Aus dem Institut für Pathologie (Univ-Prof. Dr. med. P. Ströbel) der Medizinischen Fakultät der Universität Göttingen

Dignitätsbeurteilung

## von gastrointestinalen Stromatumoren

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades der Medizinischen Fakultät der Georg-August-Universität Göttingen

VORGELEGT VON

DR. RER. NAT. BJÖRN SANDER

AUS

**BAD OEYNHAUSEN** 

GÖTTINGEN 2020

Dekan:Prof. Dr. med. W. BrückReferent/in:Prof. Dr. med. L. FüzesiKo-Referent/in:PD Dr. med. A. O. König

Datum der mündlichen Prüfung: 03.08.2021

Hiermit erkläre ich, die Dissertation mit dem Titel "Dignitätsbeurteilung von gastrointestinalen Stromatumoren" eigenständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Die Daten, auf denen die vorliegende Arbeit basiert, wurden teilweise publiziert:

**Sander B**, Cameron S, Gunawan B, Fuzesi L (2020): Optimal thresholds of risk parameters for gastrointestinal stromal tumors. Eur J Surg Oncol <u>46</u>, 180-188

Gunawan B, von Heydebreck A, **Sander B**, Schulten HJ, Haller F, Langer C, Armbrust T, Bollmann M, Gasparov S, Kovac D et al. (2007): An oncogenetic tree model in gastrointestinal stromal tumours (GISTs) identifies different pathways of cytogenetic evolution with prognostic implications. J Pathol <u>211</u>, 463-470

## Inhaltsverzeichnis

Abbil	ldungsverzeichnisIII
Tabe	llenverzeichnis IV
Abkü	rzungsverzeichnisV
1 E	inleitung1
1.1	Epidemiologie und Lokalisation von GIST1
1.2	Zelluläre Differenzierung beim GIST2
1.3	Makroskopie und Histologie beim GIST2
1.4	Genetische Alterationen3
1.5	Pädiatrische GIST4
1.6	Klassifikationssysteme zur prognostischen Einschätzung6
1.6.1	AFIP-Klassifikation und Gruppen nach Miettinen6
1.6.2	NIH-Klassifikation nach Fletcher7
1.6.3	Aktualisierte AFIP- und NIH-Klassifikationen7
1.6.4	TNM-Klassifikation8
1.7	Therapie von GIST9
1.7.1	Verlaufskontrollen versus Resektion9
1.7.2	Zielgerichtete pharmakologische Therapie10
1.8	Zytogenetische Veränderungen11
1.9	Fragestellung12
2 N	Nethoden13
2.1	Statistische Analysen13
2.1.1	Kaplan-Meyer-Kurven13
2.1.2	Ereigniszeitanalysen13

2.1.	.3	Coxsches Regressionsmodel (Cox proportional hazards model)	13
2.1.	.4	Maximal ausgewählte Rankstatistik (maximally selected rank statistics)	15
2.1.	.5	Risikoklassifikationsbäume (risk classification trees)	15
2.1.	.6	Umfassende rezidivfreie Überlebensdifferenzanalyse mit Bootstrapping	16
2.2	(	Genehmigung durch die Ethikkommission	16
3	Er	gebnisse	17
3.1	(	Göttinger GIST-Kollektiv	17
3.1.	.1	Vergleich der Risikoklassifikationssysteme	18
3.1.	.2	Analysen für den numerischen Risikoparameter Tumorgröße	23
3.1.	.3	Analysen für den numerischen Risikoparameter Mitoserate	25
3.1.	.4	Tumorgröße und Mitoserate in der Risikobewertung von GIST	27
3.1.	.5	Anzahl chromosomaler Nettoveränderungen als numerischer Risikoparameter	30
4	Di	skussion	35
4.1	9	Schwellenwerte für Tumorgröße und Mitoserate	35
4.2	I	Bewertung der Risikoschemata	36
4.3	1	Anzahl chromosomaler Nettoveränderungen als prognostischer Parameter	38
5	Zu	ısammenfassung	40
6	Lit	teraturverzeichnis	41

# Abbildungsverzeichnis

Abbildung 1: Risikobewertungssysteme für GIST	. 5
Abbildung 2: Allgemeine Charakteristika des GIST-Kollektivs	19
Abbildung 3: Vergleichende rezidivfreie Überlebensdifferenzanalysen der untersuchten	
Klassifikationsschemata	20
Abbildung 4: Risikoverhältnis als Funktion der Risikoklasse für die verschiedenen Risiko-	
klassifizierungsschemata	21
Abbildung 5: Schwellenwert-Analysen für die Tumorgröße	24
Abbildung 6: Schwellenwert-Analysen für die Mitoserate	26
Abbildung 7: Tumorgröße und Mitoserate als Risikoparameter	28
Abbildung 8: Anzahl der klonalen Nettoveränderungen als Risikoparameter.	34

## Tabellenverzeichnis

Tabelle 1: Einheitliche numerische Risikoklasse in der statistischen Auswertung14
Tabelle 2: Fallzahlen nach Risikoniveau der untersuchten Klassifikationssysteme
Tabelle 3: Multivariate Analyse mittels Cox-proportional-hazards-Modellen für die
Risikoklassifikationen und das Risiko Lokalisation22
Tabelle 4: Anzahl chromosomaler Nettoveränderungen nach Risikoniveau der
untersuchten Klassifikationssysteme
Tabelle 5: Multivariate Analyse mittels Cox-proportional-hazards-Modell für die
Risikofaktoren Größe, Mitoserate, Lokalisation und Anzahl chromosomaler
Nettoveränderungen32
Tabelle 6: Multivariate Analysen der Risikoniveau der untersuchten Risikoklassifikationen
nach Lokalisation und Anzahl chromosomaler Nettoveränderungen

# Abkürzungsverzeichnis

ABL1	Abelson murine leukemia viral oncogene homolog 1
AFIP	Armed Forces Institute of Pathology
АТР	Adenosintriphosphat
BCR	breakpoint cluster region
BRAF	v-raf murine sarcoma viral oncogene homolog B1
CGH	comparative genomic hybridization
CML	chronische myeloische Leukämie
cnv	copy number variation
coeff	coefficient
DNA	deoxyribonucleic acid
DOG1	discovered on GIST 1
ESMO	European Society for Medical Oncology
ETV6	E26 transformation-specific variant gene 6
ехр	exponential
GIST	Gastrointestinaler Stromatumor
hpf	high power field
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
MAP	mitogen activated protein
NF1	Neurofibromin 1
NIH	National Institutes of Health
NTRK3	neurotrophic receptor tyrosine kinase 3
PDGFRA	platelet-derived growth factor receptor, alpha polypeptide
RAS	rat sarcoma viral oncogene homolog
RFÜ	rezidivfreies Überleben
ROC	receiver operating characteristic
se	standard error
SCF	stem cell factor
SDH	Succinatdehydrogenase
TNM	tumor, nodes, metastases
UICC	Union for International Cancer Control

### 1 Einleitung

Gastrointestinale Stromatumoren (GIST) sind weitgehend im Gastrointestinaltrakt vorkommende mesenchymale Tumoren mit einem interindividuell stark variierenden Risiko einer Metastasierung (Miettinen et al. 1999; Nishida und Hirota 2000). Die Bewertung des Malignitätspotentials ist maßgeblich für die Therapie, da beim Vorliegen von charakteristischen Mutationen im *KIT-* oder *PDGFRA-*Gen eine erhöhte Sensitivität für eine Therapie mit Tyrosinkinaseinhibitoren und somit die Möglichkeit einer zielgerichteten Therapie besteht (Joensuu et al. 2001; Tuveson et al. 2001).

In der Bewertung des Malignitätspotentials kommt der Pathologie eine Schlüsselrolle zu. Für die Graduierung von GIST sind spezielle Klassifikationssysteme basierend auf pathologischen Parametern entwickelt und validiert worden (Fletcher et al. 2002; Franquemont 1995; Joensuu et al. 2012; Miettinen et al. 2002), während die bei anderen Sarkomen übliche Graduierung (Costa et al. 1984; Trojani et al. 1984) beim GIST nicht angewendet wird. Die in der Klinik verwendeten Risikobewertungssysteme bewerten Tumorgröße und Mitoserate sowie teils auch die Tumorlokalisation als prognostische Parameter (Fletcher et al. 2002; Miettinen et al. 2002). Ein Klassifikationssystem unter Berücksichtigung dieser Parameter ist in die TNM-Klassifikation der *Union for International Cancer Control* (UICC) aufgenommen worden (Brierley et al. 2017).

#### 1.1 Epidemiologie und Lokalisation von GIST

GIST sind in der Mehrzahl sporadische, bei Männern leicht häufiger als bei Frauen vorkommende Tumoren mit einem medianen Erkrankungsalter von über 60 Jahren und einer Inzidenz von 1-2 Fällen pro 100.000 Einwohner (Rubin et al. 2011; Tran et al. 2005). Selten kommen GIST bei Kindern vor, die sich biologisch deutlich von GIST bei Erwachsenen unterscheiden (Kaemmer et al. 2009). Mehr als die Hälfte der GIST werden im Magen diagnostiziert, ca. ein Viertel im Dünndarm und 4-5 % im Colon oder Rektum. Die seltenen extragastrointestinalen Stromatumoren kommen an verschiedenen Lokalisationen wie dem abdominalen oder retroperitonealen Bindegewebe vor (Reith et al. 2000; Soreide et al. 2016). Sogenannte Micro-GIST von weniger als 1 cm Durchmesser kommen bei bis zu

ca. 30 % der älteren Bevölkerung im oberen Gastrointestinaltrakt vor und haben im allgemeinen kein erhöhtes Malignitätspotential, selbst wenn sich in den kleinen Tumoren die spezifischen Mutationen im *KIT* oder *PDGFRA* Gen nachweisen lassen (Agaimy et al. 2007; Kawanowa et al. 2006). Für GIST aus Lokalisationen des oberen Gastrointestinaltraktes wird insgesamt ein besseres rezidivfreies Überleben als für GIST aus dem unteren Gastrointestinaltrakt beschrieben, wobei für Ösophagus-GIST aufgrund der Seltenheit dieser Lokalisation geringere Fallzahlen als für Magen- und intestinale GIST untersucht wurden (Joensuu et al. 2012) und diese Tumoren in der aktuellen TNM-Klassifikation zusammen mit Dünndarm-GIST und GIST des unteren Gastrointestinaltraktes als nicht-gastrische GIST eingeordnet werden (Brierley et al. 2017).

#### 1.2 Zelluläre Differenzierung beim GIST

Interstitielle Cajal-Zellen, die Schrittmacherzellen der glatten Muskulatur im Magen-Darm-Trakt, werden als Progenitorzellen von GIST diskutiert (Sircar et al. 1999). Cajal-Zellen wurden in der Maus charakterisiert und exprimieren die Rezeptortyrosinkinase Kit, die eine Rolle in der Entwicklung und Differenzierung der Zellen spielt (Maeda et al. 1992). GIST zeigen in der Mehrzahl der Fälle eine immunhistochemische Expression von KIT. KIT (mast/stem cell growth factor receptor) wird physiologisch durch den Liganden SCF (stem cell factor) aktiviert und leitet verschiedene Signalwege wie den RAS-MAP-Kinase-Weg ein (Lennartsson und Ronnstrand 2012). Goldstandard für die immunhistochemische Verifikation der Diagnose ist mittlerweile der Marker DOG1 (Anoctamin-1) (Espinosa et al. 2008; West et al. 2004), ein Kalzium-aktivierter Chloridkanal mit ebenfalls funktioneller Relevanz in interstitiellen Cajal-Zellen (Schroeder et al. 2008). Für DOG1 wurde insbesondere in Kollektiven mit selteneren Varianten von GIST und unabhängig vom **Mutationsstatus** eine erhöhte diagnostische Sensitivität gegenüber der immunhistochemischen Färbung von KIT gezeigt, so dass immunhistochemisch KITnegative GIST in der DOG1-Färbung häufig erfasst werden können (Liegl et al. 2009).

#### 1.3 Makroskopie und Histologie beim GIST

GIST sind unterhalb der Schleimhaut des Gastrointestinaltraktes lokalisierte, relativ scharf begrenzte Tumoren, die endoskopisch die überliegende Mucosa vorwölben und histologisch eine Proliferation aus verwirbelten oder faszikulär angeordneten Spindelzellverbänden zeigen (Shirai et al. 2001). Histologisch epitheloide Wachstumsmuster kommen gehäuft in Zusammenhang mit Mutationen im *PDGFRA*-Gen und einer Lokalisation im oberen Gastrointestinaltrakt vor (Agaimy et al. 2013; Graadt van Roggen et al. 2001).

#### 1.4 Genetische Alterationen

In sporadischen GIST erwachsener Patienten ist das KIT-Gen zu ca. 80 %, das PDGFRA Gen zu ca. 10 % mutiert (Lasota und Miettinen 2006). Typische Mutationen im *KIT*-Gen treten häufig im Exon 11 (Kodierung der Juxtamembrandomäne), sowie im Exon 9 (extrazelluläre Domäne), seltener in den Exons 13-14 (ATP-Bindedomäne der Tyrosinkinasedomäne) und 17-18 (Aktivierungsschleife der Tyrosinkinasedomäne), auf (Debiec-Rychter et al. 2006; Kinoshita et al. 2003). Im Exon 11 mutierte GIST zeigen typischerweise eine gute Sensitivität für Imatinib, während bei Mutationen im Exon 9 aufgrund einer geringeren Sensitivität eine Dosissteigerung von 400 mg auf 800 mg Imatinib empfohlen wird (Debiec-Rychter et al. 2006). KIT-Mutationen in der Tyrosinkinasedomäne sind dagegen typisch Imatinibresistent und treten de novo sowie als sekundäre resistenzvermittelnde Mutationen in vorbehandelten GIST auf, darunter auch die sekundäre T670I-Mutation (sog. gatekeeper Mutation), die den Zugang der Inhibitormoleküle an der ATP-Bindestelle verhindert (Tamborini et al. 2004). Die typische Mutation im PDGFRA Gen, D842V im Exon 18 des Gens (Medeiros et al. 2004), führt zu einer Imatinibresistenz (Hirota et al. 2003). Seltener werden Mutationen in den Exons 12 und 14 des PDGFRA-Gens nachgewiesen (Medeiros et al. 2004). Die bei GIST vorkommenden genetischen Alterationen sind in der ganz überwiegenden Mehrzahl der Fälle bei Erwachsenen somatisch erworben. Aktivierende KIT-Mutationen der Keimbahn, die unter anderem mit autosomal dominant vererbten familiären GIST assoziiert sind, sind sehr selten (Carballo et al. 2005; Li et al. 2005).

In einem Teil der Doppelwildtyp-GIST (d. h. GIST ohne Nachweis von Mutationen in *KIT* oder *PDGFRA*) kommen als Hinweis auf eine Bedeutung des *mitogen activated protein* (MAP) Kinase Signalwegs bei GIST *BRAF* V600E-Mutationen vor (Agaram et al. 2008). Im Zusammenhang mit dem MAP-Kinase-Weg sind auch Mutationen im Neurofibromin 1

(durch das *NF1* Gen kodiert) einzuordnen, einem Interaktionspartner und Regulator von RAS-Proteinen. Die im Rahmen von Neurofibromatose Typ 1 auftretenden Mutationen können auch zu einem erhöhten Risiko des Auftretens von GIST führen (Cheng et al. 2004; Kinoshita et al. 2004).

#### 1.5 Pädiatrische GIST

Pädiatrische GIST unterscheiden sich klinisch und molekularpathologisch von GIST bei Erwachsenen deutlich, darunter eine Prädominanz von Mädchen und ein Vorkommen von Lymphknotenmetastasen, die bei Erwachsenen untypisch sind (Quiroz et al. 2018). Neben den seltenen GIST-Fällen im Rahmen einer Neurofibromatose Typ 1 wird bei pädiatrischen GIST ein Verlust in Untereinheiten der Succinatdehydrogenase (SDH) bei fehlendem Nachweis der bei erwachsenen Patienten typischen Mutationen im *KIT*- oder *PDGFRA*-Gen beobachtet. Bei der Carney Triade (gastrische GIST, Paragangliome, pulmonale Chondrome) (Carney et al. 1977) handelt es sich in der Regel um ein nicht-erbliches Syndrom mit epigenetischen Veränderungen in den *SDH*-Genen (Settas et al. 2018), während beim Carney-Stratakis-Syndrom bzw. der Carney-Dyade (GIST und Paragangliome) pathogene Keimbahnvarianten in *SDH*-Genen nachgewiesen werden (McWhinney et al. 2007).



Abbildung 1: Risikobewertungssysteme für GIST

(A) AFIP-Gruppen nach Miettinen et al. 2002. (B) AFIP-Klassifikation nach Miettinen für gastrische GIST. (C) AFIP-Klassifikation nach Miettinen et al. für nicht-gastrische GIST. (D) Modifizierte AFIP(A/B)-Klassifikation für gastrische GIST. Die Einordnung von Magen-GIST zwischen 5 und 10 cm mit hoher Mitoserate von über 10 pro 50 hpf als *"very high"* im AFIP(B)-Schema ist als *"*(B)" markiert. (E) Modifizierte AFIP(A/B)-Klassifikation für nicht-gastrische GIST. Die Einordnung von nicht-gastrischen GIST zwischen 5 und 10 cm und mit einer hohen Mitoserate als *"very high"* im AFIP(B)-Schema ist als *"*(B)" markiert. (F) NIH-Klassifikation nach Fletcher et al. 2002. (G) Modifizierte NIH-Klassifikation. (H) TNM-Stadien für Magen-GIST; (I) TNM-Stadien für nicht-gastrische GIST. Die Abbildung entspricht der Abbildung 1 der begleitenden Publikation (Sander et al. 2020) mit freundlicher Genehmigung des Elsevier Verlages.

#### 1.6 Klassifikationssysteme zur prognostischen Einschätzung

Nach der ersten Risikobewertung von GIST durch Franquemont (Franquemont 1995) werden klinisch heute Klassifikationssysteme verwendet, die von Miettinen et al. am Armed Forces Institute of Pathology (AFIP) (Miettinen et al. 2002) sowie von Fletcher et al. am National Institute of Health (NIH) (Fletcher et al. 2002) entwickelt wurden (Abbildung 1). Hauptunterschied zwischen der Miettinen- und der Fletcher-Klassifikation ist die der Tumorlokalisation in der Miettinen-Klassifikation. Berücksichtigung Eine Risikobewertung nach Tumorlokalisation erfolgt auch in der aktuellen TNM-Klassifikation (Brierley et al. 2017). In einer an einer Referenzkohorte validierten Meta-Analyse zeigte die AFIP-Klassifikation in Grenzwertoptimierungskurven (receiver operating characteristic, ROC) ein gering besseres Verhältnis von Sensitivität und Spezifität als die NIH-Klassifikation (Joensuu et al. 2012). In der AFIP-Risikobewertung werden auch kleine GIST des unteren Gastrointestaltraktes mit erhöhter Mitoserate insbesondere des Colons und Rektums bereits als Hochrisikotumoren eingeordnet, die in der NIH-Klassifikation einem intermediären Risiko zugeordnet werden. Gemeinsam ist den Systemen, dass nach makroskopischer Begutachtung der Tumorgröße sowie mikroskopischer Beurteilung der Mitoserate an Hämatoxylin- und Eosin-gefärbten Schnittpräparaten eine Einordnung möglich ist. Die Bestimmung des Mutationsstatus aus Tumor-DNA ist Voraussetzung für eine bei hohem Rezidivrisiko indizierte spezifische Therapie (Casali et al. 2018). Für einzelne Alterationen, insbesondere im häufig beobachteten KIT Exon 11, wurde zwar ein prognostischer Einfluß auf das rezidivfreie Überleben beschrieben, genetische Alterationen werden jedoch derzeit nicht für die Vorhersage des Malignitätsrisikos berücksichtigt (Joensuu 2008).

#### 1.6.1 AFIP-Klassifikation und Gruppen nach Miettinen

Die AFIP-Klassifikation basiert auf einer Einteilung in acht Gruppen, die vier Risikoklassen zugeordnet sind (Miettinen et al. 2002; Miettinen et al. 2006). Zur Bestimmung der Gruppe werden die Tumoren zunächst ohne Unterscheidung nach Lokalisation entsprechend ihrer Größe und Mitoserate eingeteilt, wobei GIST mit niedriger Proliferationsaktivität (< 5 Mitosen pro Gesichtsfeld in der 40x-Vergrößerung; engl. *high power field*, hpf) je nach Größe den Gruppen 1-3b und GIST mit hoher Proliferationsaktivität (≥ 5 Mitosen pro hpf) je nach Größe den Gruppen 4-6b zugeordnet werden (Abbildung 1A); somit sind auch

kleine proliferationsaktive GIST mindestens in die Gruppe 4 einzuordnen. Die Gruppe bestimmt dann lokalisationsabhängig die Dignitätsbeurteilung mit einer Stratifizierung in vier Risikoklassen, die heute übereinstimmend mit der NIH/Fletcher-Klassifikation als *"very low"*, *"low"*, *"intermediate"* und *"high"* bezeichnet werden. Die Bezeichnung der niedrigsten Risikoklasse als *"probably benign"*, wie zunächst vorgeschlagen (Miettinen et al. 2002), ist nicht empfohlen, da für alle GIST ein zumindest niedriges Malignitätspotential angenommen wird.

Für gastrische GIST entspricht die Gruppe 1 dem Risiko *"very low"*, die Gruppen 2, 3a und 4 werden als *"low"* klassifiziert, die Gruppen 3b und 5 als *"intermediate"*, und die Gruppen 6a und 6b als *"high"*. Für nicht-gastrische GIST wird die Gruppe 1 als *"very low"* bewertet, die Gruppe 2 als *"low*, die Gruppe 3a als *"intermediate"* und die Gruppen 3b bis 6 als *"high"*, d. h. in der AFIP-Klassifikation fallen kleine, nicht-gastrische Tumoren mit erhöhter Mitoserate in die Klasse mit hohem Risiko (Abbildung 1B, C). Zur konsistenten Bestimmung der Mitoserate bewerteten Fläche unabhängig vom Gesichtsfeld des verwendeten Mikroskops werden statt der ursprünglich definierten 50 hpf heute einheitlich 5 mm<sup>2</sup> Tumorgewebe bewertet (Agaimy 2010; Brierley et al. 2017).

#### 1.6.2 NIH-Klassifikation nach Fletcher

Nach Fletcher wird das Risiko von Tumoren von weniger als 2 cm Größe und mit weniger als 5 Mitosen pro 50 hpf als *"very low"* klassifiziert, Tumoren von 2-5 cm und und mit weniger als 5 Mitosen pro 50 hpf werden in die Risikoklasse *"low"* eingeordnet, Tumoren von weniger als 5 cm und mit 5-10 Mitosen pro 50 hpf oder von 5-10 cm und mit weniger als 5 Mitosen pro 50 hpf in die Klasse *"intermediate"*, und Tumoren von 5 oder mehr cm Größe und mit 5 oder mehr Mitosen pro 50 hpf oder Tumoren von größer 10 cm oder Tumoren mit mehr als 10 Mitosen pro 50 hpf werden als *"high"* bewertet (Fletcher et al. 2002) (Abbildung 1F).

#### 1.6.3 Aktualisierte AFIP- und NIH-Klassifikationen

Sowohl die AFIP-, als auch die NIH-Klassifikation wurden in den letzten Jahren aktualisiert im Hinblick auf eine bessere Trennung in den Hochrisikogruppen. In der modifizierten NIH-Klassifikation (Huang et al. 2007) wird der 2 cm-Schwellenwert für die Tumorgröße nicht weiter verwendet, so dass die Risikoniveaus *"very low"* und *"low"* in der neue Kategorie I zusammengefasst werden. Kategorie II entspricht dem Risiko *"intermediate"* in der ursprünglichen Klassifikation. Das alte Risiko *"high"* wird in neue Kategorien III und IV aufgeteilt, wobei III die alten Kategorie *"high"* beinhaltet und die Kategorie IV Tumoren beinhaltet, die bei einer Größe über 5 cm mehr als 10 Mitosen pro 50 hpf aufweisen bzw. bei einer Größe über 10 cm mehr als 5 Mitosen pro 50 hpf.

Als Weiterentwicklung des ursprünglichen AFIP-Schemas wurden zwei modifizierte Schemata vorgeschlagen, AFIP(A) und AFIP(B) (Goh et al. 2008). Für die beiden modifizierten AFIP-Schemata werden die acht Risikogruppen nach Miettinen in vier neue Risikoniveaus eingeteilt, wobei ähnlich zum aktualisierten NIH Schema in beiden Schemata die Kategorien *"very low"* und *"low"* zusammengefasst werden. Die Hochrisikogruppe wird in eine *"high"* sowie eine *"very high"* Kategorie unterteilt, wobei die Risikokategorie *"very high"* des AFIP(B) Schemas der Kategorie IV des modifizierten NIH-Schemas entspricht. Im AFIP(A) Schema werden Tumoren mit hoher Mitoserate von über 10 pro 50 hpf und einer Größe von 5-10 cm dagegen der Risikokategorie *"high"* zugeordnet. Wie im ursprünglichen AFIP-Schema werden kleine gastrische GIST unter 2 cm in den modifizierten AFIP-Schemata als *"low"* und gastrische GIST zwischen 2-5 cm als *"intermediate"* bewertet, selbst wenn sich eine erhöhte Mitoserate von über 5 pro 50 hpf darstellt; für nicht-gastrische Lokalisationen werden dagegen auch kleine Tumoren bei erhöhter Mitoserate als *"high"* eingestuft wie bereits im ursprünglichen AFIP-Schema (Abbildung 1D, E).

#### 1.6.4 TNM-Klassifikation

Die TNM-Klassifikation der UICC definiert ähnlich AFIP-Schemata den lokalisationsabhängige Stadien (Abbildung 1H, I) (Brierley et al. 2017). Die Stadieneinteilung für Dünndarm-GIST gilt auch für GIST des Ösophagus, Kolon, Rektum und des Mesenteriums, die hier als nicht-gastrische GIST zusammengefasst werden. Gastrische GIST bis 10 cm und mit niedriger Mitoserate bis 5 pro 50 hpf werden als Stadium IA (bis 5 cm Größe) bzw. IB (größer als 5 cm, nicht größer als 10 cm) eingeordnet. Gastrische GIST bis 5 cm Größe mit erhöhter Mitoserate (mehr als 5 pro 50 hpf) werden dem Stadium II zugeordnet. Gastrische GIST größer als 10 cm mit niedriger Mitoserate bis 5 pro 50 hpf werden in das Stadium IIIA eingeordnet, ebenso Tumoren zwischen 5 und 10 cm Größe mit

erhöhter Mitoserate. Gastrische Tumoren über 10 cm und mit erhöhter Mitoserate über 5 pro 50 hpf werden dem Stadium IIIB zugeordnet. Dünndarm-GIST (nicht-gastrische GIST) bis 5 cm Größe und mit niedriger Mitoserate werden als Stadium I bewertet, zwischen 5 und 10 cm Größe als Stadium II und über 10 cm Größe als Stadium IIIA. Kleine nichtgastrische GIST bis 2 cm mit erhöhter Mitoserate > 5 werden als Stadium IIIA bewertet, alle nicht-gastrischen Tumoren über 2 cm mit erhöhter Mitoserate als Stadium IIIB (Abbildung 1H, I). Die Stadieneinteilung für gastrische GIST gilt für Magen und primäre solitäre GIST des Omentum.

#### 1.7 Therapie von GIST

Die Behandlung von GIST erfolgt nach den 2018 veröffentlichten Richtlinien der Europäischen Gesellschaft für Medizinische Onkologie (engl. *European Society for Medical Oncology*; ESMO) (Casali et al. 2018).

#### 1.7.1 Verlaufskontrollen versus Resektion

Das Standardvorgehen bei Tumoren des oberen Gastrointestinaltraktes < 2 cm ist zunächst die weitere Beobachtung, da die meisten kleinen GIST aus dem oberen Gastrointestinaltrakt ein niedriges Progressrisiko aufweisen. Für kleine rektale GIST wird, soweit unter Brücksichtigung von Co-Morbiditäten und Alter des Patienten möglich, dagegen die Exzision empfohlen, da das Risiko eines Progresses hier als generell höher anzusehen ist (Casali et al. 2018). Für GIST ab 2 cm Größe ist das Standardvorgehen die Biopsie gefolgt von der Exzision. Bei metastasiertem GIST ist eine bioptische Gewinnung von Tumormaterial zwecks molekularer Analyse und ggf. Planung einer spezifischen Therapie anzustreben. Für lokal begrenzte GIST ist die RO-Exzision des Tumors das Standardvorgehen gefolgt von einer adjuvanten Imatinib-Therapie für Imatinib-sensitive GIST mit hohem Rezidivrisiko. In individuellen Fällen kann eine R1-Resektion akzeptiert werden (z. B. lokal nicht vollständig exzidierbare Niedrigrisiko-Tumoren oder zur Reduktion des Tumorvolumens bei fortgeschrittener Erkrankung) (Casali et al. 2018).

#### 1.7.2 Zielgerichtete pharmakologische Therapie

2001 publizierten Joensuu et al. den Fall einer zum Erkrankungszeitpunkt 50-jährigen Patientin mit fortgeschrittenem und metastasiertem sowie Chemotherapie-resistentem GIST (Joensuu et al. 2001). Nach experimenteller Therapie mit dem Tyrosinkinaseinhibitor Imatinibmesylat (STI571) zeigte sich eine dramatische Reduktion der Tumormasse und somit die Vision neuer, zielgerichteter Therapieoptionen mittels spezifischen niedermolekularen (engl. small molecule) Substanzen bei soliden Tumoren. Imatinib war zuvor für Patienten mit chronischer myeloischer Leukämie (CML) als Inhibitor der Tyrosinkinaseaktivitiät des konstitutiv aktiven Fusionsgens BCR-ABL1 bei CML nach umfangreicher Suche eines niedermolekularen Inhibitors etabliert worden (Druker et al. 2001a; Druker et al. 2001b). Imatinib stellt heute die Erstlinientherapie bei GIST dar, bei primärer oder sekundärer Resistenz erfolgt die Zweitlinientherapie mit dem Tyrosinkinaseinhibitor Sunitinib (Demetri et al. 2006), wobei die Imatinib-resistente PDGFRA D842V Mutation ebenfalls nicht mit einer erhöhten Sensitivität gegen Sunitinib verbunden ist (Corless et al. 2005). Als Drittlinientherapie, d. h. bei Resistenz gegen Imatinib und Sunitinib, wird der Tyrosinkinase-Inhibitor Regorafenib verwendet (Demetri et al. 2013). Für Imatinib-sensitive GIST mit hohem Rezidivrisiko wird nach RO-Resektion eine 36-monatige Imatinib-Therapie empfohlen (Casali et al. 2018).

Die Inhibitoren Imatinib, Sunitinib und Regorafenib sind sogenannte Typ 2 Inhibitoren, die an die inaktive Konformation der Tyrosinkinase binden. Neuere Typ 1 Tyrosinkinase-Inhibitoren, die auch die aktive Kinase-Konformation binden, werden zurzeit entwickelt bzw. in initialen klinischen Studien bei GIST getestet und könnten in Zukunft die Möglichkeiten der zielgerichteten Therapie von GIST erweitern (Krishnamurty und Maly 2010).

GIST, die statt Mutationen in den Rezeptortyrosinkinase-Genen eine *BRAF* V600E Mutation aufweisen, zeigen keine erhöhte Sensitivität gegenüber Tyrosinkinase-Inhibitoren und sind Kandidaten für eine BRAF-gerichtete Therapie (Falchook et al. 2013). Neuere seltene Genfusionen bei GIST schließen z. B. die *NTRK3-ETV6*-Genfusion ein, so dass in seltenen Fällen von einer erhöhten Sensitivität dieser GIST gegenüber NTRK-Inhibitoren ausgegangen werden kann (Brenca et al. 2016).

#### 1.8 Zytogenetische Veränderungen

Charakteristische chromosomale Veränderungen bei GIST umfassen Verluste an den Chromosomenarmen 1p, 14q und 22q, wobei Verluste an 14q mit gastrischen GIST und einem eher prognostisch günstigen Verlauf, Verluste an 1p und 22q mit intestinalen GIST und einem höheren Malignitätsrisiko assoziiert sind (El-Rifai et al. 2000; Gunawan et al. 2007; Schaefer et al. 2014; Wozniak et al. 2007). Weitere charakteristische Aberrationen betreffen Verluste an 15q, 13q, 9p, 9q sowie Zugewinne an 8q (Gunawan et al. 2007; Wozniak et al. 2007).

Für das in der vorliegenden Arbeit analysierte Göttinger GIST-Kollektiv liegen mittels konventioneller komparativer genomischer Hybridisierung (engl. *comparative genomic hybridization*, CGH) erhobene zytogenetische Daten vor (Gunawan et al. 2007). Mittels konventioneller CGH können nicht balancierte chromosomale Alterationen auf einem verglichen mit heutigen hochauflösenden, Microarray-basierten CGH-Methoden begrenzten Auflösungsniveau von ca. 10 Megabasen detektiert werden, so dass diese Methode große zytogenetische Aberrationsmuster und chromosomale Instabilität erfassen kann, während Aussagen über die beteiligten Gene in der Regel nicht möglich sind (Kallioniemi et al. 1992; Pinkel et al. 1998; Solinas-Toldo et al. 1997). Der bei GIST häufig betroffene lange Arm des Chromosoms 22 kann bei einer Länge des Chromosomenarms von ca. 35 Megabasen somit mittels konventioneller CGH dargestellt werden.

#### 1.9 Fragestellung

In der klinischen Bewertung von GIST werden Risikostratifizierungen von Miettinen (AFIP) bzw. Fletcher (NIH) sowie die auf der Miettinen-Klassifikation aufbauende TNM-Klassifikation der UICC verwendet. Diese Vorhersagesysteme bauen auf empirischen Schwellenwerten auf. Die Leistung dieser Klassifikationssysteme wurde zwar bezüglich der Risikoprädiktion bestätigt (Joensuu et al. 2012), während die Auswahl der kanonischen Schwellenwerte für die numerischen Parameter Tumorgröße und Mitoserate bisher aber kaum systematisch analysiert wurde. In der vorliegenden Arbeit soll daher ausgehend von einem nicht durch Effekte adjuvanter spezifischer Therapien beeinflussten großen GIST-Kollektiv der Wert der Konsensus-Schwellenwerte für die Risikostratifizierung von GIST, die Tumorgrößen von 2 cm, 5 cm und 10 cm sowie die Mitoseraten von 5 pro 50 hpf und 10 pro 50 hpf, untersucht werden. Für diese Analyse wurden verschiedene geeignete statistische Verfahren vergleichend angewandt, wobei die zur Bearbeitung der Fragestellung ausgewählten Methoden die Möglichkeit einer Überanpassung eines Risikomodells an den der Berechnung zugrundeliegenden Datensatz statistisch berücksichtigen und minimieren. Die Etablierung statistischer Verfahren zur Stratifizierung kontinuierlicher numerischer Parameter wie der Tumorgröße und der Mitoserate wird anschließend auf einen molekularpathologischen Parameter, die Anzahl chromosomaler Imbalancen pro Chromosomenarm, angewandt.

Die vorliegende Arbeit soll folgende Fragen beantworten: Stellen die kanonischen Werte für Tumorgröße und Mitoserate optimale Schwellenwerte dar? Lassen sich Risikoschichten mit unterschiedlichem biologischem Risikopotential statistisch nachvollziehen? Welche Wertigkeit kann der Bestimmung chromosomaler Nettoveränderungen als möglicher unabhängiger Risikoparameter zur Dignitätsbestimmung bei GIST zugeordnet werden?

### 2 Methoden

#### 2.1 Statistische Analysen

Die Berechnungen wurden in der Software R (R-Core-Team 2019), Version 3.3.0, durchgeführt. Das rezidivfreie Überleben wurde als die Zeit von der Resektion bis zum ersten Auftreten eines lokalen Rezidivs oder einer Metastase definiert. Als Signifikanzniveau, d. h. Wahrscheinlichkeit, mit der die angenommene Nullhypothese, ein beobachteter Unterschied sei zufällig zustande gekommen, fälschlich abgelehnt wird, wurde ein Wert von  $\alpha$  = 0,05 angenommen.

#### 2.1.1 Kaplan-Meyer-Kurven

Überlebenszeitkurven wurden mittels des Kaplan-Meyer-Estimators berechnet (Kaplan und Meier 1958), der Standardmethode bei sogenannten rechts zensierten Daten, d. h. Überlebensdaten oder rezidivfreien Überlebensdaten, für die nicht bei allen eingeschlossenen Individuen das erwartete Ereignis (hier das Rezidiv eines primären GIST) beobachtet wurde, so dass z. B. der endliche Beobachtungszeitraum, ein Versterben vor Eintritt eines Rezidivs oder ein Ausscheiden aus der Studie im Modell berücksichtigt werden.

#### 2.1.2 Ereigniszeitanalysen

Der Einfluss von Risikofaktoren auf das rezidivfreie Überleben wurde durch Ereigniszeitanalysen mittels der Kaplan-Meyer-Formel gefolgt von statistischen Tests für Überlebenskurven-Differenzen untersucht, wobei zur Analyse auf statistisch signifikante Differenzen zwischen Überlebenskurven *Log-rank*-Tests mittels der R-Funktion *survdiff* sowie Likelihood-ratio-Tests mittels der R-Funktion *coxph*, d. h. basierend auf Coxschen Regressionsmodellen, durchgeführt wurden.

#### 2.1.3 Coxsches Regressionsmodel (Cox proportional hazards model)

Als statistisches Verfahren zur uni- oder multivariaten Abschätzung numerischer oder kategorischer Variablen auf das rezidivfreie Überleben wurden Coxsche Regressionsmodelle (Cox 1972) verwendet (*coxph* Funktion in R). Zur graphischen Darstellung des Risikoverhältnisses (*hazard ratio*) wurden die Risikokategorien der untersuchten Schemata als einheitliche numerische Parameter, d. h. von 0-3, untersucht (Tabelle 1). In der TNM-Klassifikation wurden zudem die Risiken IA, IB (für Magen-GIST) und I (für nicht-gastrische GIST) in einer Gruppe zusammengefasst.

Das Coxsche Regressionsmodel berechnet unter Berücksichtigung der Beobachtungszeit eine Risikofunktion (*hazard function*) für jedes Risiko-Covariat. Die durch den statistischen Test ermittelten Risikoverhältnisse (*hazard ratios*) für die Risikokategorien wurden graphisch dargestellt und es wurde eine lineare Regression für die numerischen Risikolevel

Einheitliche	NIH-	Modifizierte	AFIP-	Modifizierte	TNM
numerische	Klassifikation	NIH-	Klassifikation	AFIP(A)- und	Stadien
Risikoklasse		Klassifikation		AFIP(B)-	
				Klassifikation	
0	very low	I	very low	low	IA, IB, I
1	low	II	low	intermediate	II
2	intermediate	111	intermediate	high	IIIA
3	high	IV	high	very high	IIIB

#### Tabelle 1: Einheitliche numerische Risikoklasse in der statistischen Auswertung

Zur statistischen Vergleichbarkeit der verschiedenen Risikoklassifikationssysteme wurden den Risikoklassen einheitliche numerische Risikoniveaus 0-3 zugewiesen (Spalte 1). Für das TNM-System (letzte Spalte) wurden zusätzlich die Stadien IA und IB für gastrische Tumoren sowie das Stadium I für nicht-gastrische Tumoren zum Zweck der statistischen Analyse in einer Risikoklasse zusammengefasst.

1-3 mittels der R Funktion *Im* durchgeführt, um die Linearität der Risikoverhältnisse in den untersuchten Risikoschemata zu vergleichen.

#### 2.1.4 Maximal ausgewählte Rankstatistik (*maximally selected rank statistics*)

Maximal ausgewählte Rankstatistiken (*maximally selected rank statistics*) wurden mittels des *Maxstat*-Paketes in R berechnet (Lausen et al. 2004). Der *Maxstat*-Algorithmus ermittelt einen optimalen Schwellenwert oder einen signifikanten Bereich für einen numerischen Risikofaktor und berücksichtigt durch multiples Testen bedingte statistische Effekte. Der getestete Schwellenwertbereich wird innerhalb des Softwarepaketes *maxstat* durch eine standardisierte *Log-rank*-Statistik beschrieben, wobei zur Darstellung der Standardabweichung zusätzlich Zufallsstichproben mit n = 500 Repetitionen aus dem Datensatz gezogen wurden (sog. *bootstrapping*). Es wurden die Standard-Programmparameter, d. h. *Log-rank*-Statistik und die "Lau94"-*p*-Wert-Approximation (Worsley 1982) verwendet. Für den gesamten Datensatz wurde das Standard-Signifikanzniveau von p = 0,05 angenommen, für die separate Analyse von Magen-GIST und nicht-gastrischen GIST zur Berücksichtigung der nachträglichen Subgruppenbildung ein Signifikanzniveau p = 0,025.

#### 2.1.5 Risikoklassifikationsbäume (*risk classification trees*)

Regressionsbäume bzw. Risikoklassifikationsbäume (Breiman et al. 1984) wurden mittels rekursiver Partitionierung innerhalb des *Rpart*-Paketes in R berechnet und durch Kaplan-Meyer-Kurven und Überlebensdifferenzanalysen mit *Log-rank*-Tests visualisiert. In *rpart* werden Risikofaktoren als Zweige eines Baummodels von einem Satz durch den Benutzer vorgegebener möglicher Risikofaktoren hinzugefügt, bis das Risikomodell nicht weiter verbessert werden kann oder alle vorgegebenen Risikofaktoren im Modell berücksichtigt sind. Die entstandenen Risikoklassifikationsbäume sind als Entscheidungsbäume zu interpretieren, wobei die einzelnen Zweige die Entscheidungen repräsentieren und die Endpunkte die Risikoklassen. Ein vorgegebener Risikofaktor wird nur zu dem Baummodell hinzugefügt, wenn sich dadurch das Modell verbessert. Für die Berechnungen in *rpart* wurden kategorisierte Schwellenwerte für die Tumorgröße von 1-12 cm in 0,5-cm-Schritten sowie für die Mitoserate von 1-12 Mitosen pro 50 hpf in Einerschritten vorgegeben. Desweiteren wurden Modelle mit vorgegebenen kanonischen Größenwerten (2 cm, 5 cm und 10 cm) berechnet.

#### 2.1.6 Umfassende rezidivfreie Überlebensdifferenzanalyse mit Bootstrapping

Die Wertigkeit von Zwei-Schwellenwert-Modellen für die Parameter Größe oder Mitose in Risikomodellen wurde durch rezidivfreie Überlebensdifferenzanalysen mittels der R *survdiff* Funktion berechnet. Hierbei wurden alle möglichen Schwellenwertkombinationen in einem Bereich von 2-12 cm für die Tumorgröße und 2-12 Mitosen pro 50 hpf für die Mitoserate getestet, wobei vorgegeben wird, dass der zweite Schwellenwert mindestens zwei numerische Einheiten größer als der erste Schwellenwert sein muss und die gebildeten Untergruppen mindestens 5 % der Patienten beinhalten müssen. Zur Minimierung des Einflusses einzelner Fälle auf das Ergebnis wurden wiederum Zufallsstichproben (*bootstrapping*) mit *n* = 500 durchgeführt, und der mittlere  $\chi^2$ -Wert als Funktion der Kombination aus beiden Schwellenwerten graphisch dargestellt (mit optimalen Schwellenwertkombinationen als dunkle Areale).

#### 2.2 Genehmigung durch die Ethikkommission

Die Studie wurde von der Ethikkommission unter der Nummer 14/01/05 genehmigt.

### 3 Ergebnisse

#### 3.1 Göttinger GIST-Kollektiv

Das Göttinger GIST-Kollektiv, auf dessen Basis in der vorliegenden Arbeit die Stratifizierung der Risikoparameter untersucht wird, enthält GIST-Fälle aus einem bis in die 1990er Jahre zurückreichenden Zeitraum mit teils retrospektiv reklassifizierten Fällen und stellt daher heute eine besondere Quelle für eine Untersuchung der Risikobewertung beim GIST dar (Gunawan et al. 2007). Ähnliche Untersuchungen an neueren Fallkollektiven würden dagegen aufgrund des Anteils adjuvant behandelter Patienten zunehmend schwierig.

Aus dem Göttinger GIST-Kollektiv wurden 161 Imatinib-naive primäre GIST mit vorhandener Tumorgröße, Mitoserate, Tumorlokalisation sowie Untersuchung auf klonale Nettoveränderungen in der CGH-Analyse analysiert, darunter 103 gastrische und 58 nichtgastrische GIST aus dem Zeitraum 1989-2006 (Gunawan et al. 2007) mit einem medianen Beobachtungszeitraum von 34 Monaten, einem mittleren Beobachtungszeitraum von 44 Monaten und einem maximalem Beobachtungszeitraum von 156 Monaten. Die durchschnittliche Tumorgröße betrug 6,2 cm (Median: 5 cm), die durchschnittliche Mitoserate 16,5 pro 50 hpf (Median: 2 pro 50 hpf). Innerhalb der Fälle ohne beobachteten Progress betrug die Tumorgröße 5,2 cm (Median: 4 cm) und die mittlere Mitoserate 7,1 pro 50 hpf (Median: 2 pro 50 hpf). Innerhalb der Fälle mit Progress betrug die mittlere Tumorgröße 10,1 cm (Median: 9 cm) und die mittlere Mitoserate 53,3 pro 50 hpf (Median: 30 pro 50 hpf). 20 % der Fälle zeigten ein Rezidiv (Kaplan-Meyer-Kurve, Abbildung 2A). Für ein lokalisationsabhängiges Modell mit Unterteilung in Magen, Dünndarm und Colon sowie Rektum ergab sich ein signifikant verschiedenes rezidivfreies Überleben nach Lokalisation (p = 0,008 für das Gesamtmodell, Abbildung 2B). In einer multivariaten Analyse ergab sich für die Gruppe Dünndarm und Colon gegenüber Magen ein statistisch signifikanter Unterschied (p = 0,0184) und auch für das Rektum ein signifikanter Unterschied (p = 0,0066; zur Verteilung der Fälle nach den untersuchten Risikoklassifikationssystemen, siehe Tabelle 2).

	Fall	Risiko-		Modifi-				
AFIP	rali-	klasse	NIH	zierte	AFIP	AFIP(A)	AFIP(B)	TNM
Gruppe	2011	(Tab. 1)		NIH				
1	24 (0)	0	19 (0)	70 (0)	24 (0)	97 (0)	97 (0)	95 (0)*
2	46 (0)	1	39 (0)	39 (3)	73 (0)	15 (3)	15 (3)	17 (3)
За	31 (0)	2	48 (3)	20 (9)	15 (3)	34 (19)	16 (8)	21 (8)
3b	7 (3)	3	55 (30)	32 (21)	49 (30)	15 (11)	33 (22)	28 (22)
4	3 (0)							
5	11 (6)							
6а	24 (13)							
6b	15 (11)							

#### Tabelle 2: Fallzahlen nach Risikoniveau der untersuchten Klassifikationssysteme

Die Anzahl der Fälle mit Rezidiv sind jeweils in runden Klammern hinter der Gesamtzahl angegeben. Im linken Teil der Tabelle ist die Anzahl der Fälle nach AFIP-Gruppen gezeigt, im rechten Teil die jeweilige Anzahl der Fälle in den Risikoklassen der untersuchten Klassifikationssysteme. \*Dem TNM-Stadium IA sind 45 Fälle zugeordnet, dem TNM-Stadium IB 25 Fälle und dem TNM-Stadium I 25 Fälle (jeweils ohne beobachtetes Rezidiv).

#### 3.1.1 Vergleich der Risikoklassifikationssysteme

Rezidivfreie Überlebensdifferenzanalysen stellten sich in den originalen NIH- und AFIP-Klassifikationen überlappende Kurven ohne Rezidivfälle für die beiden niedrigsten Risikoklassen dar (Abbildung 3A, C), wogegen die modifizierten NIH- und AFIP-Klassifikationen sowie die TNM-Klassifikation unterschiedliche Kurven zwischen den beiden niedrigsten Risikoklassen zeigten (Abbildung 3B, D-F). Die modifizierte AFIP(B)-Klassifikation, die sich von der modifizierten AFIP(A)-Klassifikation in der Definition der höchsten Risikoklassen unterscheidet, zeigte zudem eine gering bessere Trennung der beiden höchsten Risikoklassen, *"high"* und *"very high"*, verglichen mit der AFIP(A)-Klassifikation (Abbildung 3D, E).



Abbildung 2: Allgemeine Charakteristika des GIST-Kollektivs

(A) Kaplan-Meyer-Kurve mit 95%-Konfidenzintervall (n = 161, davon 33 Fälle bzw. 20 % der Fälle mit Rezidiv). (B) Kaplan-Meyer-Kurven nach Lokalisation ohne extra-gastrointestinale GIST (n = 3) zusammengefasst in 3 Gruppen (n = 158, davon 30 Fälle mit Rezidiv), Magen und Omentum (n = 105, davon 14 Fälle bzw. 13 % der Fälle mit Rezidiv), Dünndarm und Colon (n = 43, davon 11 Fälle bzw. 26 % der Fälle mit Rezidiv) sowie Rektum (n = 10, davon 5 Fälle bzw. 50 % der Fälle mit Rezidiv). Für das Modell ergibt sich im *Log-rank*-Test p = 0,008. Im Cox proportional hazards Model ergibt sich signifikante p-Werte für Dünndarm und Colon (p = 0,027) und für das Rektum (p = 0,010) gegenüber dem durch die Gruppe Magen und Omentum gegebenen Basisrisiko.

Das Risikoverhältnis als Funktion des Risikolevels zeigt bei den verschiedenen Schemata eine unterschiedlich gute Linearität (Abbildung 4A-F). In den ursprünglichen NIH- und AFIP-Klassifikationen (Abbildung 4A, C) sind die unteren Risikoniveaus 0-1 bzw. *"very low"* und *"low"* nicht unterscheidbar, in der NIH-Klassifikation ist die *"intermediate"* Risikoklasse zudem niedriger positioniert als in der AFIP-Klassifikation, während in beiden Klassifikationen ein deutlicher Risikoanstieg zwischen den Ebenen *"intermediate"* und *"high"* sichtbar ist (Abbildung 4A, C). Die modifizierten NIH- und AFIP-Risikomodelle sowie das TNM-Schema (Abbildung 4B, D-F) zeigten eine bessere Linearität gegenüber den älteren Klassifikationen, wobei die R<sup>2</sup>-Werte als Maß für die Varianz, die durch ein lineares Modell unter Einbezug der Risikoniveaus 1-3 erklärt werden kann, bei der modifizierten NIH-Klassifikation, der AFIP(B)-Klassifikation und der TNM-Klassifikation jeweils höher war



Abbildung 3: Vergleichende rezidivfreie Überlebensdifferenzanalysen der untersuchten Klassifikationsschemata.

(A) NIH-Klassifikation, (B) modifizierte NIH-Klassifikation, (C) AFIP-Klassifikation, (D-E) die modifizierten AFIP-Schemata, AFIP(A) und AFIP(B), sowie (F) TNM-Klassifikation (horizontale Achse: Zeit in Monaten; vertikale Achse: relatives rezidivfreies Überleben). Die Abbildung entspricht dem ersten Teil der Abbildung 2 der begleitenden Publikation (Sander et al. 2020) mit freundlicher Genehmigung des Elsevier Verlages.

als in den ursprünglichen Klassifikationen. Für die AFIP(A)-Klassifikation war der R<sup>2</sup>-Wert niedriger als für die ursprüngliche AFIP-Klassifikation, wobei sich dennoch eine verbesserte Linearität zwischen den beiden unteren Risikoniveaus bei etwas dichter gegeneinander platzierten höheren Risikoniveaus zeigt (Abbildung 4D). In der AFIP(B)-Klassifikation erschien dagegen insbesondere die Linearität zwischen den beiden höheren Risikoniveaus besser als bei der AFIP(A)-Klassifikation (Abbildung 4E).



Abbildung 4: Risikoverhältnis als Funktion der Risikoklasse für die verschiedenen Risikoklassifizierungsschemata.

Für diese Analyse wurden die Risikoklassen der Schemata als lineare numerische Risikolevel (von 0 bis 3) dargestellt. Für die NIH-Klassifikation **(A)**, entsprechen die numerischen Level *"very low"*, *"low"*, *"intermediate"*, *"high"*; für die modifizierte NIH-Klassifikation **(B)** den Klassen I – IV; für die AFIP-Klassifikation **(C)** den heute gebräuchlichen Bezeichnungen *"very low"*, *"low"*, *"intermediate"*, *"high"*; für die modifizierte AFIP(A)-Klassifikation **(D)** sowie die modifizierte AFIP(B)-Klassifikation **(E)** *"low"*, *"intermediate"*, *"high"* und *"very high"*; für die TNM-Klassifikation **(F)** I, II, IIIA und IIIB. Im Vergleich zu den älteren Klassifikationen (NIH und AFIP) zeigen die neueren Schemata insgesamt einen besseren linearen Anstieg gegenüber dem niedrigsten Risikolevel sowie eine bessere Unterscheidung zwischen den ersten beiden Risikoleveln. Die R<sup>2</sup>-Werte wurden für die Ausgleichsgerade zwischen den Risikoniveaus 1-3 berechnet und sind jeweils in der unteren rechten Ecke angegeben. Die Abbildung entspricht dem zweiten Teil der Abbildung 2 der begleitenden Publikation (Sander et al. 2020) mit freundlicher Genehmigung des Elsevier Verlages.

Risiko	coef	exp(coef)	se(coef)	Z	р
NIH	2,488	12,040	0,569	4,374	1,220e-05
Lokalisation	1,073	2,925	0,350	3,066	2,170e-03
Modifizierte NIH	1,398	4,047	0,230	6,068	1,300e-09
Lokalisation	1,274	3,576	0,363	3,507	4,530e-04
AFIP	2,118	8,310	0,484	4,376	1,210e-05
Lokalisation	0,433	1,542	0,356	1,217	2,240e-01
AFIP(A)	1,186	3,275	0,191	6,203	5,550e-10
Lokalisation	0,821	2,272	0,357	2,299	2,150e-02
AFIP(B)	1,343	3,832	0,230	5,834	5,400e-09
Lokalisation	0,989	2,689	0,362	2,729	6,350e-03
TNM	1,320	3,744	0,220	6,013	1,820e-09
Lokalisation	0,365	1,441	0,368	0,992	3,210e-01

# Tabelle3:MultivariateAnalysemittelsCox-proportional-hazards-ModellenfürdieRisikoklassifikationen und das Risiko Lokalisation.

In den multivariaten Analysen wird untersucht, ob die Lokalisation einen zusätzlichen prognostischen Wert bei gegebener Risikoklassifikation hat. Es ergeben sich hochsignifikante *p*-Werte für die Risikoklassifikationen, wobei sich die Tumorlokalisation (gastrisch versus nicht-gastrisch) für die NIH-Klassifikation, die modifizierte NIH-Klassifikation sowie die modifizierten AFIP Klassifikationen, AFIP(A) und AFIP(B), als signifikanter unabhängiger Risikofaktor ergibt, während sich die Tumorlokalisation für die originale AFIP-Klassifikation und die TNM-Klassifikation nicht als signifikanter unabhängiger Risikofaktor zeigt.

Wurde die Tumorlokalisation (Magen-GIST versus nicht-gastrische GIST) multivariat mittels Coxschen Regressionsmodellen als Risikofaktor bei gegebener Risikoklassifikation untersucht, ergab sich ein signifikanter zusätzlicher prognostischer Wert für die Tumorlokalisation für das ursprüngliche sowie modifizierte NIH-Schema sowie die beiden modifizierten AFIP-Schemata (Tabelle 3) als Hinweis auf einen weiteren prognostischen Wert der Tumorlokalisation als Risikofaktor in diesen Schemata unabhängig von einer bereits erfolgten Berücksichtigung der Tumorlokalisation in den AFIP-Schemata.

#### 3.1.2 Analysen für den numerischen Risikoparameter Tumorgröße

Mittels maximal selektierter Rankstatistik (*Maxstat*-Analyse in R) wurde die prognostische Wertigkeit einer numerischen Variablen in einem aus dem Datensatz abgeleiteten Wertebereich berechnet und als standardisierte *Log-rank*-Statistik über diesen Bereich graphisch dargestellt (Abbildung 5A-C). Das Maximum der erhaltenen Kurve entspricht dem in diesem Test ermittelten optimalen Schwellenwert für einen numerischen Risikofaktor, das statistische Signifikanzniveau (definiert als  $\alpha$  = 0,05 für den gesamten Datensatz und  $\alpha$  = 0,025 für die getrennte Betrachtung von Magen-GIST und nichtgastrischen GIST) wurde zudem als graue Linie (Abbildung 5A-C) dargestellt. Für die Tumorgröße wurde so für Schwellenwerte von 3-12 cm ein signifikanter Modellbereich ermittelt mit einem kleineren Maximum bei etwa 4 cm und einem Hauptmaximum bei 11 cm (Abbildung 5A). Für Magen-GIST wurde mit dem *Maxstat*-Verfahren ein signifikantes Maximum bei 10-11 cm beobachtet (Abbildung 5B), für nicht-gastrische Tumoren ein signifikantes Hauptmaximum bei 4-5 cm sowie ein nicht-signifikantes kleineres Nebenmaximum bei 10-11 cm (Abbildung 5C).

Die Analyse mittels rekursiver Partitionierung und Darstellung eines Risikoklassifikationsbaumes (*Rpart*-Analyse in R) ermöglicht dem *Rpart*-Algorithmus, aus den möglichen Schwellenwerten hier von 2 bis 12 cm in Intervallen von 0,5 cm ein optimales Modell zu erstellen und auch die Anzahl der verwendeten Schwellenwerte zu wählen, wobei sich ein Modell mit zwei Schwellenwerten bei 4 cm und 10,5 cm (d. h. < 4 cm,  $\geq$  4 cm und < 10,5 cm,  $\geq$  10,5 cm) als bestes Modell zeigt (Abbildung 5D, mit zugehöriger Kaplan-Meyer-Analyse in Abbildung 5E), was gut zu den Ergebnissen der *Maxstat*-Analyse passt. In einer dritten, unabhängigen Analyse wurden alle möglichen Zwei-Schwellenwert-Modelle für die Tumorgröße durch rezidivfreie Überlebensdifferenzanalysen mittels *Logrank*-Statistik bewertet, wobei für jedes Modell 500 Zufallsstichproben (*bootstrapping*) analysiert wurden und der jeweilige  $\chi^2$ -Mittelwert graphisch dargestellt wurde. Konsistent mit den *Maxstat*- und *Rpart*-Analysen zeigte sich eine Region optimaler  $\chi^2$ -Werte bei 4-5 cm für den ersten Schwellenwert in Kombination mit 10-11 cm für den zweiten Schwellenwert (Abbildung 5F).



Abbildung 5: Schwellenwert-Analysen für die Tumorgröße.

(A) Die maximal selektierte Rankstatistik (*Maxstat*-Analyse) für die Tumorgröße zeigt in einem Interval von 3-12 cm signifikante Schwellenwerte (graue Linie: Signifikanzniveau von  $\alpha = 0,05$ ). (B) Die *Maxstat*-Analyse für Magen-GIST zeigt ein Maximum bei 10-11 cm (graue Linie:  $\alpha = 0,025$ ). (C) Die *Maxstat*-Analyse für nicht-gastrische GIST zeigt ein Maximum bei 4-5 cm (graue Linie:  $\alpha = 0,025$ ). (D) In der *Rpart*-Analyse wird ein Risikoklassifikationsbaum mit den Schwellenwerten 4 cm und 10,5 cm als optimales Modell ausgewählt mit (E) einem *p*-Wert von < 1e-6 in der rezidivfreien Überlebensdifferenzanalyse (RFÜ, rezidivfreies Überleben). (F-H) Graphische Darstellung der  $\chi^2$ -Werte für alle Zwei-Schwellenwert-Modelle in einem Bereich von 2-12 cm. Dunklere Bereich zeigen höhere  $\chi^2$ -Werte, d. h. bessere Schwellenwert-Kombinationen, mit Hervorhebung von Regionen optimaler Modelle in Kreisen. Die Abbildung entspricht Abbildung 3 der begleitenden Publikation (Sander et al. 2020) mit freundlicher Genehmigung des Elsevier Verlages.

Für sowohl Magen-GIST als auch nicht-gastrische GIST wurden optimale Bereiche in derselben Region beobachtet (Abbildung 5G-H), wobei für Magen-GIST der niedrige Schwellenwert deutlich unschärfer definiert erscheint, da sich in der graphischen Darstellung (Abbildung 5G) ein breites horizontales Band hoher  $\chi^2$ -Werte ergibt.

#### 3.1.3 Analysen für den numerischen Risikoparameter Mitoserate

Für die Mitoserate pro 50 hpf wurden die Analysen wie für die Tumorgröße dargelegt durchgeführt. In der *Maxstat*-Analyse ergab sich ein weiter, statistisch signifikanter Schwellenwertbereich für  $\geq$  1 pro 50 hpf mit einem optimalen Schwellenwert bei 10 pro 50 hpf (Abbildung 6A). In der separaten Berechnung der Kurven für Magen-GIST und nicht-gastrische GIST zeigten beide Kurven ein Maximum bei ca. 10 pro 50 hpf (Abbildung 6B-C), wobei die Kurve für nicht-gastrische Tumoren im Vergleich mit der Kurve für Magen-GIST eine breitere Schulter für Mitoserateen über 10 pro 50 hpf mit einem signifikanten Bereich bis über 50 pro 50 hpf zeigte (Abbildung 6C). In der *Rpart*-Analyse wählte der Algorithmus ein Zwei-Schwellenwert-Modell mit Schwellenwerten von 5 pro 50 hpf und 10 pro 50 hpf als bestes Modell (Abbildung 6D, die Kaplan-Meyer-Analyse zu diesem Modell ist in Abbildung 6E dargestellt).

In der umfassenden rezidivfreien Überlebensanalyse für Zwei-Schwellenwert-Modelle (Abbildung 6F-H), die wie oben für die Tumorgröße beschrieben durchgeführt wurde, zeigte sich für das gesamte Fallkollektiv ein erster Schwellenwert bei 5-6 Mitosen pro 50 hpf in Kombination mit einem zweiten Schwellenwert bei 10-11 Mitosen pro 50 hpf als optimale Kombination (eingekreistes Areal in Abbildung 6F) konsistent mit der *Rpart*-Analyse. In der separaten Berechnung für Magen-GIST (Abbildung 6G) und nicht-gastrische GIST (Abbildung 6H) ergab sich ein insgesamt ähnliches Bild, wobei die Region optimaler  $\chi^2$ -Werte für beide Subgruppen breiter und weniger scharf definiert erschien und sich insbesondere bei den nicht-gastrischen GIST (Abbildung 6H) für den unteren Schwellenwert (Schwellenwert 1) kein scharfes Maximum ergab, während sich für den hohen Schwellenwert (Schwellenwert 2) ein Bereich optimaler  $\chi^2$ -Werte für einen Bereich von > 9-12 pro 50 hpf ergab.



#### Abbildung 6: Schwellenwert-Analysen für die Mitoserate.

(A) Die maximal selektierte Rankstatistik (*Maxstat*-Analyse) für die Mitoserate zeigt, dass ein breiter Bereich von Werten signifikante Modelle erzeugt mit einem Maximum bei ca. 10 pro 50 hpf und einer breiten Schulter bis ca. 60 pro 50 hpf (graue Linie: Signifikanzniveau von  $\alpha = 0,05$ ). (B, C) Die *Maxstat*-Analysen für Magen-GIST und nicht-gastrische GIST verlaufen ähnlich zum Gesamtkollektiv, wobei sich insbesondere für nicht-gastrische GIST eine breite signifikante Schulter bis > 40 pro 50 hpf ergibt (graue Linie:  $\alpha = 0,025$ ). (D) In der *Rpart*-Analyse wird ein Risikoklassifikationsbaum mit den zwei Schwellenwerten 5 pro 50 hpf und 10 pro 50 hpf als optimales Modell ausgewählt mit (E) einem *p*-Wert von < 1e-6 in der rezidivfreien Überlebensdifferenzanalyse (RFÜ, rezidivfreies Überleben). (F-H) Graphische Darstellung der  $\chi^2$ -Werte für alle Zwei-Schwellenwert-Modelle in einem Bereich von 2 pro 50 hpf bis 12 pro 50 hpf. Dunklere Bereich zeigen höhere  $\chi^2$ -Werte, d. h. bessere Schwellenwert-Kombinationen, mit Hervorhebung von Regionen optimaler Modelle in Kreisen. Für nicht-gastrische GIST zeigt sich eine unscharfe Region optimaler  $\chi^2$ -Werte. \* weisen auf nicht getestete Modelle mit zu geringer Fallzahl in einer der Untergruppen hin. Die Abbildung entspricht Abbildung 4 der begleitenden Publikation (Sander et al. 2020) mit freundlicher Genehmigung des Elsevier Verlages.

#### 3.1.4 Tumorgröße und Mitoserate in der Risikobewertung von GIST

Im Vergleich von Tumorgröße und Mitoserate zeigte sich in einer Darstellung des relativen Anteils beobachteter Rezidive als Funktion des jeweiligen numerischen Parameters ein unterschiedliches Verhalten für die beiden Parameter (Abbildung 7A, B; die Anzahl aller beobachteter Rezidive wurde jeweils als 100 % festgelegt). Die für die Tumorgröße erhaltene Kurve begann bei 0 % beobachteter Rezidive, d. h. für sehr kleine GIST wurden keine Rezidive beobachtet (Abbildung 7A), während die Kurve für die Mitoserate bei  $\leq$  3 Mitosen pro 50 bei ca. 10 % der Rezidive beginnt (Abbildung 7B). Histologisch wenige oder fehlende Mitosen waren daher mit beobachteten Rezidiven in einem Teil der Fälle verbunden, insbesondere bei sehr großen und nicht im Magen lokalisierten Tumoren. Im Einzelnen handelte es sich dabei um einen Dünndarm-GIST von 11 cm ohne histologisch nachweisbare Mitosen, einen Dünndarm-GIST von 25 cm mit einer Mitose pro 50 hpf und einen extragastrointestinalen Unterbauch-GIST von 12 cm und einer Mitose pro 50 hpf. Einige Risikoklassifikationssysteme verwenden drei Schwellenwerte für die Tumorgröße 2 cm 5 cm und 10 cm. In einer Rpart-Analyse mit Vorgabe dieser drei Schwellenwerte kombiniert mit Zufallsstichproben (bootstrapping) ergab sich in ca. 67% der Programmläufe ein Risikoklassifikationsbaum, in dem alle drei Schwellenwerte durch das Programm zur Bildung eines optimalen Modells berücksichtigt wurden, so dass die Verwendung aller drei Schwellenwerte in der Rpart-Analyse statistisch nachvollziehbar ist (Abbildung 7C, mit korrespondierender Kaplan-Meyer-Analyse in Abbildung 7D mit p = 9,9e-6 im Log-rank-Test).



Abbildung 7: Tumorgröße und Mitoserate als Risikoparameter.

(A, B) Anteil der beobachteten Rezidive als Funktion der jeweiligen Parameter für (A) Tumorgröße und (B) Mitoserate. Zu Vergleichszwecken wurde die Kurve für die Tumorgröße als gestrichelte graue Kurve der Abbildung (B) hinzugefügt. Während die Kurve für die Tumorgröße im Ursprung beginnt, zeigt die Kurve für die Mitoserate bei  $\leq$  3 Mitosen pro 50 hpf in ca. 10 % bereits Rezidive. Für sehr kleine Tumoren wurden daher keine Rezidive beobachtet, während histologisch wenige Mitosen oder ein Fehlen von Mitosen in einem Teil der Fälle dennoch mit einem beobachteten Rezidiv verbunden war. **(C, D)** *Rpart*-Analyse für ein Drei-Schwellenwert-Modell für die Tumorgröße aus den Schwellenwerten 2 cm, 5 cm und 10 cm wie in einigen Risikovorhersagemodellen verwendet. Bei Vorgabe der drei Schwellenwerte für die Tumorgröße bestätigt der Risikoklassifikationsbaum **(C)** ein Modell unter Verwendung aller drei Schwellenwerte als optimales Modell mit **(D)** einem *p*-Wert von *p* = 9,9e-6 (*Log-rank*-Test). **(E)** Bezogen auf die Miettinen Gruppen zeigt sich ein Sprung im rezidivfreien Überleben von Gruppe 3a (keine beobachteten Rezidive) zu Gruppe 3b (p = 0,0001), d. h. bei niedriger Mitoserate wurden Rezidive erst bei Tumoren über 10 cm Größe beobachtet. **(F)** Bei in Gruppe 4 ebenso fehlenden Rezidiven zeigt sich ein weiterer Sprung von Gruppe 3a und 4 zusammen versus Gruppe 5 (p = 6,6e-6), d. h. kleine Tumoren (2-5 cm) mit hoher Mitoserate zeigen ein hohes Rezidivrisiko konsistent mit einer hohen Relevanz der Mitoserate für die Risikovorhersage. Die Abbildung entspricht Abbildung 5 der begleitenden Publikation (Sander et al. 2020) mit freundlicher Genehmigung des Elsevier Verlages.

Da die durchgeführten statistischen Tests für Tumorgröße und Lokalisation auf das Vorkommen bestimmter Übergänge im malignen Potential hinweisen, wurden anschließend auch die AFIP-Gruppen miteinander verglichen. Hier zeigte sich ein Sprung im Malignitätspotential zwischen den AFIP-Gruppen 3a und 3b (Abbildung 7E, p = 0,0001) sowie zwischen den Gruppen 3a und 4 zusammen versus Gruppe 5 (Abbildung 7F, p = 6,6e-6). In der kleinen Gruppe 3b zeigten 3 on 7 Fällen ein Rezidiv, in der Gruppe 5 zeigten 6 von 11 Fällen ein Rezidiv, während in den Gruppen 3a und 4 keine Rezidive zu beobachten waren. Die Gruppen 3a, 4 und 5 werden im modifizierten NIH-Schema als Risikolevel II zusammengefasst.

3.1.5 Anzahl chromosomaler Nettoveränderungen als numerischer Risikoparameter

Die Anzahl chromosomaler Nettoveränderungen bezogen auf Chromosomenarme betrug für das Fallkollektiv insgesamt im Mittel 3,9 (Median: 3), für Magen-GIST im Mittel 3,2 (Median: 2) und für nicht-gastrische GIST im Mittel 5,2 (Median: 4). Bezogen auf die Risikoklassifikationssysteme zeigten Risikoklassen mit höherem Risiko generell eine höhere Anzahl chromosomaler Nettoveränderungen, wobei nicht in jedem Fall die höchste Risikokategorie auch den höchsten Mittelwert von Nettoveränderungen aufwies (Tabelle 4). Es zeigten sich jedoch in allen Risikokategorien aller untersuchten Schemata im Mittel gleich viele oder mehr chromosomale Veränderungen innerhalb der Fälle mit beobachtetem Rezidiv als innerhalb aller Fälle einer Risikoklasse (Tabelle 4). In einer multivariaten Analyse mittels Coxschem Regressionsmodell erhielten bei Vorgabe der Risikoparameter Größe, Mitoserate, Tumorlokalisation und Anzahl klonaler Nettoveränderungen alle vier Parameter signifikante p-Werte als Hinweis darauf, dass alle Parameter unabhängigen prognostischen Wert für die Risikovorhersage haben (Tabelle 5; p = 0,0003). In der Schwellenwertanalyse mittels maximal selektierter Rankstatistik ergab sich für das gesamte Fallkollektiv ein Wert von 5 Nettoveränderungen (Abbildung 8A), für Magen-GIST von 4 Nettoveränderungen (Abbildung 8B) und für nicht-gastrische GIST von 6 Nettoveränderungen (Abbildung 8C).

AFIP Gruppe	Mittlere Anzahl CNV	Risiko- klasse (Tab. 1)	NIH	Modifi- zierte NIH	AFIP	AFIP(A)	AFIP(B)	TNM
1	1,9 (-)	0	2,0 (-)	2,8 (-)	1,9 (-)	2,6 (-)	2,6 (-)	2,7 (-)
2	3,3 (-)	1	3,4 (-)	2,5 (4,0)	2,9 (-)	5,7 (6,7)	5,7 (6,7)	5,2 (6,7)
3a	2,5 (-)	2	2,4 (4,0)	6,7 (6,7)	5,7 (6,7)	5,8 (7,2)	4,8 (5,3)	5,0 (8)
3b	4,6 (7,7)	3	6,1 (7,1)	6,3 (7,2)	5,9 (6,8)	6 (6,2)	6,4 (7,4)	6,5 (6,4)
4	2,0 (-)							
5	7,0 (6,0)							
6a	6,2 (7,5)							
6b	6,0 (6,2)							

# Tabelle 4: Anzahl chromosomaler Nettoveränderungen nach Risikoniveau der untersuchten Klassifikationssysteme.

Für die AFIP-Gruppen (erste zwei Spalten) sowie für die untersuchten Klassifikationssysteme ist jeweils die durchschnittliche Anzahl chromosomaler Nettoveränderungen bezogen auf Chromosomenarme angegeben, wobei die Werte in Klammern die durchschnittlichen Werte für Fälle mit Rezidiv angegeben (CNV, Anzahl chromosomale Nettoveränderungen bezogen auf Chromosomenarme).

Um die Wertigkeit der Anzahl klonaler Nettoveränderungen bei gegebener Risikoklassifikation durch die untersuchten Schemata sowie unter Berücksichtigung der bekannten Lokalisationsabhängigkeit chromosomaler Veränderungen (Gunawan et al. 2007) zu beurteilen, wurden weitere multivariate Coxsche Regressionsmodelle berechnet, wobei getestet wurde, ob die Parameter Tumorlokalisation und Anzahl klonaler Nettoveränderungen bezogen auf Chromosomenarme (< 5 versus  $\geq$  5) zusätzlichen unabhängigen prognostischen Wert bei gegebener Risikoklassifikation aufweisen (Tabelle 6). Für die originale NIH-Klassifikation stellten sich in der Gruppe *"high"* (höchstes Risikolevel) sowohl die Tumorlokalisation als auch die Anzahl klonaler Nettoveränderungen

	coef	exp(coef)	se(coef)	Z	р
Tumorgröße (> 5 cm)	1,18	3,2543	0,4644	2,541	0,011059
Mitoserate (> 5/50 hpf)	2,4747	11,8776	0,6393	3,871	0,000108
Lokalisation	1,2264	3,4091	0,3756	3,265	0,001095
Nettoveränderungen	1,4222	4,1461	0,3966	3,586	0,000336

## Tabelle 5: Multivariate Analyse mittels *Cox-proportional-hazards*-Modell für die Risikofaktoren Größe, Mitoserate, Lokalisation und Anzahl chromosomaler Nettoveränderungen.

In einer multivariaten Analyse zeigt sich das Risiko Anzahl chromosomaler Nettoveränderungen bezogen auf Chromosomenarme als unabhängiger signifikanter Risikofaktor (p = 0,0003) bei Tumorgröße, Mitoserate und Lokalisation (gastrisch versus nicht-gastrisch) als gegebene Risikofaktoren (coef, *coefficient* [Regressionskoeffizient]; exp, *exponential* [Exponentialfunktion]; se, *standard error* [Standardfehler]).

als unabhängige Risikoparameter dar. In der modifizierten NIH-Klassifikation stellte sich die Anzahl klonaler Nettoveränderungen in der Risikogruppe "III" (zweithöchstes Risikolevel) als unabhängiger Risikofaktor dar, in der Risikogruppe "IV" (höchstes Risikolevel) zeigten sich sowohl die Tumorlokalisation als auch die Anzahl klonaler Nettoveränderungen als unabhängige Risikofaktoren. In der originalen AFIP-Klassifikation stellte sich die Anzahl klonaler Nettoveränderungen in der Gruppe "*high*" (höchste Risikogruppe) als unabhängiger Risikoparameter dar, während die Tumorlokalisation keinen zusätzlichen prognostischen Nutzen zeigte. In der modifizierten AFIP(A)-Klassifikation wurde ein signifikanter *p*-Wert für die dritte Risikogruppe (*"high*") beobachtet, während in der höchsten Gruppe (*"very high*") weder Tumorlokalisation noch Anzahl klonaler Nettoveränderungen einen zusätzlichen prognostischen Wert zeigten. In der modifizierten AFIP(B)-Klassifikation zeigten sich in der *"very high*"-Risikogruppe signifikante p-Werte für sowohl klonale Nettoveränderungen als auch Tumorlokalisation, in der Risikogruppe *"high*" war dies nur für die Anzahl klonaler Nettoveränderungen der Fall. Für die TNM-Klassifikation

NIH	Risikoklassen	very low	low	intermediate	high	
$CNV \ge 5$	$p_{Log\text{-}rank}$	-	-	0,99	8,95e-05	
Lokalisation	$p_{Log\text{-}rank}$	-	-	0,99	0,00386	
Modif. NIH	Risikoklassen		I	П	Ш	IV
$CNV \ge 5$	$p_{Log\text{-}rank}$		-	0,99	0,0174	0,0234
Lokalisation	$p_{Log\text{-}rank}$		-	0,99	0,0583	0,0113
AFIP	Risikoklassen	very low	low	intermediate	high	
$CNV \ge 5$	$p_{Log-rank}$	-	-	0,99	6,46e-05	
Lokalisation	$p_{\scriptscriptstyle Log\text{-}rank}$	-	-	0,99	0,0544	
AFIP(A)	Risikoklassen		low	intermediate	high	very high
$CNV \ge 5$	$p_{Log-rank}$		-	0,99	0,000684	0,0759
Lokalisation	$p_{Log\text{-}rank}$		-	0,99	0,280367	0,1104
AFIP(B)	Risikoklassen		low	intermediate	high	very high
$CNV \ge 5$	$p_{ t Log ext{-rank}}$		-	0,99	0,00991	0,0237
Lokalisation	$p_{Log-rank}$		-	0,99	0,82581	0,0120
TNM	Risikoklassen		I	П	IIIA	IIIB
$CNV \ge 5$	<b>p</b> Log-rank		-	0,99	0,00662	0,0512
Lokalisation	$p_{Log-rank}$		-	0,99	0,23956	0,0525

Tabelle 6: Multivariate Analysen der Risikoniveau der untersuchten Risikoklassifikationen nach Lokalisation und Anzahl chromosomaler Nettoveränderungen.

Es werden multivariate Coxsche Regressionsmodelle berechnet, wobei getestet wird, ob die Parameter Tumorlokalisation und Anzahl klonaler Nettoveränderungen bezogen auf Chromosomenarme (< 5 versus  $\geq$  5) einen zusätzlichen unabhängigen prognostischen Wert haben. Für die originale NIH-Klassifikation stellen in der "high"-Gruppe sowohl Lokalisation als auch die Anzahl klonaler Nettoveränderungen unabhängige Risikoparameter dar. In der modifizierten NIH-Klassifikation stellt sich die Anzahl klonaler Nettoveränderungen in der Risikogruppe "III" als unabhängiger Risikofaktor dar, in der Risikogruppe "IV" zeigen sich sowohl die Tumorlokalisation als auch die Anzahl klonaler Nettoveränderungen als unabhängige Risikofaktoren. In der originalen AFIP-Klassifikation stellt sich die Anzahl klonaler Nettoveränderungen in der höchsten Risikogruppe ("high") als unabhängiger Risikoparameter dar, während die Tumorlokalisation keinen zusätzlichen prognostischen Nutzen zeigt, in der modifizierten AFIP(A)-Klassifikation ist dies für die dritte Risikogruppe (*"high"*) der Fall, während in der höchsten Gruppe (*"very high"*) weder Tumorlokalisation noch Anzahl klonaler Nettoveränderungen einen zusätzlichen prognostischen Wert zeigen. In der modifizierten AFIP(B)-Klassifikation zeigen sich in der "very high"-Risikogruppe signifikante p-Werte für sowohl klonale Nettoveränderungen als auch Tumorlokalisation, in der "high"-Risikogruppe ist dies nur für die Anzahl klonaler Nettoveränderungen der Fall. Für die TNM-Klassifikation ergibt sich ein zusätzlicher prognostischer Nutzen für die Anzahl klonaler Nettoveränderungen in der Risikogruppe "IIIA". Insgesamt ist die Anzahl chromosomaler Nettoveränderungen insbesondere geeignet, die hohen Risikoniveaus weiter aufzutrennen.



Abbildung 8: Anzahl der klonalen Nettoveränderungen als Risikoparameter.

(A) Mittels maximal selektierter Rankstatistik (*Maxstat*-Analyse in R) ergibt sich ein optimaler Schwellenwert von 5 durch CGH nachgewiesene klonale Nettoveränderungen bezogen auf Chromosomenarme bei einem weiten statistisch signifikanten Bereich von 1-9 Nettoveränderungen (A; graue Linie: Signifikanzniveau  $\alpha = 0,05$ ). Für die Subgruppe der Magen-GIST ergibt sich ein Maximum bei 4 Nettoveränderungen (B; graue Linie: Signifikanzniveau  $\alpha = 0,025$ ) und für nicht-gastrische GIST von 6 Nettoveränderungen (C; graue Linie: Signifikanzniveau  $\alpha = 0,025$ ).

ergab sich ein zusätzlicher prognostischer Nutzen für die Anzahl klonaler Nettoveränderungen in der Risikogruppe "IIIA". Insgesamt ist die Anzahl chromosomaler Nettoveränderungen daher geeignet, die hohen Risikoniveaus weiter aufzutrennen, auch in Risikogruppen, in denen eine weitere Unterscheidung nach Lokalisation keinen zusätzlichen prognostischen Nutzen zeigt.

### 4 Diskussion

Die vorliegende Arbeit untersucht die Wertigkeit von Schwellenwerten für numerische Risikoparameter für die Risikovorhersage bei GIST. Zu diesem Zweck wurde ein Fallkollektiv analysiert, das bei teils retrospektiv reklassifizierten Fällen bis in die frühe Zeit der GIST-Diagnostik zurückreicht und deren Patienten noch keine Behandlung mit Tyrosinkinaseinhibitoren erhalten haben. Das Göttinger GIST-Kollektiv stellt somit eine besondere Datenquelle bei der Frage der prognostischen Beurteilung von GIST dar.

#### 4.1 Schwellenwerte für Tumorgröße und Mitoserate

Zur statistischen Validierung von Schwellenwerten wurden unterschiedliche Ansätze verwendet inklusive der Berechnung von Regressionsbäumen, maximal selektierter Rankstatistik und rezidivfreie Überlebensdifferenzanalysen mit Zufallsstichproben (bootstrapping). Die verwendeten Verfahren berücksichtigen dabei, dass eine begrenzte Fallpopulation vorliegt und daher das Risiko einer Überanpassung an die vorliegenden Daten besteht (Fehrmann et al. 2007; Orford et al. 2002). Zusammengenommen unterstützen die Ergebnisse die Annahme optimaler Schwellenwerte von 4-5 cm sowie 10-11 cm für die Tumorgröße sowie von 5 pro 50 hpf und 10 pro 50 hpf für die Mitoserate. Die kanonischen, in der klinischen Praxis verwendeten Werte repräsentieren somit geeignete Schwellenwerte für die Zuordnung von GIST zu Risikogruppen. Da die verschiedenen statistischen Verfahren für die Tumorgröße den ersten Schwellenwert bei ca. 4-5 cm einordnen, deuten die Ergebnisse an, dass auch etwas weniger als 5 cm durchmessende Tumoren bereits biologisch ähnlich zu Tumoren sind, die den 5 cm-Schwellenwert bereits überschreiten. Die gefundenen Schwellenwerte für die Mitoserate passen ebenso gut zu den in der Praxis verwendeten Schwellenwerten von 5 und 10 pro 50 hpf, wobei sich in der rezidivfreien Überlebensdifferenzanalyse für nicht-gastrische GIST eine breite Übergangszone zeigt, was andeutet, dass die Mitoserate in dieser Gruppe keinen numerisch scharf zu definierenden Parameter darstellt (Abbildung 6H). Für das pathologische Gutachten ist es daher wichtig, genaue Werte für die gemessenen

Risikoparameter anzugeben (statt z. B.  $_{\sim} \leq 5$  cm<sup>4</sup>) und diese Parameter im Kontext der klinischen Daten, d. h. Tumorlokalisation, Alter und Co-Morbiditäten des Patienten, sowie in Kenntnis der teils bedingten Trennschärfe numerischer Schwellenwerte zu interpretieren.

Die rezidivfreien Überlebensdifferenzanalysen deuten nicht darauf hin, dass für Magen-GIST und nicht-gastrische GIST unterschiedliche numerische Schwellenwerte verwendet werden sollten. Mit Ausnahme des ursprünglichen AFIP-Schemas und des TNM-Schemas würden jedoch mehrere Schemata (NIH-Schema, modifiziertes NIH-Schema, modifizierte AFIP-Schemata) von einer zusätzlichen Berücksichtigung der Lokalisation profitieren (Tabelle 3). Selbst wenn wie im Fall der modifizierten AFIP-Schemata die Lokalisation bereits in die Dignitätsbeurteilung eingeht, sollte daher in der individuellen Bewertung eines Falles der Tumorlokalisation eine besondere Geltung zukommen, da selbst bei lokalisationsabhängig getrennten Risikobewertungsschemata nicht automatisch davon ausgegangen werden kann, dass die Lokalisation bereits mit dem vollständigen statistischen Informationsgehalt in das jeweilige Schema eingeht.

#### 4.2 Bewertung der Risikoschemata

Gegenüber den NIH-Klassifikationen bewerten die auf den AFIP-Gruppen aufbauenden Schemata (AFIP- und modifizierte AFIP-Klassifikation, TNM-Klassifikation) die Mitoserate inhärent höher als die Tumorgröße, da ein Überschreiten des 5 pro 50 hpf-Schwellenwertes eine bei gleicher Tumorgröße drei numerische Gruppen höhere Einordnung bewirkt (Abbildung 1A).

Im vorliegenden Fallkollektiv lassen sich Sprünge im rezidivfreien Überleben bezogen auf die AFIP-Gruppen nachweisen, insbesondere zwischen den AFIP-Gruppen 3a und 3b (Abbildung 7E) sowie zwischen den AFIP-Gruppen 3a/4 und 5 (Abbildung 7F). Da in der modifizierten NIH-Klassifikation die Gruppen 3a, 4 und 5 dem (zweitniedrigsten) Risikolevel "II" zugeordnet werden, ist es daher wichtig, bei Verwendung dieses Schemas auch die AFIP-Gruppe anzugeben. Der Sprung zwischen den AFIP-Gruppen 3a und 3b deutet darüber

hinaus an, dass für niedrig proliferierende GIST der 10 cm-Schwellenwert eine besondere Bedeutung in der Risikovorhersage hat. Neuere Klassifikationssysteme (die modifizierten NIH- und AFIP-Klassifikationen und die TNM-Klassifikation) zeigen generell eine bessere Linearität zwischen den Risikoniveaus und insbesondere einen auch im vorliegenden Fallkollektiv (n = 161 Fälle) nachvollziehbaren Unterschied zwischen dem niedrigsten und zweitniedrigsten Risikolevel, was konsistent mit Beobachtungen an anderen Fallkollektiven ist (Goh et al. 2008; Schmieder et al. 2016). In neueren Klassifikationen wird zudem der 2 cm-Schwellenwert für die Tumorgröße nur noch teilweise verwendet (für Magen-GIST in den modifizierten AFIP-Klassifikationen sowie für nicht-gastrische GIST in der TNM-Klassifikation; siehe Abbildung 1), was gerechtfertigt erscheint, da GIST ab einer Größe von 2 cm operativ therapiert werden und sogar eine Perforation aufweisen können. Eine statistische Analyse der Wertigkeit dieses niedrigen Schwellenwertes ist limitiert durch die Seltenheit beobachteter Rezidive in der Gruppe dieser kleinen Tumoren. In der hier durchgeführten Analyse mittels rekursiver Partitionierung und Ermittlung von Risikoklassifikationsbäumen wird der 2 cm-Schwellenwert bei wiederholter Durchführung dieser Analyse an Zufallsstichproben aus dem Datensatz (bootstrapping) in ca. 2/3 der erhaltenen Modelle berücksichtigt, was andeutet, dass dieser Schwellenwert eine auf wenige Fälle bezogene statistische Bedeutung hat.

In den AFIP- und NIH-Klassifikationen gibt es zudem eine Diskrepanz für sehr kleine Magen-GIST mit erhöhter mitotischer Aktivität, die nach AFIP-Klassifikation als *"low"*, nach NIH-Klassifikation als *"intermediate"* eingeordnet werden. Das vorliegenden Fallkollektiv enthält drei Tumoren von 2 cm oder kleiner mit einer Mitoserate > 5 pro 50 hpf, darunter zwei Magen-GIST und ein nicht-gastrischer GIST, die bei einer Beobachtungszeit von jeweils > 5 Jahren kein Rezidiv gezeigt haben. Wegen der geringen Fallzahl ist eine weitere Analyse kleiner GIST mit hoher Proliferationsaktivität an größeren Fallkollektiven notwendig.

In der NIH-Klassifikation würde eine Berücksichtigung der Tumorlokalisation in der Hochrisiko-Klasse zu einer besseren Auftrennung führen. So traten innerhalb der 31 Fälle mit Progress in der NIH-Risikogruppe *"high"* (14 Magen-GIST und 17 nicht-gastrische GIST) Rezidive im Mittel ca. 8 Monate früher in der nicht-gastrischen Gruppe als in der Gruppe der Magen-GIST auf.

### 4.3 Anzahl chromosomaler Nettoveränderungen als prognostischer Parameter

Zur Frage, ob molekularpathologische Analysen einen unabhängigen prognostischen Nutzen zur Risikovorhersage beitragen, sind die mittels konventioneller CGH erhobenen zytogenetischen Daten bemerkenswert und unterstreichen noch einmal den besonderen Wert des Göttinger GIST-Kollektives. Insbesondere zeigt sich, dass die Anzahl chromosomaler Nettoveränderungen bezogen auf Chromosomenarme einen zusätzlichen bereits gegebener Tumorgröße, prognostischen Wert bei Mitoserate und Tumorlokalisation aufweist (Tabelle 5; p = 0,0003). Desweiteren zeigt sich in allen untersuchten Risikoklassifikationen bei Annahme des ermittelten optimalen Schwellenwertes von Veränderungen an fünf Chromosomenarmen in den höheren Risikoniveaus ein zusätzlicher prognostischer Wert der Anzahl chromosomaler Nettoveränderungen, selbst wenn eine zusätzliche Auftrennung nach Lokalisation keinen weiteren prognostischen Nutzen hat (Tabelle 6). Allerdings werden zytogenetische Daten in der diagnostischen Routine nicht standardmäßig erhoben, während pathologische Basisparameter wie die Tumorgröße und Mitoserate mit einfachen Mitteln während der Standardaufarbeitung bestimmt werden können und somit die Risikoeinordnung in jedem pathologischen Institut ermöglichen. Eine Berücksichtigung von im Vergleich zu pathologischen Basisparametern aufwändigeren molekularpathologischen Parametern in der Dignitätsbeurteilung könnte zum Beispiel Bedeutung erlangen, wenn bei neueren Inhibitoren im Einzelfall der Nutzen (etwa bei individueller Unverträglichkeit oder Co-Morbiditäten) einer Therapie abgewogen werden muss.

Im klinischen Alltag wird bisher eine Mutationsanalyse zur individualisierten Therapieplanung durchgeführt, wobei für einige Treibermutationen auch Daten hinsichtlich einer Bewertung des individuellen Rezidivrisikos vorliegen (Jiang et al. 2016; Lasota et al. 2004; Wardelmann et al. 2003; Wozniak et al. 2012). Weitere Verbesserungen der Prognose könnten zukünftig durch die Berücksichtigung weiterer klinischer Parameter, z. B. intra- oder extraoperative Tumorperforation (Joensuu et al. 2012) sowie molekularer Parameter, etwa dem Nachweis chromosomaler Bruchpunkte (Gunawan et al. 2007), erreicht werden.

### 5 Zusammenfassung

Gastrointestinale Stromatumoren (GIST) sind die häufigsten mesenchymalen Neoplasien des Gastrointestinaltraktes mit variablem Malignitätspotential. Aktuelle Risikoklassifikationen berücksichtigen die Tumorgröße, Mitoserate und teils auch die Tumorlokalisation als Risikofaktoren, wobei zur Stratifizierung des Risikos die Definition von Schwellenwerten für die numerischen Parameter zu erfolgen hat. In der vorliegenden Arbeit wurde die Wertigkeit der kanonischen Schwellenwerte (2 cm, 5 cm und 10 cm für die Tumorgröße; 5 Mitosen pro 50 hpf bzw. 5 mm<sup>2</sup> Tumorgewebe und 10 Mitosen pro 50 hpf bzw. 5 mm<sup>2</sup> Tumorgewebe für die Mitoserate) mittels geeigneter statistischer Verfahren zur Optimierung numerischer Schwellenwerte für Risikoparameter untersucht. Die Analysen erfolgten am Göttinger GIST-Fallkollektiv, das bis in die Frühzeit der GIST-Diagnostik zurückreicht und Follow-up-Daten von Patienten enthält, die keine heute bei hohem Rezidivrisiko übliche individualisierte Therapie mit Tyrosinkinaseinhibitoren erhalten haben. In den verwendeten statistischen Verfahren wurden konsistent Schwellenwerte von 4-5 cm und 10-11 cm für die Tumorgröße als optimale Schwellenwerte ermittelt, sowie Werte von 5 pro 50 hpf und 10 pro 50 hpf für die Mitoserate, was die kanonischen, zur Risikostratifizierung in der klinischen Praxis verwendeten Schwellenwerte validiert. Breite Risikoübergänge werden für die Mitoserate, insbesondere bei nichtgastrischen GIST beobachtet als Hinweis auf die Bedeutung einer exakten pathologischen Berichterstattung und einer Beurteilung der Risikoparameter im Kontext der erreichbaren statistischen Trennbarkeit zur Planung einer leitliniengerechten Therapie. Neuere, modifizierte Versionen der bekannten NIH- und AFIP-Klassifikationen unterschieden sich von den älteren Schemata vor allem durch eine verbesserte Linearität der Risikoniveaus. Zusätzlich zeigte sich die Anzahl klonaler Nettoveränderungen bezogen auf Chromosomenarme als unabhängiger prognostischer Parameter bei gegebener Tumorgröße, Mitoserate und Tumorlokalisation als Hinweis auf einen möglichen zusätzlichen Nutzen bisher noch nicht für die Risikovorhersage verwendeter molekularpathologischer Daten in der Risikoklassifikation von GIST.

### 6 Literaturverzeichnis

Agaimy A (2010): Gastrointestinal stromal tumors (GIST) from risk stratification systems to the new TNM proposal: more questions than answers? A review emphasizing the need for a standardized GIST reporting. Int J Clin Exp Pathol <u>3</u>, 461-471

Agaimy A, Otto C, Braun A, Geddert H, Schaefer IM, Haller F (2013): Value of epithelioid morphology and PDGFRA immunostaining pattern for prediction of PDGFRA mutated genotype in gastrointestinal stromal tumors (GISTs). Int J Clin Exp Pathol <u>6</u>, 1839-1846

Agaimy A, Wunsch PH, Hofstaedter F, Blaszyk H, Rummele P, Gaumann A, Dietmaier W, Hartmann A (2007): Minute gastric sclerosing stromal tumors (GIST tumorlets) are common in adults and frequently show c-KIT mutations. Am J Surg Pathol <u>31</u>, 113-120

Agaram NP, Wong GC, Guo T, Maki RG, Singer S, Dematteo RP, Besmer P, Antonescu CR (2008): Novel V600E BRAF mutations in imatinib-naive and imatinib-resistant gastrointestinal stromal tumors. Genes Chromosomes Cancer <u>47</u>, 853-859

Breiman L, Friedman J, Stone CJ, Olshen RA: Classification and regression trees; 1. Auflage; Chapman & Hall/CRC, Boca Raton, Florida 1984

Brenca M, Rossi S, Polano M, Gasparotto D, Zanatta L, Racanelli D, Valori L, Lamon S, Dei Tos AP, Maestro R (2016): Transcriptome sequencing identifies ETV6-NTRK3 as a gene fusion involved in GIST. J Pathol <u>238</u>, 543-549

Brierley J, Gospodarowicz MK, Wittekind C: TNM classification of malignant tumours; 8. Auflage; John Wiley & Sons, Inc., Chichester, West Sussex, United Kingdom; Hoboken, New Jersey 2017

Carballo M, Roig I, Aguilar F, Pol MA, Gamundi MJ, Hernan I, Martinez-Gimeno M (2005): Novel c-KIT germline mutation in a family with gastrointestinal stromal tumors and cutaneous hyperpigmentation. Am J Med Genet A <u>132A</u>, 361-364

Carney JA, Sheps SG, Go VL, Gordon H (1977): The triad of gastric leiomyosarcoma, functioning extra-adrenal paraganglioma and pulmonary chondroma. N Engl J Med <u>296</u>, 1517-1518

Casali PG, Abecassis N, Aro HT, Bauer S, Biagini R, Bielack S, Bonvalot S, Boukovinas I, Bovee J, Brodowicz T et al. (2018): Gastrointestinal stromal tumours: ESMO-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol <u>29</u>, iv68-iv78

Cheng SP, Huang MJ, Yang TL, Tzen CY, Liu CL, Liu TP, Hsiao SC (2004): Neurofibromatosis with gastrointestinal stromal tumors: insights into the association. Dig Dis Sci <u>49</u>, 1165-1169

Corless CL, Schroeder A, Griffith D, Town A, McGreevey L, Harrell P, Shiraga S, Bainbridge T, Morich J, Heinrich MC (2005): PDGFRA mutations in gastrointestinal stromal tumors: frequency, spectrum and in vitro sensitivity to imatinib. J Clin Oncol <u>23</u>, 5357-5364

Costa J, Wesley RA, Glatstein E, Rosenberg SA (1984): The grading of soft tissue sarcomas. Results of a clinicohistopathologic correlation in a series of 163 cases. Cancer <u>53</u>, 530-541

Cox DR (1972): Regression Models and Life-Tables. Journal of the Royal Statistical Society: Series B (Methodological) <u>34</u>, 187-202

Debiec-Rychter M, Sciot R, Le Cesne A, Schlemmer M, Hohenberger P, van Oosterom AT, Blay JY, Leyvraz S, Stul M, Casali PG et al. (2006): KIT mutations and dose selection for imatinib in patients with advanced gastrointestinal stromal tumours. Eur J Cancer <u>42</u>, 1093-1103

Demetri GD, van Oosterom AT, Garrett CR, Blackstein ME, Shah MH, Verweij J, McArthur G, Judson IR, Heinrich MC, Morgan JA et al. (2006): Efficacy and safety of sunitinib in patients with advanced gastrointestinal stromal tumour after failure of imatinib: a randomised controlled trial. Lancet <u>368</u>, 1329-1338

Demetri GD, Reichardt P, Kang YK, Blay JY, Rutkowski P, Gelderblom H, Hohenberger P, Leahy M, von Mehren M, Joensuu H et al. (2013): Efficacy and safety of regorafenib for advanced gastrointestinal stromal tumours after failure of imatinib and sunitinib (GRID): an international, multicentre, randomised, placebo-controlled, phase 3 trial. Lancet <u>381</u>, 295-302

Druker BJ, Sawyers CL, Kantarjian H, Resta DJ, Reese SF, Ford JM, Capdeville R, Talpaz M (2001a): Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. N Engl J Med <u>344</u>, 1038-1042

Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, Lydon NB, Kantarjian H, Capdeville R, Ohno-Jones S et al. (2001b): Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med <u>344</u>, 1031-1037

El-Rifai W, Sarlomo-Rikala M, Andersson LC, Knuutila S, Miettinen M (2000): DNA sequence copy number changes in gastrointestinal stromal tumors: tumor progression and prognostic significance. Cancer Res <u>60</u>, 3899-3903

Espinosa I, Lee CH, Kim MK, Rouse BT, Subramanian S, Montgomery K, Varma S, Corless CL, Heinrich MC, Smith KS et al. (2008): A novel monoclonal antibody against DOG1 is a sensitive and specific marker for gastrointestinal stromal tumors. Am J Surg Pathol <u>32</u>, 210-218

Falchook GS, Trent JC, Heinrich MC, Beadling C, Patterson J, Bastida CC, Blackman SC, Kurzrock R (2013): BRAF mutant gastrointestinal stromal tumor: first report of regression with BRAF inhibitor dabrafenib (GSK2118436) and whole exomic sequencing for analysis of acquired resistance. Oncotarget <u>4</u>, 310-315

Fehrmann RS, Li XY, van der Zee AG, de Jong S, Te Meerman GJ, de Vries EG, Crijns AP (2007): Profiling studies in ovarian cancer: a review. Oncologist <u>12</u>, 960-966

Fletcher CD, Berman JJ, Corless C, Gorstein F, Lasota J, Longley BJ, Miettinen M, O'Leary TJ, Remotti H, Rubin BP et al. (2002): Diagnosis of gastrointestinal stromal tumors: A consensus approach. Hum Pathol <u>33</u>, 459-465

Franquemont DW (1995): Differentiation and risk assessment of gastrointestinal stromal tumors. Am J Clin Pathol <u>103</u>, 41-47

Goh BK, Chow PK, Yap WM, Kesavan SM, Song IC, Paul PG, Ooi BS, Chung YF, Wong WK (2008): Which is the optimal risk stratification system for surgically treated localized primary GIST? Comparison of three contemporary prognostic criteria in 171 tumors and a proposal for a modified Armed Forces Institute of Pathology risk criteria. Ann Surg Oncol 15, 2153-2163

Graadt van Roggen JF, van Velthuysen ML, Hogendoorn PC (2001): The histopathological differential diagnosis of gastrointestinal stromal tumours. J Clin Pathol <u>54</u>, 96-102

Gunawan B, von Heydebreck A, Sander B, Schulten HJ, Haller F, Langer C, Armbrust T, Bollmann M, Gasparov S, Kovac D et al. (2007): An oncogenetic tree model in gastrointestinal stromal tumours (GISTs) identifies different pathways of cytogenetic evolution with prognostic implications. J Pathol <u>211</u>, 463-470

Hirota S, Ohashi A, Nishida T, Isozaki K, Kinoshita K, Shinomura Y, Kitamura Y (2003): Gainof-function mutations of platelet-derived growth factor receptor alpha gene in gastrointestinal stromal tumors. Gastroenterology <u>125</u>, 660-667

Huang HY, Li CF, Huang WW, Hu TH, Lin CN, Uen YH, Hsiung CY, Lu D (2007): A modification of NIH consensus criteria to better distinguish the highly lethal subset of primary localized gastrointestinal stromal tumors: a subdivision of the original high-risk group on the basis of outcome. Surgery <u>141</u>, 748-756

Jiang Z, Zhang J, Li Z, Liu Y, Wang D, Han G (2016): A meta-analysis of prognostic value of KIT mutation status in gastrointestinal stromal tumors. Onco Targets Ther <u>9</u>, 3387-3398

Joensuu H (2008): Risk stratification of patients diagnosed with gastrointestinal stromal tumor. Hum Pathol <u>39</u>, 1411-1419

Joensuu H, Roberts PJ, Sarlomo-Rikala M, Andersson LC, Tervahartiala P, Tuveson D, Silberman S, Capdeville R, Dimitrijevic S, Druker B et al. (2001): Effect of the tyrosine kinase inhibitor STI571 in a patient with a metastatic gastrointestinal stromal tumor. N Engl J Med <u>344</u>, 1052-1056

Joensuu H, Vehtari A, Riihimaki J, Nishida T, Steigen SE, Brabec P, Plank L, Nilsson B, Cirilli C, Braconi C et al. (2012): Risk of recurrence of gastrointestinal stromal tumour after surgery: an analysis of pooled population-based cohorts. Lancet Oncol <u>13</u>, 265-274

Kaemmer DA, Otto J, Lassay L, Steinau G, Klink C, Junge K, Klinge U, Schumpelick V (2009): The Gist of literature on pediatric GIST: review of clinical presentation. J Pediatr Hematol Oncol <u>31</u>, 108-112

Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D (1992): Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science <u>258</u>, 818-821

Kaplan EL, Meier P (1958): Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association <u>53</u>, 457-481

Kawanowa K, Sakuma Y, Sakurai S, Hishima T, Iwasaki Y, Saito K, Hosoya Y, Nakajima T, Funata N (2006): High incidence of microscopic gastrointestinal stromal tumors in the stomach. Hum Pathol <u>37</u>, 1527-1535

Kinoshita K, Hirota S, Isozaki K, Ohashi A, Nishida T, Kitamura Y, Shinomura Y, Matsuzawa Y (2004): Absence of c-kit gene mutations in gastrointestinal stromal tumours from neurofibromatosis type 1 patients. J Pathol <u>202</u>, 80-85

Kinoshita K, Isozaki K, Hirota S, Nishida T, Chen H, Nakahara M, Nagasawa Y, Ohashi A, Shinomura Y, Kitamura Y et al. (2003): c-kit gene mutation at exon 17 or 13 is very rare in sporadic gastrointestinal stromal tumors. J Gastroenterol Hepatol <u>18</u>, 147-151

Krishnamurty R, Maly DJ (2010): Biochemical mechanisms of resistance to small-molecule protein kinase inhibitors. ACS Chem Biol <u>5</u>, 121-138

Lasota J, Miettinen M (2006): KIT and PDGFRA mutations in gastrointestinal stromal tumors (GISTs). Semin Diagn Pathol <u>23</u>, 91-102

Lasota J, Dansonka-Mieszkowska A, Sobin LH, Miettinen M (2004): A great majority of GISTs with PDGFRA mutations represent gastric tumors of low or no malignant potential. Lab Invest <u>84</u>, 874-883

Lausen B, Hothorn T, Bretz F, Schumacher M (2004): Assessment of optimal selected prognostic factors. Biometrical J <u>46</u>, 364-374

Lennartsson J, Ronnstrand L (2012): Stem cell factor receptor/c-Kit: from basic science to clinical implications. Physiol Rev <u>92</u>, 1619-1649

Li FP, Fletcher JA, Heinrich MC, Garber JE, Sallan SE, Curiel-Lewandrowski C, Duensing A, van de Rijn M, Schnipper LE, Demetri GD (2005): Familial gastrointestinal stromal tumor syndrome: phenotypic and molecular features in a kindred. J Clin Oncol <u>23</u>, 2735-2743

Liegl B, Hornick JL, Corless CL, Fletcher CD (2009): Monoclonal antibody DOG1.1 shows higher sensitivity than KIT in the diagnosis of gastrointestinal stromal tumors, including unusual subtypes. Am J Surg Pathol <u>33</u>, 437-446

Maeda H, Yamagata A, Nishikawa S, Yoshinaga K, Kobayashi S, Nishi K, Nishikawa S (1992): Requirement of c-kit for development of intestinal pacemaker system. Development <u>116</u>, 369-375

McWhinney SR, Pasini B, Stratakis CA, International Carney T, Carney-Stratakis Syndrome C (2007): Familial gastrointestinal stromal tumors and germ-line mutations. N Engl J Med <u>357</u>, 1054-1056

Medeiros F, Corless CL, Duensing A, Hornick JL, Oliveira AM, Heinrich MC, Fletcher JA, Fletcher CD (2004): KIT-negative gastrointestinal stromal tumors: proof of concept and therapeutic implications. Am J Surg Pathol <u>28</u>, 889-894

Miettinen M, Sarlomo-Rikala M, Lasota J (1999): Gastrointestinal stromal tumors: recent advances in understanding of their biology. Hum Pathol <u>30</u>, 1213-1220

Miettinen M, El-Rifai W, L HLS, Lasota J (2002): Evaluation of malignancy and prognosis of gastrointestinal stromal tumors: a review. Hum Pathol <u>33</u>, 478-483

Miettinen M, Makhlouf H, Sobin LH, Lasota J (2006): Gastrointestinal stromal tumors of the jejunum and ileum: a clinicopathologic, immunohistochemical, and molecular genetic study of 906 cases before imatinib with long-term follow-up. Am J Surg Pathol <u>30</u>, 477-489

Nishida T, Hirota S (2000): Biological and clinical review of stromal tumors in the gastrointestinal tract. Histol Histopathol <u>15</u>, 1293-1301

Orford JL, Lennon R, Melby S, Fasseas P, Bell MR, Rihal CS, Holmes DR, Berger PB (2002): Frequency and correlates of coronary stent thrombosis in the modern era: analysis of a single center registry. J Am Coll Cardiol <u>40</u>, 1567-1572

Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y et al. (1998): High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet <u>20</u>, 207-211

Quiroz HJ, Willobee BA, Sussman MS, Fox BR, Thorson CM, Sola JE, Perez EA (2018): Pediatric gastrointestinal stromal tumors-a review of diagnostic modalities. Transl Gastroenterol Hepatol <u>3</u>, 54

R-Core-Team: R: A language and environment for statistical computing. hrsg. https://www.R-project.org/ R Foundation for Statistical Computing, Wien, Österreich 2019

Reith JD, Goldblum JR, Lyles RH, Weiss SW (2000): Extragastrointestinal (soft tissue) stromal tumors: an analysis of 48 cases with emphasis on histologic predictors of outcome. Mod Pathol <u>13</u>, 577-585

Rubin JL, Sanon M, Taylor DC, Coombs J, Bollu V, Sirulnik L (2011): Epidemiology, survival, and costs of localized gastrointestinal stromal tumors. Int J Gen Med <u>4</u>, 121-130

Sander B, Cameron S, Gunawan B, Fuzesi L (2020): Optimal thresholds of risk parameters for gastrointestinal stromal tumors. Eur J Surg Oncol <u>46</u>, 180-188

Schaefer IM, Delfs C, Cameron S, Gunawan B, Agaimy A, Ghadimi BM, Haller F (2014): Chromosomal aberrations in primary PDGFRA-mutated gastrointestinal stromal tumors. Hum Pathol <u>45</u>, 85-97

Schmieder M, Henne-Bruns D, Mayer B, Knippschild U, Rolke C, Schwab M, Kramer K (2016): Comparison of Different Risk Classification Systems in 558 Patients with Gastrointestinal Stromal Tumors after RO-Resection. Front Pharmacol <u>7</u>, 504

Schroeder BC, Cheng T, Jan YN, Jan LY (2008): Expression cloning of TMEM16A as a calciumactivated chloride channel subunit. Cell <u>134</u>, 1019-1029

Settas N, Faucz FR, Stratakis CA (2018): Succinate dehydrogenase (SDH) deficiency, Carney triad and the epigenome. Mol Cell Endocrinol <u>469</u>, 107-111

Shirai H, Takeuchi T, Naka T, Minaghi S, Kimura A, Hamazaki S, Ito H (2001): Gastrointestinal stromal tumor of the stomach: report of a case. Surg Today <u>31</u>, 346-349

Sircar K, Hewlett BR, Huizinga JD, Chorneyko K, Berezin I, Riddell RH (1999): Interstitial cells of Cajal as precursors of gastrointestinal stromal tumors. Am J Surg Pathol <u>23</u>, 377-389

Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P (1997): Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosomes Cancer <u>20</u>, 399-407

Soreide K, Sandvik OM, Soreide JA, Giljaca V, Jureckova A, Bulusu VR (2016): Global epidemiology of gastrointestinal stromal tumours (GIST): A systematic review of population-based cohort studies. Cancer Epidemiol <u>40</u>, 39-46

Tamborini E, Bonadiman L, Greco A, Albertini V, Negri T, Gronchi A, Bertulli R, Colecchia M, Casali PG, Pierotti MA et al. (2004): A new mutation in the KIT ATP pocket causes acquired resistance to imatinib in a gastrointestinal stromal tumor patient. Gastroenterology <u>127</u>, 294-299

Tran T, Davila JA, El-Serag HB (2005): The epidemiology of malignant gastrointestinal stromal tumors: an analysis of 1,458 cases from 1992 to 2000. Am J Gastroenterol <u>100</u>, 162-168

Trojani M, Contesso G, Coindre JM, Rouesse J, Bui NB, de Mascarel A, Goussot JF, David M, Bonichon F, Lagarde C (1984): Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. Int J Cancer <u>33</u>, 37-42

Tuveson DA, Willis NA, Jacks T, Griffin JD, Singer S, Fletcher CD, Fletcher JA, Demetri GD (2001): STI571 inactivation of the gastrointestinal stromal tumor c-KIT oncoprotein: biological and clinical implications. Oncogene <u>20</u>, 5054-5058

Wardelmann E, Losen I, Hans V, Neidt I, Speidel N, Bierhoff E, Heinicke T, Pietsch T, Buttner R, Merkelbach-Bruse S (2003): Deletion of Trp-557 and Lys-558 in the juxtamembrane domain of the c-kit protooncogene is associated with metastatic behavior of gastrointestinal stromal tumors. Int J Cancer <u>106</u>, 887-895

West RB, Corless CL, Chen X, Rubin BP, Subramanian S, Montgomery K, Zhu S, Ball CA, Nielsen TO, Patel R et al. (2004): The novel marker, DOG1, is expressed ubiquitously in gastrointestinal stromal tumors irrespective of KIT or PDGFRA mutation status. Am J Pathol <u>165</u>, 107-113

Worsley KJ (1982): An Improved Bonferroni Inequality and Applications. Biometrika <u>69</u>, 297-302

Wozniak A, Sciot R, Guillou L, Pauwels P, Wasag B, Stul M, Vermeesch JR, Vandenberghe P, Limon J, Debiec-Rychter M (2007): Array CGH analysis in primary gastrointestinal stromal tumors: cytogenetic profile correlates with anatomic site and tumor aggressiveness, irrespective of mutational status. Genes Chromosomes Cancer <u>46</u>, 261-276

Wozniak A, Rutkowski P, Piskorz A, Ciwoniuk M, Osuch C, Bylina E, Sygut J, Chosia M, Rys J, Urbanczyk K et al. (2012): Prognostic value of KIT/PDGFRA mutations in gastrointestinal stromal tumours (GIST): Polish Clinical GIST Registry experience. Ann Oncol <u>23</u>, 353-360

## Danksagung

Ich danke herzlich Herrn Prof. Dr. Füzesi für die Betreuung der Arbeit und die Idee des Themas, sowie Herrn PD Dr. Gunawan und Frau PD Dr. Cameron für ihre Unterstützung. Frau Dr. von Heydebreck bin ich für die Einführung in die hier verwendeten statistischen Verfahren sehr dankbar. Die Publikation "Optimal thresholds of risk parameters for gastrointestinal stromal tumors" von B. Sander, S. Cameron, B. Gunawan, L. Füzesi in European Journal of Surgical Oncology: the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 2020;46(1):180-188 ist unter

https://doi.org/10.1016/j.ejso.2019.08.009

mit institutionellem Login zugänglich.

Die Publikation "An oncogenetic tree model in gastrointestinal stromal tumours (GISTs) identifies different pathways of cytogenetic evolution with prognostic implications" von B. Gunawan, A. von Heydebreck, B. Sander, H.-J. Schulten, F. Haller, C. Langer, T. Armbrust, M. Bollmann, S. Gasparov, D. Kovac, L. Füzesi in The Journal of Pathology 2007;211(4):463-70 ist unter

https://doi.org/10.1002/path.2128

mit institutionellem Login zugänglich.