur Vorhersage

der Zielgröße Gesamtüberleben sollen möglichst diejenigen Variablen genutzt werden, welche das Gesamtüberleben am besten bzw. vollständigsten erklären. Eine Möglichkeit ist die automatische, schrittweise Variablenselektion durch die Statistiksoftware. Dabei wird zwischen Vorwärts- und Rückwärts-Auswahl unterschieden (je nachdem werden schrittweise signifikante Prädiktoren hinzugefügt oder nicht-signifikante entfernt). Unter den automatischen Selektionsmechanismen wird die Rückwärts-Auswahl in der Literatur bevorzugt (Steyerberg 2009, S. 194–198). Die schrittweisen Methoden werden als objektiv angesehen; sie können von jedem Statistiker reproduziert werden (Steyerberg 2009, S. 196). Schrittweise Variablenselektion bietet sich dann an, wenn noch unklar ist, welche Variablen für die Modellentwicklung bedeutsam sind. Diese Methoden werden jedoch auch kritisch betrachtet. Die schrittweise Variablenselektion würde so z.B. Variablen in das Modell hinzufügen, welche rein zufällig einen erklärenden Zusammenhang mit der Zielvariablen haben. Ebenfalls kann der Algorithmus fälschlicherweise relevante Variablen unberücksichtigt lassen. Zusätzlich wird über diese Methode auch berichtet, dass sie Signifikanzniveaus verfälscht, Konfidenzintervalle zu schmal und Parameterschätzwerte zu hoch schätzt (Harrell 2015, S. 67–68). In dieser Arbeit wurde daher auf eine automatische Variablenselektion verzichtet. Die Variablen wurden aufgrund klinischer Erfahrungen und statistischer Ergebnisse ausgewählt. Die manuell erstellte Cox-Regression zeigt, dass die Bedeutung der Variable „Geschlecht“ im Zusammenspiel mit den anderen Variablen insignifikant geworden ist. Eine Multikollinearität des Geschlechts mit den anderen Variablen liegt daher nahe. Das Ergebnis unserer Cox-Regression deckt sich mit dem bereits bekannten Wissen über bedeutsame Prädiktoren: Mit steigendem Alter und TNM-Status sinkt die Überlebenswahrscheinlichkeit. Ebenfalls bestätigt sich, dass es Histologien gibt, welche zu einer schlechteren Prognose beitragen. Bedingt durch die hohe Fallzahl hat dieses Modell eine hohe Güte. Neben den Daten des Gießener Standortes konnten Daten anderer deutscher Standorte ausgewertet werden, sodass es sich um ein multizentrisches Modell handelt.

Die Cox-Regression wurde weiterhin in Form eines Nomogramms dargestellt (siehe Abbildung 11). Das Nomogramm selbst ist dabei kein eigenes Modell, sondern lediglich eine Darstellungsform. Die Technik der Nomographie wurde 1884 vom französischen Ingenieur Philbert Maurice d'Ocagne erfunden und wurde zur Vereinfachung technischer Formeln verwendet. Nomogramme haben daraufhin in vielen Bereichen von Technik und Wissenschaft Verwendung gefunden (Ali et al. 2013).

Im Gegensatz zu anderen Arbeiten wurden sowohl eine interne Validierung durchgeführt als auch eine Validierung mit Daten, welche nicht für die Modellentwicklung verwendet wurden. Darüber hinaus wurde durch die 1000-fache Modellentwicklung/Validierung die Stabilität des Modells beschrieben. Damit wurden in dieser Arbeit drei Methoden zur Überprüfung der Validität des Modells angewandt. Mit einem C-Index von 0,787 (interne Validierung) und 0,861 (95% CI: 0,808-0,913; multizentrische Validierung mit unabhängigen Daten) handelt es sich um eine gute Validität. In Kapitel 4.4 werden die Ergebnisse der Validierung in Bezug zu bereits publizierten Modellen diskutiert.

4.3 Rekursive Partitionsanalyse / Klassifikations- und Regressionsbaum

Das PEC scheint der wichtigste prognostische Faktor zu sein (er rangiert ganz oben im CART-Baum). Dies führt dazu, dass bei der Risikogruppen-Definition (siehe Abbildung 18, Tabelle 22) alle Patienten mit einem PEC zur Hochrisikogruppe gehören. Unter Kapitel 3. Material und Methoden wurden die Möglichkeiten der Validierung vorgestellt. Die 147 Fälle des Validierungsdatensatzes wurden entsprechend ihrer Charakteristiken den einzelnen Endknoten zugeordnet (anhand des von n=340 entwickelten CART). Die Kaplan-Meier-Überlebenszeitanalyse zeigt, dass sich durch schrittweises Zusammenfassen der Subgruppen (Signifikanzgrenze: $p=0,05$) dieselben Hauptrisikogruppen bilden lassen, wie bei der Trainingskohorte. Daher ist für die Zufallsstichprobe diese Form der Validierung erfolgreich. Um die Stabilität des Modells abhängig von der verwendeten Zufallsstichprobe zu beschreiben, wurde eine 1000-fache Modellentwicklung/-validierung durchgeführt. Es hat sich ebenfalls herausgestellt, dass das PEC häufiger als wichtigster Prädiktor bestätigt wurde, häufiger als z.B. das Diagnosealter. CART-Bäume eignen sich gut, um komplexe Zusammenhänge zwischen Prädiktoren und Zielgröße einfach darzustellen und eine klinische Anwendbarkeit zu ermöglichen. Die Nutzung eines CART ist aufgrund der grafischen Darstellung und fehlender mathematischer Gleichung (vgl. zur logistischen oder Cox-Regression) einfach. Dennoch können auch CART-Modelle vielschichtig und komplex werden. In diesen Fällen ist die Cox-Regression mit seinen eindeutigen *Hazard Ratios* im Vorteil (Shariat et al. 2008). CART-Bäume werden in der Literatur kontrovers diskutiert. Kattan et al. kommt zu dem Schluss, dass die althergebrachten Methoden wie z.B. die Cox-Regression einen besseren prognostischen Effekt haben (Shariat et al. 2008; Kattan 2003). Mehrere Publikationen sehen das Nomogramm als der Risikogruppen-Einteilung überlegen an. Das Zusammenfassen der Ausprägungen von Prädiktoren in

Risikogruppen führt naturgemäß zu einem Verlust an Informationen und damit auch zu einem Verlust an Vorhersagegenauigkeit. Dahingegen erhält man beim Nomogramm für jede mögliche Kombination an Patientencharakteristika eine individualisierte Überlebenswahrscheinlichkeit (Shariat et al. 2008).

Schlussfolgernd lässt sich sagen, dass mit der CART-Analyse leicht Zusammenhänge dargestellt und analysiert werden können.

4.4 Überblick über bereits publizierte Modelle

In der Vergangenheit wurden mehrere Überlebenszeitmodelle für Speicheldrüsenkarzinompatienten veröffentlicht:

Vander Poorten-Score:

Als grundlegend im Parotiskarzinom-*Modeling* wird der Vander Poorten-Score gesehen. Obwohl dieses Modell das rezidivfreie Überleben als Grundlage nimmt, wird es hier wegen der Bedeutung für das gesamte Parotis-*Modeling* vorgestellt. Dieser wird als erster publizierter und validierter prognostischer *Score* angesehen. Die niederländische Arbeitsgruppe veröffentlichte 1999 zwei Prognosesysteme (Prä-OP und Post-OP). Eine Gewichtung der Prädiktoren T-Stadium, N-Stadium, Schmerzen, Diagnosealter, Gesichtsinvasion, Facialisnervfehlfunktion, Perineuralwachstum und Resektionsrand führen zu einem prognostischen Wert, anhand dessen eine Risikogruppe zugeordnet wird. Kritisch zu hinterfragen ist dabei, ob die erklärenden Prädiktoren T-Stadium, Gesichtsinvasion, Perineuralwachstum und Facialisnervfehlfunktion jeweils redundante Rückschlüsse auf die Tumorgroße zulassen. Dann würde es sich um Multikollinearität handeln, welche zu einer ungenauen Schätzung des statistischen Prädiktoreinflusses führen würde. 2003 und 2017 wurde der Score mit guten Ergebnissen validiert (Poorten et al. 2009; Paderno et al. 2018; Vander Poorten et al. 2003). Es ist jedoch festzuhalten, dass das Vander Poorten-Modell mit n=151 eine viel schlechtere Datengrundlage hat als die in dieser Arbeit entwickelten Modelle.

Memorial Sloan Kettering Cancer Center-Nomogram for Overall Survival (MSKCC):

Die Arbeitsgruppe um Ali et al. veröffentlichte 2014 das erste Nomogramm zur Vorhersage der postoperativen 10-Jahres-Gesamtüberlebensrate für Parotiskarzinompatienten (Ali et al. 2014). Das Nomogramm wurde 2018 von Hay et al. validiert und weist einen C-Index von 0,833 (95% CI 0.758-0.908) auf (Hay et al. 2018). Die verwendeten Prädiktoren sind das Vorhandensein des klinischen T4-Stadiums, die

Malignität, das Vorhandensein von Perineuralinvasion und maximalem Tumorausmaß. Auffällig ist, dass die Variablen maximales Tumorausmaß und klinisches T4-Stadium grundsätzlich Rückschlüsse auf dieselben Tumorcharakteristika zulassen (Größe des Tumors). Es ist wahrscheinlich, dass dadurch zwei Variablen mit einer hohen biologischen und statistischen Korrelation in die Modellbildung integriert wurden, was die Modellinterpretation schwieriger macht (Multikollinearität).

Das MSKCC-Modell wurde aus 301 heterogenen Speicheldrüsenkarzinomen trainiert (266 sind in der Parotis lokalisiert, 30 in der Glandula submandibularis, 5 in der Glandula sublingualis). Unser Modell wurde lediglich aus Parotiskarzinomfällen (n=487) trainiert. Dies ist sinnvoll, da schon mehrfach festgestellt wurde, dass es statistisch signifikante Überlebenszeitunterschiede abhängig von der Tumorlokalisation gibt (siehe Tabelle 4).

Li et al.

Die chinesische Arbeitsgruppe veröffentlichte 2017 ein Nomogramm (n=4218, SEER Database) zur Vorhersage des Gesamtüberlebens und validierte dies (n=244). Es hat einen C-Index von 0,829 (95 % CI, 0,783–0,869) (Li et al. 2017). Das Li-Nomogramm wurde genau wie das MSKCC-Nomogramm aus Speicheldrüsenkarzinomen (unabhängig ihrer Drüsenlokalisation) entwickelt (Li et al. 2017). Im Gegensatz zu den anderen hier vorgestellten Nomogrammen wird hierbei die gewählte Therapie als Prädiktor einbezogen, worauf wir bewusst verzichtet haben, damit das Modell zum Diagnosezeitpunkt anwendbar ist. Darüber hinaus ergibt sich in der klinischen Realität die gewählte Therapie aus den anderen erklärenden Variablen (TNM, Alter, Histologie etc.); damit ist die Therapie nicht als unabhängige Variable anzusehen. Ein Einbezug der Therapie als Variable in die Modelle würde aufgrund der Multikollinearität die Aussagekraft verschlechtern.

Mit einem Wert von 0,861 (95% CI: 0,808-0,913) hat das Modell dieser Arbeit im Vergleich zum MSKCC- und dem Li-Nomogramm einen höheren C-Index und weist damit eine bessere diskriminative Fähigkeit auf. Es existieren noch weitere prognostische Modelle, welche sich jedoch nicht auf das Gesamtüberleben fokussieren, sondern auf die Rezidivwahrscheinlichkeit, andere Ereignisse oder sie gelten lediglich für eine bestimmte Histologie (Sun et al. 2019; Carrillo et al. 2007; Lu et al. 2017) und sind daher nicht vergleichbar. Es gibt nur wenige publizierte Modelle zum Abschätzen des Gesamtüberlebens.

4.5 Praktische Anwendbarkeit der Ergebnisse

Die Ergebnisse dieser Arbeit können zum Abschätzen der Prognose von Parotiskarzinompatienten verwendet werden. Das Nomogramm bietet eine Möglichkeit, eine individuelle Prognose abzuschätzen; die CART-Analyse die Möglichkeit, Risikogruppen zu definieren. In den letzten Jahren wurden zahlreiche *Online-Tools* entwickelt, welche Forschern und Ärzten das Anwenden der statistischen Modelle erleichtert (Lee et al. 2016). Exemplarisch sei hier das *Web-Tool* des Sloan Memorial Kettering Cancer Center zur Vorhersage des *Outcomes* von Prostata-Karzinomen vor Prostatektomie (https://www.mskcc.org/nomograms/prostate/pre_op) oder des National Health Service (NHS) für Brustkrebs (<https://breast.predict.nhs.uk/tool>) genannt. 2013 wurden bereits veröffentlichte *Web-Tools* zur Krebsprognose gesichtet und ausgewertet (Rabin et al. 2013). Die Vermittlung von Krebsdiagnose und -prognose benötigt viel Feingefühl und stellt für Ärzte und Patienten eine Herausforderung dar. Daraus ergeben sich zahlreiche ethische Fragen: Für welchen Anwenderkreis sind die *Web-Tools* gedacht? Sollte jeder Mensch Zugang zu diesen Anwendungen erhalten? Was passiert mit einem Patienten, der so seine eigene Prognose errechnet? Können Nicht-Geschulte das Ergebnis im Hinblick auf die Schwächen von statistischen Modellen und Datenauswertung richtig interpretieren?

Die Autoren beklagen, dass es bisher keine gemeinsam formulierten Richtlinien für die Entwicklung und Veröffentlichung von *Web-Tools* gibt (Rabin et al. 2013). Aufgrund unserer ethischen Bedenken und der unter Kapitel 4.7 formulierten statistischen Schwächen dieser Arbeit haben wir auf eine Veröffentlichung des *Web-Tools* verzichtet. 2015 wurde eine Checkliste (TRIPOD) zur transparenten Berichterstattung über Entwicklung und Validierung von prognostischen Modellen veröffentlicht. Ursächlich war gemäß den Autoren die Auswertung von Publikationen von prognostischen Modellen, welche nur unzureichend dokumentiert (z.B. Infos über Einschlusskriterien, Prädiktorenauswahl, Zusammensetzung der Kohorten, Validierung etc.) waren, um eine professionelle Einschätzung der Modellgüte vorzunehmen zu können (Collins et al. 2015). Die Checkliste wurde für diese Arbeit angewandt (siehe „Anhang“).

4.6 Ausblick

Der methodische Überbegriff zu dieser Arbeit lautet „*Machine Learning*“. Darunter werden Methoden verstanden, bei der ein Modell aus Daten „lernt“ und die Erkenntnis auf neue Daten anwendbar macht. CART und die neuronalen Netzwerke sind Methoden, welche sich schon lange etabliert haben. Es ist attraktiv, diese Methoden auf die dokumentierten Daten anzuwenden, um Prognosen abzuschätzen und

Zusammenhänge darzustellen (Kourou et al. 2015). Mit mehr Fallzahlen, besserem Wissen und Diagnostizieren von Prädiktoren (z.B. molekulare Marker) werden auch die prognostischen Modelle zunehmend besser. *Machine Learning*-Verfahren werden vermutlich künftig verstärkt in der Medizin angewandt. Eine PubMed-Recherche offenbart die Vielzahl an veröffentlichten Publikationen hierzu, welche aktuell hauptsächlich Brust- und Prostatakarzinome thematisieren (Jović et al. 2017; Montazeri et al. 2016; Nindrea et al. 2018).

In der Literatur herrscht die Meinung vor, dass die etablierten Prognosemodelle wie TNM/UICC nicht mehr ausreichend sind, da sie lediglich anatomische Prädiktoren (Tumorgroße, Lymphknoten, Fernmetastasen) einbeziehen und andere Prädiktoren wie Alter, Histologie usw. unberücksichtigt lassen. Vielmehr werden Modelle, bei denen man bei Bedarf neue Prädiktoren flexibel hinzufügen oder entfernen kann (z.B. in Form eines Nomogramms) als Zukunft der Prognoseschätzung gesehen (Kreppel et al. 2013). In den letzten Jahren wurden zahlreiche prognostische Modelle für diverse medizinische Fragestellungen veröffentlicht. Die Dokumentation der Entwicklung dieser Modelle ist oftmals mangelhaft, sodass eine Bewertung hinsichtlich der Aussagekraft schwierig ist. Mit dem *TRIPOD-Statement* wurde erstmals eine international anerkannte Leitlinie zur transparenten Berichterstattung veröffentlicht. Frei zugängliche prognostische Modelle im Sinne von Web-Applikationen werden in Zukunft zunehmen. Bereits heute sind zahlreiche Anwendungen zur Überlebensvorhersage für jeden zugänglich. Es bleibt fraglich, welche psychischen Folgen das Nutzen dieser Modelle für ungeschulte Laien wie Patienten oder deren Angehörige hat. Aus dem Klinikalltag ist bekannt, dass die Vermittlung von Tumordiagnose und -prognose ein hohes Maß an Professionalität und Feinfühligkeit verlangt. Von einer unreflektierten Veröffentlichung solcher webbasierten prognostischen Modelle ist daher abzuraten.

4.7 Stärken und Grenzen der Arbeit

Bei dieser Arbeit handelt es sich um eine retrospektive Datenanalyse, da die Datensammlung bereits abgeschlossen war, als die Fragestellung der Untersuchung formuliert wurde. Durch den langen Zeitraum konnte, verglichen mit bereits publizierten Modellen und in Anbetracht der Seltenheit von Parotiskarzinomen, eine relativ große Anzahl an Fällen herangezogen werden. Die Fallanzahl ist für die Modellqualität ausschlaggebend. Der lange Beobachtungszeitraum ist jedoch auch eine Schwäche, da sich während des Beobachtungszeitraums Therapieprotokolle, diagnostische Technik,

und Klassifikationen (WHO, TNM etc.) geändert haben, was zu Unschärfen führen könnte. Eine prognostische Studie ist daher optimaler Weise prospektiv konzipiert. Dabei ist allen Beteiligten die Zielsetzung der Datenerhebung bewusst und die Datenqualität kann dementsprechend kontrolliert werden (Steyerberg 2009, S. 35–36). Bei einer retrospektiven Datengrundlage ist unter anderem die Datenqualität limitierend, z.B. im Hinblick auf das dokumentierte Ereignis (Tod, Todesursache, Rezidiv) bzw. andere Prädiktoren. So konnte in dieser Arbeit nur das Gesamtüberleben verwendet werden. Bei einer höheren Datenqualität wäre z.B. die Verwendung des DFS (*disease-free survival*; Zeitraum ab Ende der Primärtherapie bis zum Wiederauftreten der Krankheit) oder des PFS (*progression-free survival*; Zeitraum nach Therapieende, in dem kein Voranschreiten der Krankheit feststellbar ist) möglich. Ebenfalls konnten nur die „üblichen“ im klinischen Verlauf dokumentierten Prädiktoren ausgewertet werden. Möchte man die Daten hinsichtlich eines ganz bestimmten Prädiktors untersuchen (z.B. einer HPV-Assoziation der Karzinome), könnte es sein, dass nur wenige Fälle zur Verfügung stehen, da diese Information nicht routinemäßig erfasst und dokumentiert wird. Leider haben wir nur eine begrenzte Auswahl an Variablen, die standortübergreifend einheitlich dokumentiert wurden. Mehrere Variablen, welche in der Literatur als statistisch signifikant erkannt wurden (wie z.B. Perineuralinvasion, Biomarker (siehe Tabelle 4), konnten daher nicht einbezogen werden. Es ist anzunehmen, dass sich die Rahmenbedingungen je nach Klinikstandort unterscheiden (andere Therapieprotokolle, diagnostisches Vorgehen, demographische Struktur etc.). Daher erhöht diese Vielfältigkeit die Allgemeingültigkeit des Modells (Steyerberg 2009). Weiterhin waren Datensätze zahlreicher Fälle unvollständig bzw. fehlerhaft dokumentiert und konnten nicht ausgewertet werden. Für *Machine-Learning*-Prozesse gilt, je mehr Fälle verfügbar sind, desto besser kann das Modell „lernen“ (Kourou et al. 2015). Eine grundsätzliche Schwäche beim prognostischen *Modeling* ist, dass die generierten Vorhersagewerte stark davon abhängen, mit welchen Daten das Modell trainiert wird. So wird dieses Modell für Extremfälle (z.B. sehr alte bzw. sehr junge Patienten) ungenauer abschneiden, als für Fälle, welche sich öfter in unserer Kohorte wiederfinden. Diese Schwächen müssen jedoch in Kauf genommen werden, um überhaupt statistische Analysen durchführen zu können, da ansonsten zu wenig Datenmaterial zur Verfügung stehen würde.

Das Aufteilen des Datensatzes in Trainings- und Validierungsdatsatz ist sowohl Stärke als auch Schwäche der Vorgehensweise. Durch den *Datasplit* bleiben 147 Fälle (Validierungsdatsatz) für das Trainieren des Modells ungenutzt. Dadurch wird aber auch eine Validierung möglich, bei der im Gegensatz zu anderen Validierungsmethoden zuvor ungenutzte Daten zur Verfügung stehen. Das Berücksichtigen des Standorts bei

der Zusammensetzung der Stichproben ist ebenfalls sowohl eine Stärke (mehrere geographische Stichproben) als auch Schwäche (ggf. Heterogenität der erhobenen Daten). Aufgrund des Datasplits haben wir außerdem die Möglichkeit, vielfach Zufallsstichproben zu erstellen und daraus Modelltrainings und Validierungen (1000 Cox-Regressionen und 1000 CART-Bäume) durchzuführen, um die Modellstabilität festzustellen.

Mit den Variablen (Alter, T-Stadium, N-Stadium, M-Stadium, Histo) werden lediglich fünf Variablen verwendet, wodurch dem Phänomen der Überanpassung vorgebeugt wird. Überanpassung (*Overfitting*) beschreibt ein Phänomen, welches sich aus der Modellbildung ergibt. Obwohl das Modell regelkonform anhand der vorliegenden Daten entwickelt worden ist, sind seine Vorhersagen für neue Fälle nichtzutreffend. Die Ursachen liegen in der Verwendung von zu vielen erklärenden Variablen. Das entstandene Modell funktioniert dadurch gut für die Trainingsdaten, versagt jedoch bei der Anwendung auf neue Daten. Die vielen Variablen erhöhen zwar die theoretische *Performance* des Modells, erniedrigen aber dessen Allgemeingültigkeit. Überanpassung kann vermieden werden, in dem die Anzahl an Modellparametern möglichst niedrig gehalten wird (Steyerberg 2009, S. 83-90). Mit der Mindestanzahl von Ereignissen pro Ausprägung (z.B: bei den fünf histologischen Ausprägungen) wird die Aussagekraft erhöht (>10 Ereignisse pro Ausprägung)(Steyerberg 2009, S. 198).

5 Zusammenfassung

Ziel der vorliegenden Arbeit war zu untersuchen, welche der im Krankheitsverlauf dokumentierten Variablen einen prognostischen Einfluss auf das Gesamtüberleben von Parotiskarzinompatienten haben. Es wurde ein großes Kollektiv (n=688, aus 9 deutschen Klinikstandorten, 1972-2018) an Daten von Patienten, welche an einem Parotiskarzinom erkrankt sind hinsichtlich prognostischer Faktoren retrospektiv analysiert. Unter Anwendung traditioneller (Cox-Regression) und neuer (*Machine-Learning*-) Methoden wurden Prognosemodelle entwickelt, Risikogruppen identifiziert und in klinisch anwendbare Formen (Nomogramm, „Entscheidungsbaum“, *Web-App*) umgesetzt. Die hierbei angewendeten Methoden sind die Kaplan-Meier-Überlebenszeitanalyse, der Chi2-Test, die Cox-Regression, das Nomogramm und der Klassifikations- und Regressionsbaum (CART; rekursives Partitionieren). In der univariaten Analyse zeigen die Variablen TNM, Histologie, Diagnosealter, Geschlecht und Malignität einen statistisch signifikanten Einfluss auf die Überlebenszeit. Das multivariate Modell (n=340) wurde aus den Variablen TNM, Histologie und Diagnosealter entwickelt und als Nomogramm sowie *Web-App* in eine klinisch anwendbare Form gebracht. In der CART-Analyse (n=340) wurde das Vorhandensein eines Plattenepithelkarzinoms als wichtigster prognostischer Faktor identifiziert, wichtiger als das Diagnosealter. Die histologische Diagnose sollte daher bei der Therapieentscheidung und Prognosestellung maßgebliche Berücksichtigung finden. Die Modelle wurden sowohl Modell-intern, als auch mit Fällen (n=147) validiert, welche nicht in die Modellentwicklung eingeflossen sind. Mit einem C-Index von 0,861 (95% CI: 0,808-0,913) weist das hier vorgestellte Nomogramm eine bessere diskriminative Fähigkeit auf als zwei andere publizierte Modelle. Mit Nomogramm, *Web-App* und CART-Baum wurden drei Möglichkeiten angewandt, komplexe Überlebensmodelle in eine leicht-nutzbare Form zu übertragen, deren Anwendung jedoch Ärzten vorbehalten sein sollte. Auf eine Veröffentlichung der *Web-App* wurde aus ethischen Bedenken im Hinblick auf die eventuelle Nutzung durch Patienten verzichtet.

6 Summary

The topic of the study was to investigate the predictive influence of the variables documented in the clinical course on the overall survival of parotis carcinoma patients. A large dataset (n=688 from 9 german hospitals, 1972-2018) from patients suffering from parotid carcinoma was retrospectively analyzed with regards to prognostic factors and survival analysis. Using traditional (Cox regression) and new (machine learning) methods, prognostic models were developed, risk groups identified and transformed into clinically applicable forms (nomogram, "decision tree", web app). The methods used in this thesis were the Kaplan-Meier survival analysis, the Chi2-test, the Cox regression, the nomogram and the classification and regression tree (CART; recursive partitioning).

In the univariate analysis, the variables TNM, histology, age at diagnosis, sex and malignancy revealed a statistically significant influence on the survival time. The multivariate model (n=340) was developed from the variables TNM, histology, age at diagnosis and was transformed as a nomogram and web app into a clinically applicable form. In the CART analysis (n=340) the presence of squamous cell carcinoma was identified as the most important prognostic factor, even more important than the age at diagnosis. The histological diagnosis should therefore be a key consideration for the therapy decision and prognosis. The models were validated both internally and with cases (n=147) which were not included in the model training. The nomogram presented has a C-index of 0.861 (95% CI: 0.808-0.913), which has a better discriminative ability than two other previously published models. With nomogram, web app and CART tree, three possibilities are used in this thesis to transfer complex survival models into an easy usable form. Due to ethical concerns regarding the inappropriate use by patients, the web app was not published, therefore the application should only be reserved for physicians.

Abkürzungsverzeichnis

ACC	Adenoid-zystisches Karzinom
AcinicCa	Azinuszellkarzinom
AdenoCa.....	Adenokarzinom
AJCC	American Joint Committee on Cancer
BCAC	Basalzelladenokarzinom
BN	Bonn
CaexPA	Karzinom ex pleomorphes Karzinom
CART.....	Classification- and Regression-Tree
CI.....	Konfindenzintervall
CMAP	compound muscle action potential
CT.....	Computer-Tompgrafie
cTNM.....	klinische Tumorklassifikation
DFS	Disease Free Survival
EBV	Epstein-Barr-Virus
EF.....	Erfurt
EMC	Epithelia-myoepitheliales Karzinom
G	Gera
GI.....	Gießen
GTDS	Gießener Tumordokumentationssystem
HIV	Humanes Immundefizienz-Virus
HPV	Humanes Papillomvirus
HR	Hazard Ratio
ICD-O	International Classification of Diseases for Oncology
IDC	Invasiv-duktales Karzinom
J	Jena
K	Köln
KIT.....	Tyrosinkinase KIT
MC.....	Myoepitheliales Karzinom
MEC	Mukoepidermoidkarzinom
MRT.....	Magnet-Resonanz-Tomografie
MSKCC.....	Memorial Sloan Kettering Cancer Center

NDH.....	Nordhausen
NHS.....	National Health System
NOS.....	not otherwise specified
PEC.....	Plattenepithelkarzinom
PET.....	Positronen-Emissions-Tomografie
PFS.....	Progression Free Survival
pTNM.....	pathologische Tumorklassifikation
SCC.....	Squamous cell carcinoma
SEER.....	Surveillance, Epidemiology, and End Results
SHL.....	Suhl
TNM.....	T=Tumor, N=Nodus, M=Metastase
TRIPOD.....	Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis
UICC.....	Union internationale contre le cancer
UKGM.....	Universitätsklinikum Gießen und Marburg
UL.....	Ulm
WHO.....	World Health Organization

Abbildungsverzeichnis

Abbildung 1: Anteile der Klinikstandorte n=487.	15
Abbildung 2: Häufigkeitsverteilung der fünf häufigsten Histologien n=487.	16
Abbildung 3: Kaplan-Meier-Überlebenszeitanalyse für alle Fälle (n=487). Zensierte Fälle durch senkrechte Striche markiert.	22
Abbildung 4: Kaplan-Meier-Überlebenszeitanalyse differenziert nach Standort, n=487 (*p: Log-rank-Test über alle Schichten; paarweise Vergleiche in Tabelle 10), zensierte Fälle durch senkrechte Striche markiert.	22
Abbildung 5: Kaplan-Meier-Überlebenszeitanalyse (p: Log-rank-Test; zensierte Fälle durch senkrechte Striche markiert): a) Geschlecht (n=487), b) Altersgruppe (n=487). 24	
Abbildung 6: Kaplan-Meier-Überlebenszeitanalyse differenziert nach Histologie n=487 (p: Log-rank-Test jeweils zwischen benachbarten Gruppen wie dargestellt), zensierte Fälle durch senkrechte Striche markiert.	25
Abbildung 7: Kaplan-Meier-Überlebenszeitanalyse differenziert nach seltener Histologie n=109	26
Abbildung 8: Kaplan-Meier-Überlebenszeitanalyse (p: Log-rank-Test; zensierte Fälle durch senkrechte Striche markiert): a) T-Stadium (n=487), b) N-Stadium (n=487), c) M-Stadium (n=487), d) UICC-Staging (n=485).....	27
Abbildung 9: Kaplan-Meier-Überlebenszeitanalyse differenziert nach Malignität n=316 (p: Log-rank-Test), zensierte Fälle durch senkrechte Striche markiert.	28
Abbildung 10: Kaplan-Meier-Überlebenszeitanalyse (p: Log-rank-Test, zensierte Fälle durch senkrechte Striche markiert): a) Therapie(n=487), b) radikales OP-Vorgehen (n=413), c) Neck-Dissection (n=431), d) Rezidiv (n=466)	29
Abbildung 11: Nomogramm zur Vorhersage des 5-/10-Jahres-Gesamtüberlebens bei Parotiskarzinomen. Jedem Prädiktor wird dabei eine Punktzahl zugeordnet, die auf der oberen Skala abgelesen wird. Anhand der Gesamtpunktzahl auf der unteren Skala kann das 5- und 10- Jahres-Gesamtüberleben entsprechend abgelesen werden.	35
Abbildung 12: Kalibrierungsdiagramm: Die gestrichelte Linie steht für die perfekte Kalibrierung von beobachtetem und vorhergesagtem Überleben; die durchgezogene Linie steht für die Kalibrierung von beobachtetem und vorhergesagtem Überleben des Nomogramms für die Validierungsdaten.....	36
Abbildung 13: Klassifikations- und Regressions-Baum (CART) (n=340) mit Signifikanzniveau $p=0,05$, die Farben zeigen die Zugehörigkeit zu den hier definierten Risikogruppen an (rot= Hoch-Risiko-Gruppe, gelb= mittlere Risiko-Gruppe, grün= Niedrig-Risiko-Gruppe).....	38
Abbildung 14: Kaplan-Meier-Überlebenszeitanalyse differenziert nach Subgruppen/Endknoten von Training (n=340)	39
Abbildung 15: Kaplan-Meier-Überlebenszeitanalyse differenziert nach erstellten Hauptrisikogruppen aus den Subgruppen von Training (n=340) und log-rank Test	40
Abbildung 16: Kaplan-Meier-Überlebenszeitanalyse für die einzelnen Subgruppen/Endknoten (5-7, 9-11) von Validierung n=147	42
Abbildung 17: Kaplan-Meier-Überlebenszeitanalyse differenziert nach erstellten Hauptrisikogruppen aus den Subgruppen von Validierung n=147 und log-rank-Test(p-Wert)	43

Abbildung 18: Flussdiagramm: Fallzuteilung von Training (n=340) auf die einzelnen Hauptrisikogruppen	44
Abbildung 19: Screenshot der Webanwendung zeigt Fallbeispiel: 60-jähriger Patient mit einem T2, N1, M0 Plattenepithelkarzinom	45

Tabellenverzeichnis

Tabelle 1: TNM-Klassifikation für große Speicheldrüsen (8. Auflage) (Brierley et al. 2017).....	3
Tabelle 2: UICC-Staging für große Speicheldrüsen (8. Auflage) (Brierley et al. 2017). 3	
Tabelle 3: Übersicht der unterschiedlichen WHO-Klassifikationen (Seifert et al. 1990; Barnes et al. 2005; El-Naggar et al. 2017).....	4
Tabelle 4: Übersicht über prognostische Faktoren für die Überlebenszeit von Parotiskarzinomen in der Literatur	6
Tabelle 5: Häufigkeiten der histologischen Entitäten in n=688	14
Tabelle 6: Übersicht über Gesamtkohorte (n=487).	17
Tabelle 7: Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich des Geschlechts.....	18
Tabelle 8: Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich der Altersgruppe.....	19
Tabelle 9: Häufigkeitsverteilung patienten- und tumorbezogener Merkmale bezüglich der Histologie	20
Tabelle 10: P-Werte (Log-Rank-Test; paarweise Vergleiche der Kaplan-Meier-Kurven in Abbildung 4) differenziert nach Standort, Grün: Statistisch signifikant($p < 0,05$), *: Ohne p-Wert, da alle Fälle zensiert.....	23
Tabelle 11: P-Werte (Log-Rank-Test; paarweise Vergleiche der Kaplan-Meier-Kurven in Abbildung 8) differenziert nach häufigsten Histologien, Grün: Statistisch signifikant ($p < 0,05$).....	25
Tabelle 12: P-Werte (Log-Rank-Test; paarweise Vergleiche der Kaplan-Meier-Kurven in Abbildung 9 differenziert nach seltener Histologie, Grün: Statistisch signifikant ($p < 0,05$)	26
Tabelle 13: Univariate Analyse von Einflussfaktoren auf das Gesamtüberleben von Patienten der fünf häufigsten Parotiskarzinome (n=487)	31
Tabelle 14: Übersicht Trainings- und Validierungskohorte (n=487).....	32
Tabelle 15: Multivariate Cox-Regression mit allen univariat-signifikanten Faktoren (n=340).....	33
Tabelle 16: Finale Cox-Regression (n=340)	34
Tabelle 17: Übersicht über Trainings- und Validierungskohorte für die CART-Analyse	37
Tabelle 18: P-Werte (Log-Rank-Test; paarweise Vergleiche der KM der Subgruppen/Endknoten von Training in Abbildung 22), Grün: $P > 0,05$	39
Tabelle 19: Häufigkeit der Variablen die in den 1000 trainierten CART-Modellen an erster Stelle aufgerufen wurden.....	41
Tabelle 20: Fallzusammensetzung: Verteilung der Fälle/Ereignisse auf die einzelnen Subgruppen/Endknoten.....	41
Tabelle 21: P-Werte (Log-Rank-Test; paarweise Vergleiche der KM von Subgruppen/Endknoten von Validierung in Abbildung 24) ,Grün: $P > 0,05$	42

Tabelle 22: Risikogruppentabelle für Parotiskarzinome	44
Tabelle 23: Transparent Reporting of a multivariable rediction model for Individual Prognosis Or Diagnosis- (TRIPOD)-Tabelle	72
Tabelle 24: Häufigkeitsverteilung der Histologie bezüglich patientenbezogener Merkmale	73
Tabelle 25: Häufigkeitsverteilung des T-Stadiums bezüglich patientenbezogener Merkmale	74
Tabelle 26: Häufigkeitsverteilung des N-Stadiums bezüglich patientenbezogener Merkmale	75
Tabelle 27: Häufigkeitsverteilung des M-Stadiums bezüglich patientenbezogener Merkmale	76
Tabelle 28: Häufigkeitsverteilung des UICC-Stage bezüglich patientenbezogener Merkmale	77
Tabelle 29: Häufigkeitsverteilung der Malignität bezüglich patientenbezogener Merkmale	78

Literaturverzeichnis

- Ali, Safina; Bryant, Robert; Palmer, Frank L.; DiLorenzo, Monica; Shah, Jatin P.; Patel, Snehal G.; Ganly, Ian (2015): Distant Metastases in Patients with Carcinoma of the Major Salivary Glands. In: *Annals of Surgical Oncology* 22 (12), S. 4014–4019. DOI: 10.1245/s10434-015-4454-y.
- Ali, Safina; Palmer, Frank L.; Yu, Changhong; DiLorenzo, Monica; Shah, Jatin P.; Kattan, Michael W. et al. (2013): A predictive nomogram for recurrence of carcinoma of the major salivary glands. In: *JAMA otolaryngology-- head & neck surgery* 139 (7), S. 698–705. DOI: 10.1001/jamaoto.2013.3347.
- Ali, Safina; Palmer, Frank L.; Yu, Changhong; DiLorenzo, Monica; Shah, Jatin P.; Kattan, Michael W. et al. (2014): Postoperative nomograms predictive of survival after surgical management of malignant tumors of the major salivary glands. In: *Annals of Surgical Oncology* 21 (2), S. 637–642. DOI: 10.1245/s10434-013-3321-y.
- Barnes, Eveson; Reichart; Sidransky (2005): Pathology and Genetics of Head and Neck Tumours. WHO Classification of Tumours, 3rd Edition, Volume 9.
- Becker, Christoph; Dahlem, Kilian Konrad Kenjiro; Pfeiffer, Jens (2017): Prognostic value of comorbidities in patients with carcinoma of the major salivary glands. In: *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 274 (3), S. 1651–1657. DOI: 10.1007/s00405-016-4404-8.
- Bhattacharyya, Neil; Fried, Marvin P. (2005): Determinants of survival in parotid gland carcinoma: a population-based study. In: *American Journal of Otolaryngology* 26 (1), S. 39–44. DOI: 10.1016/j.amjoto.2004.06.017.
- Bjørndal, Kristine; Krogdahl, Annelise; Therkildsen, Marianne Hamilton; Overgaard, Jens; Johansen, Jørgen; Kristensen, Claus A. et al. (2012): Salivary gland carcinoma in Denmark 1990-2005: outcome and prognostic factors. Results of the Danish Head and Neck Cancer Group (DAHANCA). In: *Oral oncology* 48 (2), S. 179–185. DOI: 10.1016/j.oraloncology.2011.09.005.
- Bjørndal, Kristine; Larsen, Stine R.; Therkildsen, Marianne H.; Kristensen, Claus A.; Charabi, Birgitte; Andersen, Elo et al. (2016): Does age affect prognosis in salivary gland carcinoma patients? A national Danish study. In: *Acta Oncologica* 55 (sup1), S. 19–22. DOI: 10.3109/0284186X.2015.1114671.
- Breiman, Leo (1984): Classification and regression trees. New York: Chapman & Hall/CRC.
- Brierley, James; Gospodarowicz, Mary K.; Wittekind, Christian (Hg.) (2017): TNM classification of malignant tumours. Eighth edition. Chichester, West Sussex, UK, Hoboken, NJ: John Wiley & Sons Inc.
- Calearo, Carlo; Storchi, Omero Fini; Pasture, Antonio; Polli, Gianni (1998): Parotid Gland Carcinoma: Analysis of Prognostic Factors. In: *Annals of Otolaryngology, Rhinology & Laryngology* 107 (11), S. 969–973. DOI: 10.1177/000348949810701112.
- Carlson, Julie; Licitra, Lisa; Locati, Laura; Raben, David; Persson, Fredrik; Stenman, Göran (2013): Salivary gland cancer: an update on present and emerging therapies. In: *American Society of Clinical Oncology educational book. American Society of Clinical Oncology. Annual Meeting*, S. 257–263. DOI: 10.1200/EdBook_AM.2013.33.257.
- Carrillo, José F.; Vázquez, Rafael; Ramírez-Ortega, Margarita C.; Cano, Ana; Ochoa-Carrillo, Francisco J.; Oñate-Ocaña, Luis F. (2007): Multivariate prediction of the probability of recurrence in patients with carcinoma of the parotid gland. In: *Cancer* 109 (10), S. 2043–2051. DOI: 10.1002/cncr.22647.
- Cederblad, Lena; Johansson, Silvia; Enblad, Gunilla; Engström, Mats; Blomquist, Erik (2009): Cancer of the parotid gland; long-term follow-up. A single centre experience on recurrence and survival. In: *Acta oncologica (Stockholm, Sweden)* 48 (4), S. 549–555. DOI: 10.1080/02841860802680419.

- Chakrabarti, Swagnik; Nair, Deepa; Malik, Akshat; Qayyumi, Burhanuddin; Nair, Sudhir; Agrawal, Jai Prakash; Chaturvedi, Pankaj (2018): Prognostic factors in parotid cancers: Clinicopathological and treatment factors influencing outcomes. In: *Indian journal of cancer* 55 (1), S. 98–104. DOI: 10.4103/ijc.IJC_503_17.
- Chang; Cheng; Allaire (2019): Web Application Framework for R. Package 'shiny' (Version 1.4.0), zuletzt geprüft am 03.11.2019.
- Cheung, Michael C.; Franzmann, Elizabeth; Sola, Juan E.; Pincus, David J.; Koniaris, Leonidas G. (2011): A comprehensive analysis of parotid and salivary gland cancer: worse outcomes for male gender. In: *The Journal of surgical research* 171 (1), S. 151–158. DOI: 10.1016/j.jss.2009.11.721.
- Chou, Wen-Chi; Chang, Kai-Ping; Lu, Chang-Hsien; Chen, Miao-Fen; Cheng, Yu-Fan; Yeh, Kun-Yun et al. (2017): Complementary role of the Memorial Sloan Kettering Cancer Center nomogram to the American Joint Committee on Cancer system for the prediction of relapse of major salivary gland carcinoma after surgery. In: *Head & neck* 39 (5), S. 860–867. DOI: 10.1002/hed.24702.
- Collins, Gary S.; Reitsma, Johannes B.; Altman, Douglas G.; Moons, Karel G. M. (2015): Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. In: *Annals of internal medicine* 162 (1), S. 55–63. DOI: 10.7326/M14-0697.
- Edge, Stephen B. (2010): AJCC cancer staging manual. 7th ed. New York: Springer.
- Ellwanger, Ulf (2001): Anwendung künstlicher neuronaler Netze für die Prognoseschätzung bei Verlaufsdaten: FU Berlin.
- El-Naggar, Adel K.; Chan, John K. C.; Grandis, Jennifer R.; Takata, Takashi; Slootweg, Pieter J. (Hg.) (2017): WHO classification of head and neck tumours. 4th edition. Lyon: International Agency for Research on Cancer (World Health Organization classification of tumours, 9 (in the 4th edition)).
- Erovic, Boban M.; Shah, Manish D.; Bruch, Guillem; Johnston, Meredith; Kim, John; O'Sullivan, Brian et al. (2015): Outcome analysis of 215 patients with parotid gland tumors: a retrospective cohort analysis. In: *Journal of Otolaryngology - Head & Neck Surgery* 44 (1), S. 43. DOI: 10.1186/s40463-015-0097-z.
- Ettl, Tobias; Gosau, Martin; Brockhoff, Gero; Schwarz-Furlan, Stephan; Agaimy, Abbas; Reichert, Torsten E. et al. (2014): Predictors of cervical lymph node metastasis in salivary gland cancer. In: *Head & neck* 36 (4), S. 517–523. DOI: 10.1002/hed.23332.
- Ettl, Tobias; Schwarz-Furlan, Stephan; Gosau, Martin; Reichert, Torsten E. (2012): Salivary gland carcinomas. In: *Oral and maxillofacial surgery* 16 (3), S. 267–283. DOI: 10.1007/s10006-012-0350-9.
- Freling, Nicole; Crippa, Flavio; Maroldi, Roberto (2016): Staging and follow-up of high-grade malignant salivary gland tumours: The role of traditional versus functional imaging approaches - A review. In: *Oral oncology* 60, S. 157–166. DOI: 10.1016/j.oraloncology.2016.04.016.
- Guntinas-Lichius, Orlando; Silver, Carl E.; Thielker, Jovanna; Bernal-Sprekelsen, Manuel; Bradford, Carol R.; Bree, Remco de et al. (2018): Management of the facial nerve in parotid cancer: preservation or resection and reconstruction. In: *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 275 (11), S. 2615–2626. DOI: 10.1007/s00405-018-5154-6.
- Hantschel, Oliver; Rix, Uwe; Superti-Furga, Giulio (2008): Target spectrum of the BCR-ABL inhibitors imatinib, nilotinib and dasatinib. In: *Leukemia & lymphoma* 49 (4), S. 615–619. DOI: 10.1080/10428190801896103.
- Harbo, G.; Bundgaard, T.; Pedersen, D.; Sogaard, H.; Overgaard, J. (2002): Prognostic indicators for malignant tumours of the parotid gland. In: *Clin Otolaryngol* 27 (6), S. 512–516. DOI: 10.1046/j.1365-2273.2002.00625.x.
- Harrell, Frank E. (2015): Regression Modeling Strategies. Cham: Springer International Publishing.

- Harrell Jr., Frank E. (2019): Regression Modeling Strategies. Package 'rms' (Version 5.1-3.1), zuletzt geprüft am 03.11.2019.
- Hay, Ashley; Migliacci, Jocelyn; Zanoni, Daniella Karassawa; Patel, Snehal; Yu, Changhong; Kattan, Michael W.; Ganly, Ian (2018): Validation of nomograms for overall survival, cancer-specific survival, and recurrence in carcinoma of the major salivary glands. In: *Head & neck* 40 (5), S. 1008–1015. DOI: 10.1002/hed.25079.
- Hothorn, Torsten; Hornik, Kurt; Strobl, Carolin; Zeileis, Achim (2019a): A Laboratory for Recursive Partytioning. Package 'party' (Version 1.3-3), zuletzt geprüft am 03.11.2019.
- Hothorn, Torsten; Hornik, Kurt; Zeileis, Achim (2006): Unbiased Recursive Partitioning: A Conditional Inference Framework. In: *Journal of Computational and Graphical Statistics* 15 (3), S. 651–674. DOI: 10.1198/106186006X133933.
- Hothorn, Torsten; Seibold, Heidi; Zeileis, Achim (2019b): A Toolkit for Recursive Partytioning. Package 'partykit' (Version 1.2-5), zuletzt geprüft am 03.11.2019.
- Ihrler, S.; Guntinas-Lichius, O.; Haas, C.; Mollenhauer, M. (2018): Neues zu Tumoren der Speicheldrüsen : WHO-Klassifikation 2017. In: *Der Pathologe* 39 (1), S. 11–17. DOI: 10.1007/s00292-017-0407-5.
- Israel, Yair; Rachmiel, Adi; Gourevich, Konstantin; Nagler, Rafael (2019): Survival Probabilities Related to Histology, Grade and Stage in Patients With Salivary Gland Tumors. In: *Anticancer research* 39 (2), S. 641–647. DOI: 10.21873/anticancer.13158.
- Jang, Jeon Yeob; Choi, Nayeon; Ko, Young-Hyeh; Chung, Man Ki; Son, Young-Ik; Baek, Chung-Hwan et al. (2018): Treatment outcomes in metastatic and localized high-grade salivary gland cancer: high chance of cure with surgery and post-operative radiation in T1–2 N0 high-grade salivary gland cancer. In: *BMC cancer* 18. DOI: 10.1186/s12885-018-4578-0.
- Johnston, Meredith L.; Huang, Shao Hui; Waldron, John N.; Atenafu, Eshetu G.; Chan, Kelvin; Cummings, Bernard J. et al. (2016): Salivary duct carcinoma: Treatment, outcomes, and patterns of failure. In: *Head & neck* 38 Suppl 1, E820-6. DOI: 10.1002/hed.24107.
- Jović, Srđan; Miljković, Milica; Ivanović, Miljan; Šaranović, Milena; Arsić, Milena (2017): Prostate Cancer Probability Prediction By Machine Learning Technique. In: *Cancer investigation* 35 (10), S. 647–651. DOI: 10.1080/07357907.2017.1406496.
- Kane, W. J.; McCaffrey, T. V.; Olsen, K. D.; Lewis, J. E. (1991): Primary Parotid Malignancies: A Clinical and Pathologic Review. In: *Archives of Otolaryngology–Head & Neck Surgery* 117 (3), S. 307–315. DOI: 10.1001/archotol.1991.01870150075010.
- Kattan, Michael W. (2003): Comparison of Cox regression with other methods for determining prediction models and nomograms. In: *The Journal of urology* 170 (6 Pt 2), S6-9; discussion S10. DOI: 10.1097/01.ju.0000094764.56269.2d.
- Keller, Gunter; Steinmann, Diana; Quaas, Alexander; Grünwald, Viktor; Janssen, Stefan; Hussein, Kais (2017): New concepts of personalized therapy in salivary gland carcinomas. In: *Oral oncology* 68, S. 103–113. DOI: 10.1016/j.oraloncology.2017.02.018.
- Kourou, Konstantina; Exarchos, Themis P.; Exarchos, Konstantinos P.; Karamouzis, Michalis V.; Fotiadis, Dimitrios I. (2015): Machine learning applications in cancer prognosis and prediction. In: *Computational and structural biotechnology journal* 13, S. 8–17. DOI: 10.1016/j.csbj.2014.11.005.
- Kreppel, Matthias; Scheer, Martin; Meyer, Moritz; Stenner, Markus; Wedemeyer, Inga; Drebber, Uta et al. (2013): Comparison of TNM-based stage grouping versus UICC/AJCC stage grouping (7th edition) in malignant parotid gland tumors. In: *Oral oncology* 49 (9), S. 903–910. DOI: 10.1016/j.oraloncology.2013.06.009.
- Lee, Yong Ho; Bang, Heejung; Kim, Dae Jung (2016): How to Establish Clinical Prediction Models. In: *Endocrinology and metabolism (Seoul, Korea)* 31 (1), S. 38–44. DOI: 10.3803/EnM.2016.31.1.38.
- Li, Yun; Ju, Jun; Liu, Xiaoxiao; Gao, Tao; Wang, Zhidong; Ni, Qianwei et al. (2017): Nomograms for predicting long-term overall survival and cancer-specific survival in patients with

major salivary gland cancer: a population-based study. In: *Oncotarget* 8 (15), S. 24469–24482. DOI: 10.18632/oncotarget.14905.

Lima, Roberto A.; Tavares, Marcos R.; Dias, Fernando L.; Kligerman, Jacob; Nascimento, Marilene F.; Barbosa, Mauro M. et al. (2005): Clinical prognostic factors in malignant parotid gland tumors. In: *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 133 (5), S. 702–708. DOI: 10.1016/j.otohns.2005.08.001.

Lu, Chang-Hsien; Liu, Chien-Ting; Chang, Pei-Hung; Hung, Chia-Yen; Li, Shau-Hsuan; Yeh, Ta-Sen et al. (2017): Develop and validation a nomogram to predict the recurrent probability in patients with major salivary gland cancer. In: *Journal of Cancer* 8 (12), S. 2247–2255. DOI: 10.7150/jca.19461.

Masconi, Katya L.; Matsha, Tandi E.; Echouffo-Tcheugui, Justin B.; Erasmus, Rajiv T.; Kengne, Andre P. (2015): Reporting and handling of missing data in predictive research for prevalent undiagnosed type 2 diabetes mellitus: a systematic review. In: *The EPMA journal* 6 (1), S. 7. DOI: 10.1186/s13167-015-0028-0.

Montazeri, Mitra; Montazeri, Mohadeseh; Montazeri, Mahdieh; Beigzadeh, Amin (2016): Machine learning models in breast cancer survival prediction. In: *Technology and health care : official journal of the European Society for Engineering and Medicine* 24 (1), S. 31–42. DOI: 10.3233/THC-151071.

National Comprehensive Cancer Network (2019): Clinical Practice Guidelines in Oncology. Head and Neck Cancers (Version 1.2018), zuletzt aktualisiert am 02.11.2019.

Nindrea, Ricvan Dana; Aryandono, Teguh; Lazuardi, Lutfan; Dwiprahasto, Iwan (2018): Diagnostic Accuracy of Different Machine Learning Algorithms for Breast Cancer Risk Calculation: a Meta-Analysis. In: *Asian Pacific journal of cancer prevention : APJCP* 19 (7), S. 1747–1752. DOI: 10.22034/APJCP.2018.19.7.1747.

Paderno, Alberto; Tomasoni, Michele; Mattavelli, Davide; Battocchio, Simonetta; Lombardi, Davide; Nicolai, Piero (2018): Primary parotid carcinoma: analysis of risk factors and validation of a prognostic index. In: *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery*. DOI: 10.1007/s00405-018-5122-1.

Pan, Sai Yi; Groh, Margaret de; Morrison, Howard (2017): A Case-Control Study of Risk Factors for Salivary Gland Cancer in Canada. In: *Journal of cancer epidemiology* 2017, S. 4909214. DOI: 10.1155/2017/4909214.

Pohar, Surjeet; Gay, Hiram; Rosenbaum, Paula; Klish, Darren; Bogart, Jeffrey; Sagerman, Robert et al. (2005): Malignant parotid tumors: presentation, clinical/pathologic prognostic factors, and treatment outcomes. In: *International journal of radiation oncology, biology, physics* 61 (1), S. 112–118. DOI: 10.1016/j.ijrobp.2004.04.052.

Poorten, Vincent Vander; Hart, Augustinus; Vauterin, Tom; Jeunen, Gert; Schoenaers, Joseph; Hamoir, Marc et al. (2009): Prognostic index for patients with parotid carcinoma: international external validation in a Belgian-German database. In: *Cancer* 115 (3), S. 540–550. DOI: 10.1002/cncr.24015.

Rabin, Borsika A.; Gaglio, Bridget; Sanders, Tristan; Nekhlyudov, Larissa; Dearing, James W.; Bull, Sheana et al. (2013): Predicting cancer prognosis using interactive online tools: a systematic review and implications for cancer care providers. In: *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 22 (10), S. 1645–1656. DOI: 10.1158/1055-9965.EPI-13-0513.

Rayess, Hani M.; Dezube, Aaron; Bawab, Ibrahim; Raza, S. Naweed; Yoo, George H.; Lin, Ho-Sheng; Jacobs, John R. (2017): Tumor Differentiation as a Prognostic Factor for Major Salivary Gland Malignancies. In: *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 157 (3), S. 454–461. DOI: 10.1177/0194599817700593.

Reiß, Michael (2009): Facharztwissen HNO-Heilkunde. Differenzierte Diagnostik und Therapie. Berlin, Heidelberg: Springer Berlin Heidelberg. Online verfügbar unter <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10330996>.

Royston, Patrick; Altman, Douglas G. (2013): External validation of a Cox prognostic model: principles and methods. In: *BMC medical research methodology* 13, S. 33. DOI: 10.1186/1471-2288-13-33.

Seethala, Raja R. (2009): An update on grading of salivary gland carcinomas. In: *Head and neck pathology* 3 (1), S. 69–77. DOI: 10.1007/s12105-009-0102-9.

Seethala, Raja R.; Altemani, Albina; Ferris, Robert L.; Fonseca, Isabel; Gnepp, Douglas R.; Ha, Patrick et al. (2019): Data Set for the Reporting of Carcinomas of the Major Salivary Glands: Explanations and Recommendations of the Guidelines From the International Collaboration on Cancer Reporting. In: *Archives of pathology & laboratory medicine* 143 (5), S. 578–586. DOI: 10.5858/arpa.2018-0422-SA.

Seifert, G.; Brocheriou, C.; Cardesa, A.; Eveson, J. W. (1990): WHO International Histological Classification of Tumours Tentative Histological Classification of Salivary Gland Tumours. In: *Pathology - Research and Practice* 186 (5), S. 555–581. DOI: 10.1016/S0344-0338(11)80220-7.

Shariat, Shahrokh F.; Karakiewicz, Pierre I.; Suardi, Nazareno; Kattan, Michael W. (2008): Comparison of nomograms with other methods for predicting outcomes in prostate cancer: a critical analysis of the literature. In: *Clinical cancer research : an official journal of the American Association for Cancer Research* 14 (14), S. 4400–4407. DOI: 10.1158/1078-0432.CCR-07-4713.

Spiro, R. H. (1986): Salivary neoplasms: overview of a 35-year experience with 2,807 patients. In: *Head & neck surgery* 8 (3), S. 177–184.

Steyerberg, Ewout W. (2009): Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating. New York, NY: Springer-Verlag New York (Statistics for Biology and Health). Online verfügbar unter <http://dx.doi.org/10.1007/978-0-387-77244-8>.

Su, Yu-Xiong; Roberts, Dianna B.; Hanna, Ehab Y.; El-Naggar, Adel; Saylam, Güleser; Frank, Steven J. et al. (2015): Risk Factors and Prognosis for Myoepithelial Carcinoma of the Major Salivary Glands. In: *Annals of Surgical Oncology* 22 (11), S. 3701–3707. DOI: 10.1245/s10434-015-4389-3.

Sun, Jian; Sun, Yang; Yang, Fei; Zhou, Qianrong; Liu, Wenjuan; Cheng, Yong et al. (2019): Nomogram to predict the prognosis of parotid gland mucoepidermoid carcinoma: a population-based study of 1306 cases. In: *PeerJ* 7 (1), e7237. DOI: 10.7717/peerj.7237.

Terhaard, Chris H. J.; Lubsen, H.; van der Tweel, I.; Hilgers, F. J. M.; Eijkenboom, W. M. H.; Marres, H. A. M. et al. (2004): Salivary gland carcinoma: independent prognostic factors for locoregional control, distant metastases, and overall survival: results of the Dutch head and neck oncology cooperative group. In: *Head Neck* 26 (8), 681-92; discussion 692-3. DOI: 10.1002/hed.10400.

Therneau; Terry M; Atkinson, Beth; Ripley, Brian: Recursive Partitioning and Regression Trees. Package 'rpart' (Version 4.1-15) 2019, zuletzt geprüft am 03.11.2019.

Therneau, Terry M.; Lumley, Thomas (2019): Survival Analysis. Package 'survival' (Version 2.44-1.1).

Vander Poorten, Vincent L. M.; Balm, Alfonsus J. M.; Hilgers, Frans J. M.; Tan, I. Bing; Loftus-Coll, Barbara M.; Keus, Ronald B. et al. (1999): The development of a prognostic score for patients with parotid carcinoma. In: *Cancer* 85 (9), S. 2057–2067. DOI: 10.1002/(SICI)1097-0142(19990501)85:9<2057::AID-CNCR24>3.0.CO;2-F.

Vander Poorten, Vincent L. M.; Hart, Augustinus A. M.; van der Laan, Bernardus F. A. M.; Baatenburg de Jong, Robert J.; Manni, Johannes J.; Marres, Henri A. M. et al. (2003): Prognostic index for patients with parotid carcinoma: external validation using the nationwide 1985-1994 Dutch Head and Neck Oncology Cooperative Group database. In: *Cancer* 97 (6), S. 1453–1463. DOI: 10.1002/cncr.11254.

- Wagner, Steffen; Wittekindt, Claus; Sharma, Shachi Jenny; Wuerdemann, Nora; Jüttner, Theresa; Reuschenbach, Miriam et al. (2017): Human papillomavirus association is the most important predictor for surgically treated patients with oropharyngeal cancer. In: *Br J Cancer* 116, 1604 EP -. DOI: 10.1038/bjc.2017.132.
- Wells, Brian J.; Chagin, Kevin M.; Nowacki, Amy S.; Kattan, Michael W. (2013): Strategies for handling missing data in electronic health record derived data. In: *EGEMS (Washington, DC)* 1 (3), S. 1035. DOI: 10.13063/2327-9214.1035.
- Xiao, Christopher C.; Baker, Andrew B.; White-Gilbertson, Shai J.; Day, Terry A. (2016): Prognostic Factors in Myoepithelial Carcinoma of the Major Salivary Glands. In: *Otolaryngology-head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 154 (6), S. 1047–1053. DOI: 10.1177/0194599816637605.
- Zhan, Kevin Y.; Lentsch, Eric J. (2016): Basal cell adenocarcinoma of the major salivary glands: A population-level study of 509 cases. In: *The Laryngoscope* 126 (5), S. 1086–1090. DOI: 10.1002/lary.25713.
- Zhao, Jianqiang; Wang, Jiafeng; Yu, Chang; Guo, Liang; WANG, KEJING; Liang, Zhong; Lou, Jianlin (2013): Prognostic factors affecting the clinical outcome of carcinoma ex pleomorphic adenoma in the major salivary gland. In: *World journal of surgical oncology* 11 (1), S. 180. DOI: 10.1186/1477-7819-11-180.

Anhang

Der Anhang enthält zusätzliche Informationen und Auswertungen, welche im Zusammenhang mit der Arbeit stehen.

Anhang A: TRIPOD-Tabelle

Tabelle 23: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis-(TRIPOD)-Tabelle

Section/Topic		Checklist Item		Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	Front Page
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	59 (Summary)
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	1-8 / 52-53
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	8
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	9-13/31/37
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	9-10/14-17
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	9
	5b	D;V	Describe eligibility criteria for participants.	10
	5c	D;V	Give details of treatments received, if relevant.	17
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	10
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	/
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	9-10
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	/
Sample size	8	D;V	Explain how the study size was arrived at.	9-10
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	11
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	11-13
	10c	V	For validation, describe how the predictions were calculated.	12-13/35-36/40-43
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	36/40-43
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	/
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	38-44
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	32/37
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	22
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	9-10/17
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	17/32
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	14-45
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	21-31

Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	34-35/38
	15b	D	Explain how to use the prediction model.	7-8/ 35/38
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	34
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	/
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	55-57
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	52-53
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	46-55
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	54-57
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	/
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	/

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

Anhang B: Häufigkeitsverteilungen der patientenbezogenen Merkmale zueinander

Histologie

Tabelle 24: Häufigkeitsverteilung der Histologie bezüglich patientenbezogener Merkmale

Histologie		AdenoCa		AcinicCa		ACC		MEC		PEC		p-Wert
		n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	
Geschlecht	W	39	33,3	66	68,0	53	57,6	55	53,9	25	31,6	<0,001
	M	78	66,7	31	32,0	39	42,4	47	46,1	54	68,4	
Alter (Median:62,8)	<Median	45	38,5	53	54,6	57	62,0	73	71,6	15	19,0	<0,001
	>Median	72	61,5	44	45,4	35	38,0	29	28,4	64	81,0	
T-Stage	T1/2	53	45,3	65	67,0	42	45,7	76	74,5	39	49,4	<0,001
	T3/4	64	54,7	32	33,0	50	54,3	26	25,5	40	50,6	
N-Stage	N0	55	47,0	71	73,2	71	77,2	73	71,6	36	45,6	<0,001
	≥N1	62	53,0	26	26,8	21	22,8	29	28,4	43	54,4	
M-Stage	M0	101	86,3	93	95,9	84	91,3	91	89,2	73	92,4	0,175
	M1	16,0	14	4,0	4	8,0	9	11,0	11	6,0	7,6	
UICC	I/II	36,0	31	54,0	56	35,0	38	61,0	60	19,0	24,4	<0,001
	III/IV	80	69	43	44,3	57	62	41	40,2	59	75,6	
Malignität	Low Grade	40	44,9	35	66,0	23	47,9	53	70,7	23	45,1	0,002
	High Grade	49	55,1	18	34,0	25	52,1	22	29,3	28	54,9	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: Statistisch signifikant ($p \leq 0,05$)

Adenokarzinom-Patienten sind eher männlich (66,7%), älter (61,5%) und haben ein höheres TNM-Stadium (T3/4: 54,7%, N1: 53%, M1: 12,8%), UICC-Stadium (III/IV: 68,1%) sowie eher *high-grade* Tumoren (56,2%).

Bei Azinuszellkarzinom-Patienten ist besonders die Dominanz des weiblichen Geschlechts auffällig (68%). Sie sind ebenfalls jünger (54,6%), haben ein niedrigeres

TNM-Stadium (T1/2: 67%, N0: 73,2%, M0: 95,9%), UICC-Stadium (I/II: 56%) und haben eher *low-grade* Tumoren (66%).

Patienten mit einem adenoid-zystischen Karzinom sind eher weiblich (57,6%), jünger (62%), haben ein höheres T-Stadium (T3/4: 54,3%), ein niedrigeres N-Stadium (N0: 77,2%), ein niedrigeres M-Stadium (M0: 9%), dahingegen aber ein höheres UICC-Stadium (III/IV: 62%) und eher *high-grade* Tumoren (52,1%).

Mukoepidermoidkarzinom-Patienten sind eher weiblich (53,9%) und jünger (71,6%). Auffallend ist, dass die Tumoren meist klein sind (T1/2: 74,5%) Sie haben außerdem ein niedrigeres N- und M-Stadium (N0: 71,6%, M0: 89,2%), UICC-Stadium (I/II: 60%) und eher *low-grade* Tumoren (70,7%).

Plattenepithelkarzinom-Patienten dahingegen sind eher männlich (68,4%) und haben ein höheres TNM-Stadium (T3/4: 50,6%, N1: 54,4%, M1: 7,6%), UICC-Stadium (III/IV: 75,6%) sowie eher *high-grade* Tumoren (54,9%). Auffallend ist, dass PEC-Patienten häufig der höheren Altersgruppe angehören (81%).

Statistisch signifikante Unterschiede ergaben sich für die Häufigkeitsverteilung der Variablen Geschlecht, Alter, T- und N-Stadium, UICC-*Staging* und Malignität.

T-Stadium

Tabelle 25: Häufigkeitsverteilung des T-Stadiums bezüglich patientenbezogener Merkmale

T-Stadium		T1/2		T3/4		p-Wert
		n	(%)	n	(%)	
Geschlecht	W	150	54,5	88	41,5	0,004
	M	125	45,5	124	58,5	
Alter (Median:62,8)	<Median	147	53,5	96	45,3	0,074
	>Median	128	46,5	116	54,7	
Histologie	AdenoCa	53	19,3	64	30,2	<0,001
	AcinicCa	65	23,6	32	15,1	
	ACC	42	15,3	50	23,6	
	MEC	76	27,6	26	12,3	
	PEC	39	14,2	40	18,9	
N	N0	213	77,5	93	43,9	<0,001
	≥N1	62	22,5	119	56,1	
M	M0	255	92,7	187	88,2	0,088
	M1	20	7,3	25	11,8	
UICC	I/II	205	74,5	0	0,0	<0,001
	III/IV	70	25,5	210	100,0	
Malignität	Low Grade	107	63,7	67	45,3	0,001
	High Grade	61	36,3	81	54,7	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: Statistisch signifikant ($p \leq 0,05$)

Patienten mit kleineren Tumoren (T1/2) sind im Vergleich zu größeren Tumoren (T3/4) eher weiblich (54,5% vs. 41,5%), und z.B. öfter vom Mukoepidermoidkarzinom betroffen

(27,6% vs. 12,3%) aber seltener vom Plattenepithelkarzinom (14,2% vs. 18,9%) Sie haben ebenfalls ein niedrigeres N-Stadium (77,5% vs. 43,9%) sowie UICC-*Staging* (74,9% vs. 0%). T3/4-Tumoren können per Definition kein UICC-*Stage* I/II sein, dadurch erklären sich die 0%. Kleinere Tumoren sind eher low-grade (63,1%) als größere (45,3%). Die Häufigkeitsverteilung ist für die Variablen Geschlecht, Histologie, N-Stadium, UICC-*Staging* und Malignität statistisch signifikant.

N-Stadium

Tabelle 26: Häufigkeitsverteilung des N-Stadiums bezüglich patientenbezogener Merkmale

N-Stadium		N0		≥N1		p-Wert
		n	(%)	n	(%)	
Geschlecht	W	170	55,6	68	37,6	<0,001
	M	136	44,4	113	62,4	
Alter (Median:62,8)	<Median	171	55,9	72	39,8	0,001
	>Median	135	44,1	109	60,2	
Histologie	AdenoCa	55	18,0	62	34,3	<0,001
	AcinicCa	71	23,2	26	14,4	
	ACC	71	23,2	21	11,6	
	MEC	73	23,9	29	16,0	
	PEC	36	11,8	43	23,8	
T	T1/2	213	69,6	62	34,3	<0,001
	T3/4	93	30,4	119	65,7	
M	M0	291	95,1	151	83,4	<0,001
	M1	15	4,9	30	16,6	
UICC	I/II	205	67,0	0	0,0	<0,001
	III/IV	101	33,0	179	100,0	
Malignität	Low Grade	127	67,9	47	36,4	<0,001
	High Grade	60	32,1	82	63,6	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: Statistisch signifikant ($p \leq 0,05$)

Patienten mit Lymphknotenmetastasen ($\geq N1$) sind im Vergleich zu Patienten ohne Lymphknotenmetastasen (N0) eher männlich (62,4% vs. 44%), älter (60,2% vs. 44,1%), öfter z.B. vom Plattenepithelkarzinom betroffen (23,8% vs. 11,8%) und haben ein höheres T- und M-Stadium (T3/4: 65,7% vs. 30,4%, M1: 16,6% vs. 4,6%) sowie UICC-*Staging* (III/IV: 100% vs. 32,7%). $\geq N1$ -Tumoren sind eher *high-grade* Tumoren (63,6% vs. 32,6%). Auffällig ist, dass die Häufigkeitsverteilung für alle einbezogenen Variablen statistisch signifikant ist.

M-Stadium

Tabelle 27: Häufigkeitsverteilung des M-Stadiums bezüglich patientenbezogener Merkmale

M-Stadium		M0		≥M1		p-Wert
		n	(%)	n	(%)	
Geschlecht	W	223	50,5	15	33,3	0,029
	M	219	49,5	30	66,7	
Alter (Median:62,8)	<Median	223	50,5	20	44,4	0,443
	>Median	219	49,5	25	55,6	
Histologie	AdenoCa	101	22,9	16	35,6	0,175
	AcinicCa	93	21,0	4	8,9	
	ACC	84	19,0	8	17,8	
	MEC	91	20,6	11	24,4	
	PEC	73	16,5	6	13,3	
T	T1/2	255	57,7	20	44,4	0,088
	T3/4	187	42,3	25	55,6	
N	N0	291	65,8	15	33,3	<0,001
	≥N1	151	34,2	30	66,7	
UICC	I/II	205	46,6	0	0,0	<0,001
	III/IV	235	53,4	45	100,0	
Malignität	Low Grade	167	58,4	7	23,3	<0,001
	High Grade	119	41,6	23	76,7	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: Statistisch signifikant ($p \leq 0,05$)

Patienten mit Fernmetastasen (M1) sind im Vergleich zu Patienten mit M0 eher männlich (65,9% vs. 49,7%) und haben ein höheres N-Stadium ($\geq N1$: 68,2% vs. 34,1%), sowie UICC-*Staging* (III/IV: 100% vs. 53,3%). Die Häufigkeitsverteilung hinsichtlich des M-Stadiums ist für die Variablen Geschlecht, N-Stadium, UICC-*Staging* und Malignität statistisch signifikant.

UICC-Stage

Tabelle 28: Häufigkeitsverteilung des UICC-Stage bezüglich patientenbezogener Merkmale

UICC-Stage		I/II		III/IV		p-Wert
		n	(%)	n	(%)	
Geschlecht	W	119	58,0	118	42,1	0,001
	M	86	42,0	162	57,9	
Alter (Median:62,8)	<Median	118	57,6	125	44,6	0,005
	>Median	87	42,4	155	55,4	
Histologie	AdenoCa	36	17,6	80	28,6	<0,001
	AcinicCa	54	26,3	43	15,4	
	ACC	35	17,1	57	20,4	
	MEC	61	29,8	41	14,6	
	PEC	19	9,3	59	21,1	
T	T1/2	205	100,0	70	25,0	<0,001
	T3/4	0	0,0	210	75,0	
N	N0	205	100,0	101	36,1	<0,001
	≥N1	0	0,0	179	63,9	
M	M0	205	100,0	235	83,9	<0,001
	M1	0	0,0	45	16,1	
Malignität	Low Grade	88	73,3	84	43,3	<0,001
	High Grade	32	26,7	110	56,7	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: Statistisch signifikant ($p \leq 0,05$)

Patienten mit höherem UICC-Stage (III/IV) sind im Vergleich zu Patienten mit Stage I/II eher männlich (57,7% vs. 42,2%), haben ein höheres T-Stadium (T3/4: 75,3% vs. 0%), ein höheres N-Stadium ($\geq N1$: 64,2% vs. 0%) und ein höheres M-Stage (M1: 15,8% vs. 0%). Patienten mit einem niedrigen UICC-Stage(I/II) können per Definition nicht T3/4, $\geq N1$ oder M1 sein, weswegen sich die 0% bei der Verteilung erklären. Patienten mit einem höheren UICC-Stage sind eher von *high-grade* Tumoren (57% vs. 27,3%) betroffen. Auffällig ist die statistische Signifikanz aller Variablen bei der Häufigkeitsverteilung.

Malignität

Tabelle 29: Häufigkeitsverteilung der Malignität bezüglich patientenbezogener Merkmale

Malignität		Low Grade		High Grade		p-Wert
		n	(%)	n	(%)	
Geschlecht	W	88	50,6	58	40,8	0,084
	M	86	49,4	84	59,2	
Alter (Median:62,8)	<Median	92	52,9	61	43	0,079
	>Median	82	47,1	81	57	
Histologie	AdenoCa	40	23,0	49	34,5	0,002
	AcinicCa	35	20,1	18	12,7	
	ACC	23	13,2	25	17,6	
	MEC	53	30,5	22	15,5	
T	PEC	23	13,2	28	19,7	0,001
	T1/2	107	61,5	61	43,0	
	T3/4	67	38,5	81	57,0	
N	N0	127	73,0	60	42,3	<0,001
	≥N1	47	27,0	82	57,7	
M	M0	167	96,0	119	83,8	<0,001
	M1	7	4,0	23	16,2	
UICC	I/II	88	51,2	32	22,5	<0,001
	III/IV	84	48,8	110	77,5	

p-Wert: Chi2-Test, Signifikanzniveau = 0,05; **fett**: Statistisch signifikant ($p \leq 0,05$)

Patienten mit *high-grade* Tumoren sind im Vergleich zu Patienten mit *low-grade* Tumoren seltener z.B. am Mukoepidermoidkarzinom erkrankt (15,4% gegenüber 30,6%), haben ein höheres T-Stadium (T3/4: 56,6% vs. 38,7%), ein höheres N-Stadium (≥N1: 57,3% vs. 27,2%) und ein höheres M-Stage (M1: 16,1% vs. 3,5%). Besonders markant ist, dass 77,5% aller an *high-grade*-diagnostizierten Patienten ein vorangeschrittenes UICC-*Staging* (III/IV) aufweisen. Die Häufigkeitsverteilung hinsichtlich der Malignität ist für die Variablen Histologie, TNM und UICC-*Staging* statistisch signifikant.

Publikationsverzeichnis

Teile der Dissertation wurden veröffentlicht:

Wittekindt C. (2019) Outcome nach Therapie von Parotis-Malignomen in Deutschland, Vortrag bei „90. Jahresversammlung der Deutschen Gesellschaft für Hals-Nasen-Ohren-Heilkunde, Kopf- und Hals-Chirurgie e.V.“, Berlin, 29.05.-01.06.2019

Erklärung zur Dissertation

„Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne unzulässige Hilfe oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nichtveröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten sowie ethische, datenschutzrechtliche und tierschutzrechtliche Grundsätze befolgt. Ich versichere, dass Dritte von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen, oder habe diese nachstehendspezifiziert. Die vorgelegte Arbeit wurde weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt. Alles aus anderen Quellen und von anderen Personen übernommene Material, das in der Arbeit verwendet wurde oder auf das direkt Bezug genommen wird, wurde als solches kenntlich gemacht. Insbesondere wurden alle Personen genannt, die direkt und indirekt an der Entstehung der vorliegenden Arbeit beteiligt waren. Mit der Überprüfung meiner Arbeit durch eine Plagiatserkennungssoftware bzw. ein internetbasiertes Softwareprogramm erkläre ich mich einverstanden.“

Ort, Datum _____ Unterschrift _____

Danksagung

Zunächst möchte ich Prof. Dr. med. Claus Wittekindt, meinem Doktorvater, für die Bereitstellung des Themas, Unterstützung, Kritik, Ideen und die Motivation danken.

Einen besonderen Dank möchte ich an Dr. rer. nat. Steffen Wagner richten für die hervorragende Betreuung meiner Arbeit, das zeitaufwändige Korrekturlesen und die schnellen, detaillierten Antworten bei Rückfragen.

Für die sachkundige Unterstützung bei Fragen zur statistischen Auswertung danke ich Dr. Jörn Pons-Kühnemann, Leiter der Medizinischen Statistik im Institut für Medizinische Informatik der Justus-Liebig-Universität Gießen sowie seinem wissenschaftlichen Mitarbeiter Jonas Tschammer.

Ebenfalls möchte ich Dr. Julia Koch danken, welche sich bereits bei einem vorangegangenen Promotionsprojekt mit der Datenanalyse der malignen Speicheldrüsentumoren beschäftigt hat und auf deren Erkenntnisse ich aufbauen konnte.

Nicht zuletzt danke ich meiner Familie und meinen Freunden, die mir all das bisher Erlebte und Erreichte ermöglicht haben und die mich während der Durchführung der Arbeit persönlich unterstützt und motiviert haben.