



Technische Universität München

Institut für diagnostische und interventionelle Radiologie

Klinikum rechts der Isar

**Nicht-invasive Subtypisierung des Pankreaskarzinoms mittels
Algorithmen des maschinellen Lernens**

Sebastian Ziegelmayer

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Medizin (Dr.med.) genehmigten Dissertation.

Vorsitzender: Prof. Dr. Jürgen Schlegel

Prüfer der Dissertation:

1. Priv.-Doz. Dr. Rickmer Braren
2. Prof. Dr. Oliver Hayden

Die Dissertation wurde am 10.08.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 16.02.2021 angenommen.

Inhaltsverzeichnis

1	Originalarbeiten	III
2	Abkürzungsverzeichnis	IV
3	Einleitung	1
3.1	Das Pankreatische Duktale Adenokarzinom	1
3.2	Tumorale Heterogenität.....	4
3.3	Funktionsweisen der Bildgebungsmodalitäten	7
3.4	Radiomics.....	9
3.5	Maschinelles Lernen.....	11
3.6	Zielsetzung der Arbeit.....	12
4	Methoden	13
4.1	Studiendesign.....	13
4.2	Klinische und histologische Parameter.....	14
4.3	Bestimmung des immunhistochemischen Subtyps	17
4.4	Bildgebung.....	18
4.5	Segmentierung der Pankreaskarzinome	19
4.6	Extraktion der Merkmale.....	20
4.7	Präprozessierung der extrahierten Merkmale	21
4.8	<i>Machine-learning</i> -Modellierung	24
5	Zusammenfassung und Ergebnisse der veröffentlichten Publikationen	28
5.1	Publikation I	28
5.2	Publikation II	30
5.3	Schlussfolgerung	31
6	Diskussion	32
6.1	Diskussion der Methoden	32
6.2	Diskussion der Ergebnisse.....	39
6.3	Limitationen	41
7	Ausblick	43
8	Zusammenfassung	44
9	Literaturverzeichnis	46
10	Abbildungsverzeichnis	67
11	Tabellenverzeichnis	68
12	Danksagung	69

1 Originalarbeiten

Die Methoden und Ergebnisse der vorliegenden Arbeit wurden veröffentlicht in:

Publikation I:

A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response to gemcitabine-based versus FOLFIRINOX chemotherapy.

Kaissis G*, **Ziegelmayer S***, Lohöfer F*, et al. PLoS One.

Veröffentlicht: 2. Oktober 2019

DOI:10.1371/journal.pone.0218642

Publikation II:

Image-Based Molecular Phenotyping of Pancreatic Ductal Adenocarcinoma.

Kaissis G*, **Ziegelmayer S***, Lohöfer F*, et al. Journal of Clinical Medicine.

Veröffentlicht: 7. März 2020

DOI:10.3390/jcm9030724

* geteilte Erstautorenschaft

2 Abkürzungsverzeichnis

ADC	engl. <i>apparent diffusion coefficient</i>
ADEX	abberant-differenziert-endokrin-exokriner Subtyp
CA-19-9.....	Carbohydrate-Antigen 19-9
CEA.....	Carcinoembryonales Antigen
CT	Computertomographie
ECOG	Eastern Cooperative Oncology Group
G	Grading
HNF1a.....	Hepatozyten-Kernfaktor 1 alpha
HU.....	Hounsfieldeinheit
ICC.....	Intraklassen-Korrelationskoeffizient
KRT81	Zytokeratin 81
M	Fernmetastasierung
MRT	Magnetresonanztomographie
Non-QM	nicht quasi-mesenchymaler Subtyp
PanIn.....	pankreatische intraepitheliale Neoplasie
PDAC	engl. pancreatic ductal adenocarcinoma
pN	pathologischer Lymphknotenstatus
pT.....	pathologische Tumorausbreitung
QM	quasi-mesenchymaler Subtyp
R	Resektionsstatus
ROC-AUC	Fläche unter der Grenzwertoptimierungskurve
Voxel	engl. <i>volume picture element</i>

3 Einleitung

3.1 Das Pankreatische Duktale Adenokarzinom

3.1.1 Epidemiologie

Der Begriff „Pankreaskarzinom“ kann im Allgemeinen für unterschiedliche maligne Neoplasien der Bauchspeicheldrüse genutzt werden. Allerdings hat sich die Nomenklatur vorwiegend für das Pankreatische Duktale Adenokarzinom (*engl. pancreatic ductal adenocarcinoma*, PDAC), das 90-95% der bösartigen Neuerkrankungen ausmacht, durchgesetzt. Im Folgenden wird daher der Begriff „Pankreaskarzinom“ für diese Entität verwendet.

Das Pankreaskarzinom ist die vierthäufigste Krebstodesursache in Deutschland. Die Fünf-Jahres-Überlebensrate beträgt für Männer und Frauen neun Prozent (Rawla et al., 2019). Somit hat das Pankreaskarzinom die niedrigste Fünf-Jahres-Überlebensrate aller Krebserkrankungen. Die Anzahl der Neuerkrankungen und Sterbefälle hat in den letzten Jahrzehnten stetig zugenommen. Jährlich erkranken rund 18.000 Personen in Deutschland an einem Pankreaskarzinom, hierbei liegt das mittlere Erkrankungsalter für Männer bei 72 Jahren und für Frauen bei 75 Jahren (Zentrum für Krebsregisterdaten im Robert Koch-Institut, 2016).

3.1.2 Pathogenese und Risikofaktoren

Die Pathogenese des Pankreaskarzinoms ist ein komplexer, in Teilen noch unverstandener Prozess. Modellhaft kann die Pathogenese als eine fortschreitende Akkumulation von Mutationen in den dukталen Zellen des Pankreas verstanden werden, wobei es über präkanzerotische Veränderungen, wie die pankreatische intraepitheliale Neoplasie (PanIn) zur schrittweisen Zellentartung kommt (Yachida et al., 2010; Ying et al., 2016). Hierbei gehört das Pankreaskarzinom, ähnlich wie z.B. das Glioblastom, zu den Entitäten, die bereits früh eine ausgeprägte genetische, morphologische und biochemische intratumorale Heterogenität der Tumorzellklone und des Tumorstromas aufweisen (Bernard et al., 2019; Dirkse et al., 2019). Diese Heterogenität scheint zentral für die Therapieresistenz und die schlechte Prognose zu

sein und bildet die biologische Grundlage der vorliegenden Arbeit. Auf die tumorale Heterogenität wird deshalb gesondert in einem folgenden Abschnitt eingegangen.

Neben gängigen Risikofaktoren wie Alkoholkonsum, Rauchen, Adipositas und Diabetes mellitus begünstigen die chronische Pankreatitis und genetische Syndrome die Entstehung eines Pankreaskarzinoms (Bartsch et al., 2012; Kenner et al., 2016; Parkin et al., 2011). Ferner erhöhen Keimbahnmutationen, wie beispielsweise die Mutation von BRCA, das Erkrankungsrisiko. Der Nachweis dieser Mutation (auf Keimbahnebene oder somatisch) ermöglicht bei <10% der Patienten eine gezielte Tumorthherapie mit dem PARP-Inhibitor Olaparib (Golan et al., 2019).

3.1.3 Diagnostik und Therapie

Eine frühe Diagnosestellung wird durch die Symptomarmut des Tumors in frühen Stadien erschwert (Kelsen et al., 1997). Häufig sind unspezifische Symptome, wie Oberbauchschmerzen, Nachtschweiß, ungewollter Gewichtsverlust und Ikterus indirekte Hinweise auf das Vorliegen eines Pankreaskarzinoms.

Die Diagnosestellung erfolgt gemäß einer Stufendiagnostik (Leitlinienprogramm Onkologie, Deutsche Krebsgesellschaft, Deutsche Krebshilfe, 2013). Ergeben die Anamnese und die körperliche Untersuchung einen hinreichenden Verdacht, so erfolgt zunächst die sonographische Untersuchung des Abdomens. Zeigen sich hierbei keine richtungsweisenden Befunde, wird eine Computertomographie (CT) oder alternativ eine Magnetresonanztomographie (MRT) des Abdomens empfohlen, wobei letztere hauptsächlich in der lokoregionären Ausbreitungsdiagnostik zur Anwendung kommt und hier aufgrund des exzellenten Kontrastverhaltens der CT überlegen ist. Sehr hochauflösend kann die lokoregionäre Ausbreitung des Tumors durch eine endosonographische Untersuchung erfasst werden. Weiterhin wird ergänzend die Bestimmung der Tumormarker CA-19-9 und CEA im Serum empfohlen. Im Rahmen der Ausbreitungsdiagnostik, sowie präoperativ, wird die computertomographische Untersuchung des Thorax und des Abdomens mit Kontrastmittel empfohlen. Eine biopsische Sicherung des Tumors wird in inoperablen Fällen empfohlen (Luz et al., 2014).

An onkologischen Zentren werden die therapeutischen Optionen nach erfolgter Ausbreitungsdiagnostik in einer interdisziplinären Tumorkonferenz diskutiert und es wird eine Therapieempfehlung festgelegt. Die Resektion des Tumors ist bei fehlender

Metastasierung die einzige kurative Therapiemöglichkeit. Entscheidend für die Resektabilität eines Tumors sind neben der Tumorausdehnung und der Gefäßinfiltration durch den Tumor auch Allgemeinzustand und Komorbiditäten des Patienten (Tempero et al., 2017). Die gängigen Operationsmethoden sind die partielle Duodenopankreatektomie sowie die pyloruserhaltende Duodenopankreatektomie. Resektate werden pathologisch aufgearbeitet. Hierbei werden das TNM-Stadium, der Differenzierungsgrad, die Perineuralscheiden-, Lymphgefäß- und Blutgefäßinvasion sowie der Resektionsstatus des Tumors bestimmt. Ist zuvor eine Chemotherapie verabreicht worden, wird zusätzlich das Ansprechen im Sinne einer Tumorregression nach Evans beurteilt.

Für das Pankreaskarzinom besteht Evidenz, dass adjuvante und palliative Chemotherapie zu einem signifikanten Überlebensvorteil führt. Die Leitsubstanzen sind Gemcitabin und die Kombinationstherapie FOLFIRINOX, bestehend aus 5-Fluoruracil, Folinsäure, Irinotecan und Oxaliplatin. Sind die Resektionsränder tumorfrei, erhalten Patienten in kurativer Intention eine adjuvante Chemotherapie mit Gemcitabin. Der Stellenwert von Gemcitabin wurde bereits 2009 durch Ueno et al. begründet und 2013 durch Oettle et al. im Rahmen der multizentrischen CONKO-001-Studie bestätigt, die eine signifikante Verlängerung des progressionsfreien Überlebens und des Gesamtüberlebens durch Gemcitabin-Monotherapie zeigen konnte (Oettle et al., 2013; Ueno et al., 2009). Neuere Studien zeigen eine Verlängerung des Gesamtüberlebens und progressionsfreien Überlebens unter adjuvanter Therapie mit FOLFIRINOX gegenüber Gemcitabin, allerdings auf Kosten einer erhöhten Toxizität (Conroy et al., 2018). Zum jetzigen Zeitpunkt ist noch keine Aktualisierung der Leitlinie in Bezug auf diese Ergebnisse erfolgt.

Erweist sich der Tumor als nicht resektabel, oder finden sich Fernmetastasen in der Ausbreitungsdiagnostik, erhalten Patienten eine Chemotherapie in neoadjuvanter oder palliativer Intention. Nicht resezierbare oder metastasierte Tumoren werden mit einer Kombination aus Gemcitabin und nab-Paclitaxel, oder mit FOLFIRINOX therapiert. Die Therapie mit Gemcitabin und nab-Paclitaxel erwies sich hierbei effektiver als eine Gemcitabin-Monotherapie (Blomstrand et al., 2019; Von Hoff et al., 2013). Ebenso konnte bei Patienten im metastasierten Stadium beobachtet werden, dass unter Therapie mit FOLFIRINOX ein signifikant längeres Gesamtüberleben besteht, als unter einer Gemcitabin-Monotherapie (Conroy et al., 2011). Für neoadjuvante Therapieansätze besteht zunehmende Evidenz: In einer retrospektiven Studie konnte

durch neoadjuvante Therapie mit FOLFIRINOX die Wahrscheinlichkeit einer R0-Resektion signifikant erhöht werden (Scheufele et al., 2019). Weiterhin zeigte eine Phase-II-Studie, dass neoadjuvante Therapiekonzepte zu einer Erhöhung der Konversionsrate bei der chirurgischen Exploration führen und damit eine Verlängerung des Gesamtüberlebens erreicht werden kann (Kunzmann et al., 2018). Neuere Konzepte eruieren die Möglichkeit einer Resektion im Stadium der Oligometastasierung (Kandel et al., 2018).

Trotz der zunehmenden Intensivierung der Chemotherapieschemata und der stetigen Optimierung des operativen Vorgehens, konnte in den letzten 30 Jahren keine durchgreifende Verlängerung des Überlebens erreicht werden (Hall et al., 2018). Dieses Erkenntnis legt nahe, dass das Überleben durch weitere tumor- und patientenspezifische Faktoren beeinflusst wird, die offenbar über die momentan erfassten pathologischen und klinischen Parameter hinausgehen.

3.2 Tumorale Heterogenität

3.2.1 Das Evolutionsmodell der Karzinogenese

Die tumorale Heterogenität ist ein tumorentitätenübergreifendes Phänomen und beschreibt, dass sowohl zwischen einzelnen Tumoren wie auch innerhalb eines individuellen Tumors verschiedene Tumorzellsubklone existieren, welche genetische, metabolische und/oder strukturelle Unterschiede aufweisen. Dies konnte bereits für eine Vielzahl von Tumorentitäten nachgewiesen werden, beispielsweise für das Nierenzellkarzinom, das Glioblastom, das Kolonkarzinom und das Mammakarzinom (Gerlinger et al., 2012; Parsons et al., 2008; Sjöblom et al., 2006). Für die allgemeine Entstehung der tumoralen Heterogenität postulierte Nowell bereits 1976 das Prinzip der Tumorevolution. Die Grundannahme dieses Modells ist, dass sich im Rahmen des Tumorwachstums, bedingt durch genetische Instabilität, verschiedene Tumorzellsubklone aus einer einzelnen Tumorzelle entwickeln (Nowell, 1976). Die Evolution von Tumorzellklonen findet im Kontext der Anatomie und der Physiologie des ursprünglichen Gewebes statt, welche nach evolutionstheoretischen Gesichtspunkten einen Selektionsdruck auf den Tumorphänotyp darstellen (Gatenby & Gillies, 2008). Gemäß der Evolutionstheorie können sich Tumorzellsubklone, welche besser an das extratumorale Milieu angepasst sind, durchsetzen. Dieses

fundamentale Prinzip zeigt, dass spezifische externe Einflüsse, wie beispielsweise Chemotherapeutika einen Selektionsdruck auf die Tumorzellklone ausüben. Sind durch genetische Instabilität Klone entstanden, die resistent gegen das verwendete Chemotherapeutikum sind, so können diese selektiert werden und den Therapieerfolg erschweren (Faltas et al., 2016). Gemäß diesem Prinzip der Tumorevolution wird weiterhin deutlich, dass unterschiedliche Tumorzellklone innerhalb des Tumormilieus gleichzeitig proliferieren und koexistieren (Anderson et al., 2011; Maley et al., 2006). Basierend auf der genetischen Instabilität entstehen schlussfolgernd Tumoren, die intertumoral und intratumoral verschiedene Eigenschaften besitzen. Dies hat zur Folge, dass Tumoren sich in ihrem Therapieansprechen auf bestimmte Chemotherapeutika unterscheiden. Kovac et al. konnten weiterhin zeigen, dass treibende Mutationen nur in bestimmten Tumorebenen vorhanden waren (Kovac et al., 2015). Diese Entdeckung bedingt, dass bei Tumorbiopsien, die routinemäßig aus Einzelarealen erfolgen, die Heterogenität des Tumors nicht hinreichend abgebildet wird (Gerlinger et al., 2012). Zusammenfassend lässt sich die Tumorevolution als ein repetitiver Prozess der klonalen Expansion, der genetischen Diversifizierung und der klonalen Selektion im Zusammenspiel mit dem extratumoralen Milieu verstehen (Greaves & Maley, 2012).

Das Pankreaskarzinom ist eine außerordentlich heterogene Tumorentität, für die eine hohe genetische und epigenetische Varianz nachgewiesen werden konnte (N. A. Juiz et al., 2019; Roberts et al., 2016; Witkiewicz et al., 2015). Diese Alterationen führen zu vielfältigen Veränderungen der strukturellen und metabolischen Eigenschaften des Pankreaskarzinoms. Beispielweise zeigen sich Unterschiede in der desmoplastischen Reaktion von Pankreaskarzinomen, die zur Bildung unterschiedlicher Stromamengen führt (Torphy et al., 2018). Das gebildete Stroma besteht aus Fibroblasten, Myofibroblasten, Pankreassternzellen, Immunzellen, Blutgefäßen, extrazellulärer Matrix und löslichen Proteinen wie Zytokinen und Wachstumsfaktoren (Feig et al., 2012; Waghray et al., 2013). Die klinische Bedeutung dieser Stromareaktion ist noch nicht vollständig geklärt. Eine frühere Studie zeigte, dass die desmoplastische Reaktion sowohl die Perfusion wie auch die Anflutung von Therapeutika verschlechtert und als Barriere wirkt (Provenzano et al., 2012). Neuere Ergebnisse deuten allerdings darauf hin, dass die Stromareaktion möglicherweise auch protektiv wirkt und das Gesamtüberleben verlängert (Özdemir et al., 2014). Die Studie von Moffit et al. postuliert, dass bestimmte stromaspezifische Genexpressionscluster existieren, die

eine Unterscheidung zwischen „normalem“ und „aktiviertem“ Stroma erlauben, wobei das Überwiegen von Letzterem mit einer schlechteren Prognose einherzugehen scheint (Moffitt et al., 2015). Auffällig ist weiterhin, dass besagte Genexpressionscluster nicht von den Tumorzellen an sich, sondern von den stromal lokalisierten Zellen exprimiert werden. Dies lässt darauf schließen, dass auch die Tumorentwicklung des Pankreaskarzinoms im Wechselspiel mit dem extratumoralen Milieu stattfindet.

3.2.2 Die Subtypisierung des Pankreaskarzinoms

Durch die technischen Fortschritte in der Genexpressionsanalyse und der RNA-Sequenzierung wurden innerhalb der letzten Jahre zahlreiche morpho-molekulare und transkriptomische Subtypen des Pankreaskarzinoms definiert. Richtungsweisend war hierbei die Entdeckung von Collisson et al., welche zwischen dem quasi-mesenchymalen, dem klassischen und dem exokrin-ähnlichen Subtyp differenzieren (Collisson et al., 2011). Diesen Tumorsubtypen konnten charakteristische Gensignaturen zugewiesen werden, die mit einem unterschiedlichen Gesamtüberleben und Therapieansprechen korrelieren. Der klassische Subtyp exprimiert vor allem adhäsionsassoziierte und epitheliale Gene, wohingegen sich der quasi-mesenchymale Subtyp durch mesenchymassozierte Gene auszeichnet (Collisson et al., 2011). In einer späteren Studie konnten durch Bailey et al. mittels RNA-Sequenzierung vier Subtypen differenziert werden (Bailey et al., 2016). In dieser Studie konnte ein squamöser Subtyp, ein immunogener, ein aberrant-differenziert-endokrin-exokriner Subtyp (ADEX), sowie ein Pankreasvorläufer-Subtyp definiert werden. Eine weitere 2015 von Moffitt et al. vorgeschlagene Subtypisierung unterscheidet zwischen einem klassischen und einem basal-ähnlichen Subtyp (Moffitt et al., 2015). Für die postulierten Subtypen konnten deutliche Überlappungen festgestellt werden, insbesondere zwischen dem squamösen Subtyp, dem quasi-mesenchymalen Subtyp und dem basal-ähnlichen Subtyp sowie zwischen dem Pankreasvorläufer-Subtyp und dem klassischen Subtyp (Raphael et al., 2017). Diese Subtypen konnten in weiterführenden Studien mit der differentiellen Expression der Surrogatmarker Zytokeratin 81 (KRT81) und Hepatozyten-Kernfaktor 1 alpha (HNF1a) immunhistochemisch korreliert werden (Muckenhuber et al., 2018; Noll et al., 2016). KRT81 exprimierende Tumoren (KRT81+) zeigen ein hohes Maß an Korrelation mit dem quasi-mesenchymalen (QM) Subtyp und weisen ein schlechteres

Gesamtüberleben als KRT81 negative (KRT81-) Tumoren auf. Des Weiteren zeigte sich ein besseres Ansprechen dieser Tumorzellen auf Gemcitabin als auf FOLFIRINOX *in vitro* (Aung et al., 2018; Collisson et al., 2011). Für den exokrin-ähnlichen bzw. ADEX-Subtyp konnte HNF1a als Surrogatprotein bestimmt werden (KRT81-/HNF1a⁺). Im klassischen Subtyp lassen sich weder KRT81 noch HNF1a anfärben (KRT81-/HNF1a⁻). Allerdings stellen neue Erkenntnisse in Frage, ob der ADEX- bzw. der exokrin-ähnliche Subtyp tatsächliche Subtypen des Pankreaskarzinoms definieren, oder ob es sich hierbei um eine Kontamination mit Azinuszellen des normalen Pankreasparenchyms handelt (Puleo et al., 2018). Aufgrund dieser Erkenntnisse und der hohen prognostischen Relevanz des QM-Subtyps gegenüber den anderen Subtypen, wurde in der vorliegenden Arbeit zwischen quasi-mesenchymal (QM) und nicht-quasi-mesenchymal (Non-QM) differenzierten Tumoren unterschieden. Der Non-QM-Subtyp umfasst hierbei den klassischen und exokrin-ähnlichen Subtyp. Im Gegensatz zum QM-Subtyp weisen Patienten mit Non-QM-Subtyp ein längeres Gesamtüberleben auf und die Tumoren sind sensitiver auf eine Therapie mit FOLFIRINOX (Aung et al., 2018).

3.3 Funktionsweisen der Bildgebungsmodalitäten

3.3.1 Computertomographie

Die Computertomographie (CT) ist ein röntgenstrahlenbasiertes Schnittbildverfahren, welches eine überlagerungsfreie Darstellung der Gewebe und Organe im zweidimensionalen Raum erlaubt (Goldman, 2007). Durch die Erfassung mehrerer Schichten entsteht aus einem Bildpixel ein Volumen, welches als Voxel (engl. Abk. für *volume picture element*) bezeichnet wird. Ein Computertomograph besteht aus einer sog. *Gantry* (Gehäuse), in der die Röntgenröhre und das Detektorsystem integriert sind. Je nach Generation sind Detektorsystem und Röntgenröhre unterschiedlich angeordnet. Das Detektorsystem kann aus mehreren Detektorreihen aufgebaut sein und erfasst die Röntgenstrahlung, welche den Körper durchdringt. Hierbei wird für jeden Einfallswinkel ein Schwächungsprofil erfasst, aus dem jedem Voxel ein bestimmter Dichtewert zugeordnet werden kann. Die Bildrekonstruktion kann durch zwei Verfahren erfolgen, die gefilterte Rückprojektion und die iterative Rekonstruktion (Seeram, 2009). Das Ausmaß der Schwächung der Röntgenstrahlen ist abhängig von der Dichte des durchstrahlten Gewebes, der Ordnungszahl der Atome und von der Strahlenenergie und wird mit dem linearen Schwächungskoeffizienten beschrieben.

Diese Absorption kann durch den Einsatz von Kontrastmittel modifiziert werden. Um die Absorption der Gewebe vergleichen zu können, wird für die Computertomographie die Hounsfieldeinheit (HU) verwendet (Hounsfield, 1980). Diese ist einheitlich als der relative Schwächungskoeffizient des durchstrahlten Objekts im Verhältnis zum Schwächungskoeffizienten von Wasser definiert. Die Computertomographie wird in der medizinischen Bildgebung vielfältig eingesetzt. Insbesondere in der Ausbreitungsdiagnostik von Tumorerkrankungen ist sie ein fester Bestandteil. Für das Pankreaskarzinom wird sie als initiale Bildgebung sowie zur Ausbreitungsdiagnostik verwendet. Hierbei erlaubt die Computertomographie eine exzellente Beurteilung der Lagebeziehung des Tumors zu den umgebenden Gefäßen, die von hoher Relevanz für die Resektabilität eines Tumors ist (Noda et al., 2018).

3.3.2 Diffusionsgewichtete Magnetresonanztomographie

Die diffusionsgewichtete Magnetresonanztomographie ist eine Sequenz der Magnetresonanztomographie (MRT), welche die zufällige Bewegung von Wassermolekülen misst (Brown'sche Molekularbewegung). Die freie Diffusion von Wassermolekülen wird in biologischen Geweben durch den molekularen Aufbau des Gewebes beeinflusst. So interagieren Makromoleküle, Zellmembranen und weitere Gewebebestandteile mit den Wassermolekülen und beschränken deren Diffusion (Koh & Collins, 2007). So beeinflusst z.B. der Zellgehalt die Restriktion der Brown'schen Molekularbewegung (Gauvain et al., 2001; Sugahara et al., 1999), da Zellmembranen als Barrieren für die Wassermoleküle wirken. Weiterhin besitzen zellreiche Gewebe einen proportional geringeren extrazellulären Raum, wodurch die Diffusion weiter eingeschränkt wird. Bei niedrigem Zellgehalt ist der extrazelluläre Raum deutlich größer, folglich können die Wassermoleküle freier diffundieren und die Brown'sche Molekularbewegung ist erhöht (Koh & Collins, 2007). Darüber hinaus verändern Faktoren wie die Beschaffenheit der extrazellulären Matrix die Diffusion. Über die Amplitude und die Länge der Diffusionsgradienten sowie das Zeitintervall zwischen den Gradienten lässt sich die Sensitivität der Sequenz gegenüber der Diffusionseinschränkung variieren. Diese drei Faktoren sind proportional zum sogenannten B-Wert. Durch Variation des B-Wertes kann das Maß der Diffusion mit unterschiedlicher Sensitivität erfasst werden. Aus den Signalintensitäten aller B-Werte in einem Voxel wird mit Hilfe eines mathematischen Modells ein Diffusionskoeffizient berechnet. Aufgrund der Tatsache, dass dieser durch die Vielzahl an oben genannten Faktoren beeinflusst wird, wird von einem scheinbaren Diffusionskoeffizienten (engl.

apparent diffusion coefficient, ADC) gesprochen (Le Bihan et al., 1988). Durch diese Methode lassen sich Bewegungen von Molekülen in einem Abstand von 8 µm messen, somit werden Informationen über die Vorgänge auf zellulärer Ebene erfasst (Neil, 1997). Es muss allerdings beachtet werden, dass auch die Perfusion und die intrazelluläre Diffusion, insbesondere bei niedrigen B-Werten, einen Teil des Signals ausmachen. Mehrere Studien konnten bereits zeigen, dass ein Zusammenhang zwischen den ADC-Werten und verschiedenen klinischen und histologischen Parametern besteht. So konnte für das Prostatakarzinom eine Korrelation zwischen dem Gleasonscore und dem ADC-Wert gefunden werden (Nagarajan et al., 2012). Weiterhin konnten Brustkrebssubtypen anhand des ADC-Wertes differenziert werden (Martincich et al., 2012). Auch für das Pankreaskarzinom konnte gezeigt werden, dass der ADC-Wert mit dem Zellgehalt des Tumors korreliert. Ebenso konnte ein hoher regionaler Zellgehalt und der damit verbundene niedrige ADC-Wert mit einem verkürzten Gesamtüberleben korreliert werden (Heid et al., 2017).

3.4 Radiomics

Vor dem Hintergrund der individualisierten Medizin und durch unser zunehmendes Verständnis über die tumorale Heterogenität werden Methoden nötig, mit welchen sich der gesamte Tumor routinemäßig und effizient analysieren lässt. Wie oben beschrieben, ist die Bildgebung ein zentraler Bestandteil im Patientenmanagement. Allerdings werden zur routinemäßigen Beurteilung des Tumors zurzeit lediglich Parameter wie die Größe des Tumors (RECIST) verwendet. So werden potenziell wichtige Informationen über relevante Eigenschaften des Tumors nicht erfasst.

Im Einklang mit den Fortschritten auf dem Gebiet der medizinischen Bildgebung etablierte sich in den letzten Jahren ein neues Forschungsfeld, welches sich mit der Erfassung von quantifizierbaren Bilddaten befasst. Dieses Forschungsfeld wird Radiomics genannt. Die Kernidee von Radiomics besteht darin, aus Bildern verwertbare Daten zu extrahieren, um aus der Analyse dieser Daten neue Erkenntnisse zu gewinnen. Typischerweise wird hierbei eine Vielzahl an Merkmalen (engl. *features*) aus diagnostischen Aufnahmen extrahiert und diese zur Prädiktion von verschiedenen Endpunkten verwendet. Der Begriff Radiomics wurde erstmalig 2012 in einer Publikation verwendet (Lambin et al., 2012). Seitdem ist es zu einer rapiden Weiterentwicklung dieses Forschungsgebiets gekommen. Im Folgenden wird der typische Arbeitsablauf einer Radiomics-Studie skizziert:

Für die Prädiktion jedweder Endpunkte ist es nötig, ein geeignetes Zielvolumen in den Bilddaten zu definieren (Bildsegmentierung). Aus den so definierten Zielvolumina werden dann die quantitativen Merkmale extrahiert. Diese Merkmale werden in semantische und agnostische Merkmale unterteilt (Gillies et al., 2016). Semantische Merkmale sind von einem Radiologen erfassbar und enthalten Deskriptoren wie Größe, Lokalisation, Nekrose und Vaskularität eines Tumors. Die agnostischen Merkmale werden durch mathematische Formeln direkt aus den Voxeln des Zielvolumens errechnet. Sie lassen sich in Statistiken erster, zweiter und höherer Ordnung einteilen. Die Statistik erster Ordnung beschreibt Merkmale, welche die Verteilung der Signalintensitäten ohne ihre räumliche Beziehung erfassen und basieren auf dem Histogramm des Zielvolumens. Zu den häufig erhobenen Merkmalen gehört der Median, das Minimum und Maximum, sowie die Asymmetrie und die Flachheit des Histogramms (Gillies et al., 2016). Merkmale der Statistik zweiter Ordnung, auch Texturmerkmale genannt, beschreiben die Zusammenhänge zwischen Voxeln mit ähnlichen oder unterschiedlichen Kontrastwerten (Haralick et al., 1973). Zusätzlich kann die Anzahl der Merkmale durch Modifikation des Bildes, beispielsweise durch Applikation bestimmter Filter, vergrößert werden (Randen & Husøy, 1999). Diese Merkmale werden den Merkmalen höherer Ordnung zugeteilt (Gillies et al., 2016). Die extrahierten Merkmale werden in einem nächsten Schritt prozessiert, um möglichst unabhängige und reproduzierbare Merkmale zu erhalten (Lambin et al., 2012). So wird aus einem optisch wahrnehmbaren Bild ein quantitativer Datensatz, der mit verschiedenen Endpunkten korreliert werden kann. Diese Art der Analyse lässt sich auf jedes Krankheitsbild anwenden, wird jedoch vornehmlich zur Charakterisierung von Tumorerkrankungen genutzt (Gillies et al., 2016).

Ein verwandtes Feld, welches die Verbindungen zwischen Tumorgenetik und Bildgebung untersucht, wird Radiogenomics genannt. Hierbei werden die extrahierten Merkmale mit der Genexpression und den Genmustern der einzelnen Tumoren korreliert (Rutman & Kuo, 2009). In mehreren Studien konnte gezeigt werden, dass besagte Korrelationen für diverse Tumorentitäten existieren; beispielsweise konnten bildmorphologische Phänotypen des Glioblastoms bestimmten Genexpressionsmustern zugewiesen (Kickingeder et al., 2016) und aus bildmorphologischen Merkmalen die Überexpression von EGF-Rezeptoren abgeleitet werden (Diehn et al., 2008). Diese Erkenntnisse untermauern die Hypothese, dass aufgrund der engen Korrelation zwischen Genetik (Genom) und Morphologie

(Phenom), radiologische Bilddaten einen Rückschluss auf die zugrunde liegende molekulare Zusammensetzung von z.B. Tumoren ermöglichen.

3.5 Maschinelles Lernen

Als *maschinelles Lernen* wird der Prozess beschrieben, in welchem ein programmierter Algorithmus aus einem Erfahrungsschatz, meistens also einem Datensatz, lernt, eine vorgegebene Aufgabe mit einer bestimmten Genauigkeit zu lösen. Hierbei ist zu erwarten, dass die Genauigkeit der Lösung mit wachsendem Erfahrungsschatz, also zunehmender Datenmenge, zunimmt (Mitchell, 1997). Im Gegensatz zu klassischen Computerprogrammen, die rigide nach implementierten Befehlen arbeiten, erfolgt die Aufgabenlösung durch Algorithmen des maschinellen Lernens, ohne diese *a priori* definierten Befehle (Samuel, 2000). Diese Art des Lernens, die ein Teilbereich der künstlichen Intelligenz ist, ist geeignet Aufgaben zu lösen, deren Komplexität die menschliche Analysefähigkeit übersteigt und deren Gesetzmäßigkeit nicht in feste Befehle gefasst werden kann. Grundsätzlich wird zwischen überwachtem (engl. supervised), nicht überwachtem (engl. unsupervised) und verstärkendem (engl. reinforcement) Lernen unterschieden. Die in dieser Arbeit verwendeten Algorithmen folgen dem Prinzip des überwachten Lernens, auf welches im Folgenden näher eingegangen wird. Überwachtes Lernen beschreibt den Prozess, in welchem ein Algorithmus den mathematischen Zusammenhang zwischen einer Inputvariable (x , z.B. Bilddaten) und einer bekannten Outputvariable (y , z.B. Überleben) modelliert. Dieser Zusammenhang kann als $f(x)=y$ beschrieben werden. Der Algorithmus etabliert an einem Trainingsdatensatz, bestehend aus verschiedenen Input-Output-Paaren die mathematische Funktion, welche diesen Zusammenhang hinreichend beschreibt. Diese Funktion kann daraufhin zur Prädiktion der z.B. noch nicht eingetretenen Outputvariablen „Überleben des Patienten“ von der zugehörigen neuen Inputvariablen (CT Bildgebung vor Therapiebeginn) angewandt werden. Eine nähere Beschreibung der in dieser Arbeit verwendeten Algorithmen erfolgt im Methodenteil. In der medizinischen Forschung werden Algorithmen des maschinellen Lernens mannigfaltig eingesetzt (Coudray et al., 2018; Gibbons et al., 2017; Ong et al., 2012). Insbesondere in den bildgebenden Fachrichtungen konnten eine Vielzahl an Studien den Nutzen solcher Algorithmen demonstrieren (Polan et al., 2016; Setio et al., 2016). Im Gebiet Radiomics können Algorithmen für verschiedene Aufgaben, wie beispielsweise die Präprozessierung von Merkmalen oder für die Segmentierung der Zielvolumina verwendet werden. Häufig kommen sie für die Prädiktion des

definierten Endpunkts zum Einsatz, wobei die extrahierten Bildmerkmale als Input fungieren.

3.6 Zielsetzung der Arbeit

Das Pankreaskarzinom ist eine hochgradig heterogene Tumorentität. Diese Heterogenität manifestiert sich auf genetischer, molekularer und struktureller Ebene. Ein erster Schritt die tumorale Heterogenität zu kategorisieren, war die Entdeckung molekularer Subtypen, die sich in ihrem Überleben und Therapieansprechen unterscheiden. Momentan erfolgt die Klassifikation der Subtypen durch Analyse der genetischen Tumorprofile, oder durch immunhistochemische Färbungen feingeweblicher Tumorschnitte. Der kosten- und zeitintensive Charakter, sowie die Invasivität (Notwendigkeit der Probengewinnung mittels Biopsie) und das Risiko einer unvollständigen Repräsentation (sogenanntes *undersampling*) dieser Analysen verdeutlichen, dass praktikablere Methoden zur Erfassung des Subtyps nötig sind. Ziel dieser Arbeit ist die Entwicklung von Algorithmen des maschinellen Lernens, welche eine Vorhersage der molekularen Subtypen des Pankreaskarzinoms durch Modellierung quantitativer Merkmale aus verschiedenen Bildgebungsmodalitäten ermöglichen. In der vorliegenden Arbeit wurden hierfür präoperative CT-Aufnahmen und diffusionsgewichtete MRT-Aufnahmen verwendet.

4 Methoden

4.1 Studiendesign

Die retrospektive Erhebung, Verarbeitung und Analyse der Patientendaten wurde von der Ethikkommission der medizinischen Fakultät der Technischen Universität München genehmigt (Protokollnummer 180/17). Die jeweilige Kohortenbildung ist der Abbildung 1 zu entnehmen.

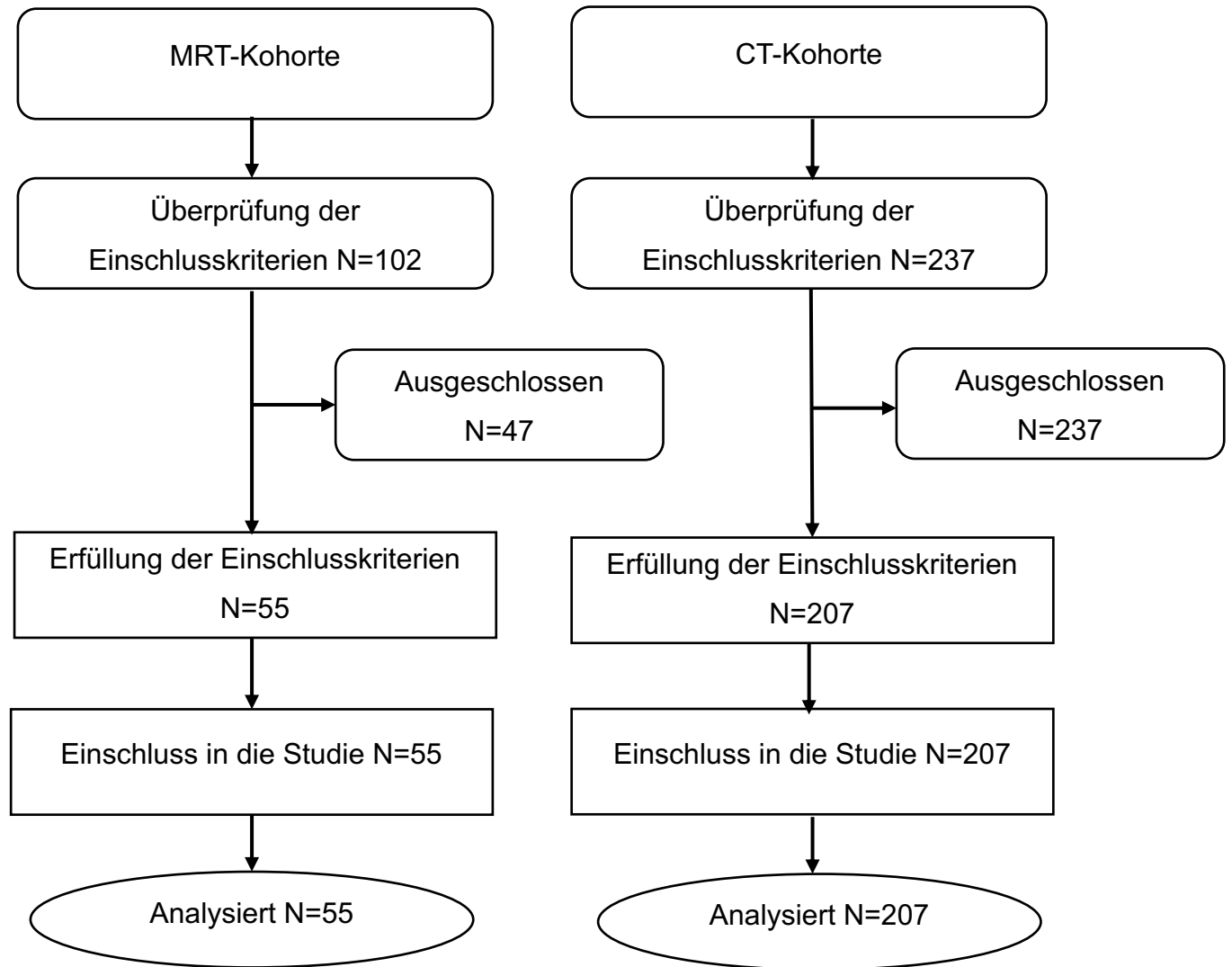


Abbildung 1 Flussdiagramm über die Rekrutierung der Patienten der MRT- und CT-Kohorte

Eine Recherche des PACS (engl. Picture Archiving and Communication System) des Klinikums rechts der Isar ergab 102 Patienten mit Verdacht auf ein duktales Adenokarzinom des Pankreas, die im Zeitraum von Januar 2010 bis Dezember 2016 eine diffusiongewichtete MRT-Untersuchung erhielten und 237 Patienten mit Verdacht auf ein duktales Adenokarzinom des Pankreas, die im Zeitraum von Oktober 2006 bis März 2019 eine CT-Untersuchungen hatten.

Aus der MRT-Kohorte wurden 47 Patienten aus folgenden Gründen ausgeschlossen:

- Kein PDAC im histopathologischen Abschlussbericht (8 Patienten)
- Unzureichende oder unvollständige Bildgebung (14 Patienten)
- Tod innerhalb von zwei Monaten nach initialer Bildgebung (4 Patienten)
- Neoadjuvante Chemotherapie oder vorausgehende Therapie anderweitiger Tumorerkrankungen (8 Patienten)
- Histopathologisch nicht klassifizierbarer Tumorsubtyp (13 Patienten)

Aus der CT-Kohorte wurden 30 Patienten aus folgenden Gründen ausgeschlossen:

- Vorherige Chemotherapie (9 Patienten)
- Unzureichende oder unvollständige Bildgebung (6 Patienten)
- Anderweitige Malignomerkkrankung bei Diagnosestellung (3 Patienten)
- Postoperativer Nachbeobachtungszeitraum unter 2 Wochen (12 Patienten)

Die Analyse erfolgte insgesamt an 262 Patienten, deren Tumoren in kurativer Intention reseziert wurden. Die Tumoren der Patienten wurden abhängig von ihrem immunhistochemischen Subtyp in die zwei Gruppen QM und Non-QM eingeteilt. In der MRT-Kohorte waren 27 Tumoren vom QM-Subtyp (KRT81+) und 28 vom Non-QM-Subtyp (KRT81-). In der CT-Kohorte waren 45 Tumoren vom QM-Subtyp, 136 Tumoren vom Non-QM-Subtyp und für 26 Tumoren war der Subtyp nicht klassifizierbar (KRT81+/HNF1a+). Die Erhebung der klinischen Parameter und *follow-up*-Daten erfolgte über das Krankenhausinformationssystem des Klinikums rechts der Isar und das bayerische Krebsregister.

4.2 Klinische und histologische Parameter

Die erhobenen klinischen und histologischen Parameter, sowie deren Verteilung für die jeweilige Kohorte finden sich in Tabelle 1 und 2. Zusätzlich wurde das mediane Gesamtüberleben sowie, für die MRT-Kohorte, das mediane progressionsfreie Überleben erhoben. Letzteres ist definiert als der Zeitraum zwischen der Diagnosestellung und dem Auftreten eines Rezidivs, neuer Lymphknoten- oder Fernmetastasen. In der MRT-Kohorte verstarben alle Patienten im Erhebungsintervall, sodass keine zensierten Überlebenszeiten für die Kohorte vorlagen.

Tabelle 1 Verteilung der klinischen, histologischen Parameter und der Mittelwerte der Überlebenszeiten der MRT-Kohorte, aufgeteilt nach QM und Non-QM. Quelle: (Kaissis et al., 2019)

Variable	QM (N=27)	Non-QM (N=28)
Alter in Jahren	Mittelwert: 67 ± 11.7	Mittelwert: 65 ± 10.5
Geschlecht	Männlich: 15 Weiblich: 12	Männlich: 15 Weiblich: 13
pT	T1: 3 T2: 2 T3: 22	T1: 3 T2: 3 T3: 22
pN	N0: 6 N1: 21	N0: 8 N1: 20
M	M0: 22 M1: 5	M0: 23 M1: 5
G	G2: 16 G3: 11	G2: 15 G3: 13
R	R0: 20 R1: 7	R0: 21 R1: 7
Adjuvante Chemotherapie	Gemcitabin-basiert: 25 Nicht erhalten: 2	Gemcitabin-basiert: 26 Nicht erhalten: 2
Palliative Chemotherapie	Gemcitabin-basiert: 14 FOLFIRINOX: 13	Gemcitabin-basiert: 16 FOLFIRINOX: 12
ECOG-Status	0: 11 1: 16	0: 13 1: 15
Medianes Gesamtüberleben in Monaten	7.00	22.00
Medianes progressionfreies Überleben in Monaten	6.30	18.90
Zensiert	0	0
Tumorvolumen (ml)	16.4 ± 15.6	15.0 ± 14.0
Lymphknoten-Ratio	0.12 ± 0.07	0.10 ± 0.07

Tabelle 2 Verteilung der klinischen, histologischen Parameter und der Mittelwerte der Überlebenszeiten der CT-Kohorte, aufgeteilt nach QM, Non-QM und unklassifizierbare Tumoren. Quelle: (Kaissis et al., 2020)

Variable	QM (N=45)	Non-QM (N=136)	KRT81+/HNF1a+ (N=26)
Geschlecht	Männlich: 25 Weiblich: 20	Männlich: 75 Weiblich: 61	Männlich: 14 Weiblich: 12
Alter in Jahren	Mittelwert: 68 Spannweite: 42-87	Mittelwert: 67 Spannweite: 45-90	Mittelwert: 72 Spannweite: 53-90
pT	T1: 1 T2: 8 T3: 31 T4: 5	T1: 3 T2: 19 T3: 103 T4: 11	T1: 1 T2: 4 T3: 18 T4: 3
pN	N0: 13 N1: 32	N0: 31 N1: 105	N0: 5 N1: 21
M	M0: 39 M1: 6	M0: 125 M1: 11	M0: 25 M1: 1
G	G1: 1 G2: 21 G3: 23	G1: 6 G2: 60 G3: 70	G1: 4 G2: 13 G3: 9
R	R0: 20 R1: 25	R0: 68 R1: 68	R1: 15 R0: 11
CA-19-9	Normal: 5 Erhöht: 20 N.A.: 20	Normal: 22 Erhöht: 43 N.A.: 71	Normal: 2 Erhöht: 5 N.A.: 19
CEA	Normal: 12 Erhöht: 8 N.A.: 25	Normal: 38 Erhöht: 12 N.A.: 86	Normal: 3 Erhöht: 1 N.A.: 22
Adjuvante-Chemotherapie	Gemcitabin: 16 FOLFIRINOX: 3 Nicht erhalten: 26	Gemcitabin: 68 FOLFIRINOX: 2 Nicht erhalten: :66	Gemcitabin: 12 FOLFIRINOX: 2 Nicht erhalten: 12
ECOG-Status	0: 36 1: 9	0: 106 1: 30	0: 19 1: 7
Medianes Gesamtüberleben in Monaten	9.5	16.5	14.6
Zensiert	14	39	6
Tumorvolumen (ml)	15 ± 10	36 ± 16	26 ± 17
Lymphknoten-Ratio	0.13 ± 0.06	0.11 ± 0.03	0.14 ± 0.05
Tumor Lokalisation	Caput/Corpus: 44 Schwanz: 1	Caput/Corpus: 133 Schwanz: 3	Caput/Corpus: 25 Schwanz: 1

4.3 Bestimmung des immunhistochemischen Subtyps

Die Bestimmung des immunhistochemischen Subtyps erfolgte für sämtliche Tumoren in beiden Kohorten am Institut für Pathologie der Technischen Universität München. Die immunhistochemische Färbung wurde nach dem Protokoll wie in (Muckenhuber et al., 2018; Noll et al., 2016) durchgeführt. Zusammenfassend wurden aus jedem Tumorpräparat mindestens zwei, falls möglich drei repräsentative Gewebezyylinder mit einem Durchmesser von 1mm ausgestanzt (Noll et al., 2016; Stenzinger et al., 2013). Polyklonale Anti-HNF1a-Antikörper und monoklonale Anti-Keratin-81-Antikörper wurden für die immunhistochemische Färbung verwendet. Wie in (Muckenhuber et al., 2018) beschrieben wurde für die Definition von KRT81+ Tumoren ein Cut-off von 30 % exprimierender Tumorzellen festgelegt. Für HNF1a+ wurde die Färbung bei ausreichender Vergrößerung visuell in *schwach*, *mäßig* und *stark* ausgeprägt, eingeteilt. Tumoren mit mäßiger oder starker Färbung wurden als HNF1a+ beschrieben (Muckenhuber et al., 2018). Für die weitere Analyse wurden QM-Subtyp und Non-QM-Subtyp folgendermaßen definiert: QM-Subtyp=KRT81+/HNF1a- und Non-QM-Subtyp=KRT81-/HNF1a+ oder KRT81-/HNF1a-. KRT81+/HNF1a+ Tumoren wurden als immunhistochemisch nicht klassifizierbar eingestuft. In Abbildung 2 sind repräsentative Fotografien der jeweiligen Subtypen dargestellt.

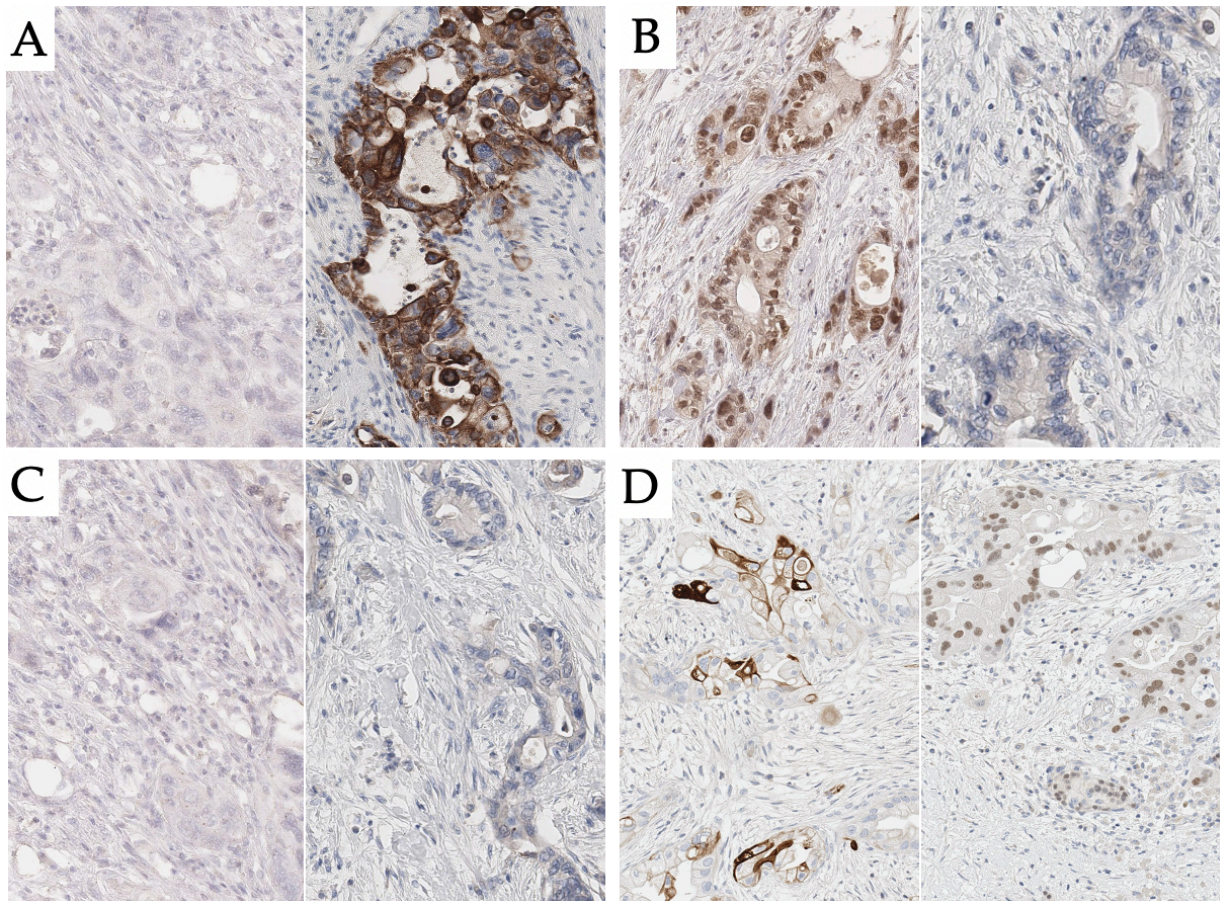


Abbildung 2 Fotografien repräsentativer immunhistochemisch gefärbter Tumorproben KRT81+/HNF1a- (QM) [A], KRT81-/HNF1a+ (Non-QM) [B], KRT81-/HNF1a-(Non-QM) [C] und KRT81+/HNF1a+ (nicht klassifizierbar) [D]. Quelle: (Kaissis et al., 2020)

4.4 Bildgebung

4.4.1 Publikation I

Alle MRT-Aufnahmen wurden an einem 1,5-Tesla-MRT-System (Siemens Magnetom Avanto, Release VB17) durchgeführt. Folgende Sequenzen wurden für jeden Patienten erfasst:

- T2 gewichtete Spin-Echo-Sequenz mit einer Schichtdicke von 5mm
- T1 gewichtete Gradienten-Echo-Sequenz mit einer Schichtdicke von 5mm
- T1 gewichtete kontrastverstärkte Sequenz mit arterieller, portalvenöser, venöser und spätvenöser Phase
- diffusionsgewichtete Sequenz mit B-Werten von 0, 50, 300 und 600
- T2 gewichtete Magnetresonanz-Cholangiopankreatikographie

Die diffusionsgewichtete Sequenz wurde echoplanar ausgelesen und eine ADC-map wurde aus den verschiedenen B-Werten errechnet. Die Rekonstruktion der ADC-map erfolgte auf einer Matrix von 192x192 Voxeln mit einer Voxelgröße von 1,5x1,5x5,5mm. Die Datenerhebung erfolgte in Expiration. Das Bildprotokoll sowie

die technischen Soft- und Hardwarespezifikationen des MRT-Gerätes blieben während der Datenerfassung unverändert.

4.4.2 Publikation II

Die CT-Aufnahmen der Pankreaskarzinompatienten wurden an folgenden Geräten durchgeführt:

- Siemens Somatom Definition AS (Siemens Healthineers, Erlangen, Deutschland)
- Philips iCT (Philips Healthcare, Best, Niederlande)
- Philips IQON Spectral CT (Philips Healthcare, Best, Niederlande)

Die verwendeten CT-Parameter finden sich in Tabelle 3.

Tabelle 3 Akquisitionparameter der verwendeten CT-Geräte

Gerät	Siemens Somatom	Philips iCT	Philips IQON Spectral CT
Röhrenspannung [kV]	120	120	120
Zeilen	64	256	64
Schichtdicke	3mm	3mm	3mm
Patientenanzahl	87	79	41

Bei jedem Patienten wurde in Inspiration eine kontrastmittelgestützte CT-Aufnahme des Abdomens und Beckens in nativer, arterieller und portalvenöser Phase akquiriert. Die arterielle Phase wurde anhand eines Bolus-Trackings bestimmt und die venöse Phase wurde 70 Sekunden nach Kontrastmittelapplikation akquiriert.

4.5 Segmentierung der Pankreaskarzinome

4.5.1 Publikation I

Die Datensätze der Patienten wurden pseudonymisiert und in die Segmentierungssoftware ITK-SNAP (Version 3.8.0) (Yushkevich et al., 2006) importiert. Die Tumoren wurden unter Berücksichtigung aller verfügbaren bildgebenden Sequenzen identifiziert und im b600-Bild manuell segmentiert. Die so generierte Segmentierungsmaske wurde auf die ADC-map transferiert. Die Bilddaten wurden nach zwei Wochen erneut segmentiert. Die Segmentierungen wurden im

Konsens mit einem erfahrenen Radiologen (Facharzt) qualitätskontrolliert und die optimale Segmentierung für jeden Tumor ausgewählt.

4.5.2 Publikation II

Die Segmentierung der CT-Daten erfolgte nach identischem Vorgehen. Die Tumore wurden in der portalvenösen Phase unter Berücksichtigung der nativen und arteriellen Aufnahmen segmentiert. Eine erneute Segmentierung erfolgte für jeweils 20 zufällig ausgewählte Tumoren vom QM-Subtyp, vom Non-QM-Subtyp und von den unklassifizierbaren Tumoren.

4.6 Extraktion der Merkmale

4.6.1 Publikation I

Für die Extraktion der Merkmale wurde die Programmiersprache Python (Version 3.7.3) mit dem Paket PyRadiomics (Version 2.1.1) (Van Griethuysen et al., 2017) verwendet. Zunächst wurde die Plausibilität der Intensitätswerte der Segmentierungen überprüft und die Spannweite der Intensitätswerte der einzelnen Pixel auf 32 Bins aufgeteilt (Vallières et al., 2015). Die Bilddaten wurden daraufhin mit B-Spline Interpolation auf eine isotrope Voxelgröße von 3x3x3mm rekonstruiert. Aufgrund der quantitativen Eigenschaft des ADC-Wertes wurde auf eine Normalisierung der Intensitätswerte verzichtet. Die Merkmale wurden gemäß den Richtlinien der Biomarker Standardisierungs-Initiative (ISBI) extrahiert (Zwanenburg et al., 2017).

Folgende Merkmalkategorien wurden extrahiert:

First order statistics
Shape features 3D
Shape features 2D
Gray Level Cooccurrence Matrix (GLCM)
Gray Level Dependence Matrix (GLDM)
Gray Level Run Length Matrix (GLRLM)
Gray Level Size Zone Matrix (GLSZM)
Neighbouring Gray Tone Difference Matrix (NGTDM)

Die Anzahl der extrahierten Merkmale wurde weiterhin durch die Verwendung verschiedener Filter erweitert. Folgende Filter wurden verwendet:

Laplacian of Gaussian(LoG)-Filter (Marr-Hildreth-Operator) mit einem Sigmawert von 3.0
Wavelet-Transformation (Coiflet-wavelet)
Quadrat der absoluten Werte des Originalbildes
Quadratwurzel der absoluten Werte des Originalbildes
Logarithmus der absoluten Werte des Originalbildes
Exponentialfunktion der absoluten Werte des Originalbildes
Berechnung der Gradientengröße im Originalbild
Berechnung von Local Binary Patterns (Banerjee et al., 2013; Ojala et al., 1994, 1996)

Insgesamt wurden 1606 Merkmale extrahiert. Die Definition und die Berechnung der einzelnen Merkmale ist der Dokumentation von PyRadiomics zu entnehmen.

4.6.2 Publikation II

Die Extraktion der Merkmale der CT-Datensätze erfolgte nach ähnlicher Vorgehensweise. Hiervon abweichend wurden die Dichtewerte auf 25 Bins aufgeteilt. Zusätzlich wurden die Bilder mit Sigmawerten von 1.0 und 2.0 LoG-filtriert. Insgesamt wurden 1474 Merkmale extrahiert.

4.7 Präprozessierung der extrahierten Merkmale

4.7.1 Publikation I

Merkmale mit einer Varianz von null und Merkmale mit fehlenden Werten wurden aus der Analyse entfernt. Die Werte der einzelnen Merkmale wurden zwischen null und eins normalisiert. Die Merkmale der ersten und wiederholten Segmentierung wurden mit dem Intraklassen-Korrelationskoeffizient (ICC) verglichen (McGraw & Wong, 1996; Shrout & Fleiss, 1979). Merkmale mit einem ICC unter 0.8 wurden entfernt. Insgesamt verblieben nach Präprozessierung die folgenden 40 Merkmale:

original_firstorder_Entropy
original_GLCM_Joint_Energy
wavelet-HLL_gldm_SmallDependenceEmphasis
wavelet-LHL_firstorder_Kurtosis

lbp-3D-k_glcm_Contrast
log-sigma-1-0-mm-3D_glcm_Imc2
wavelet-HHL_glszm_SmallAreaEmphasis
lbp-3D-k_firstorder_Variance
original_glszm_LargeAreaLowGrayLevelEmphasis
wavelet-HHL_glszm_ZonePercentage
log-sigma-1-0-mm-3D_glcm_Imc1
wavelet-HHL_glcm_Imc2
lbp-3D-k_firstorder_90Percentile
log-sigma-3-0-mm-3D_gldm_DependenceVariance
wavelet-LHH_firstorder_Kurtosis
wavelet-LHH_glcm_Imc2
lbp-3D-m2_firstorder_Kurtosis
wavelet-LHH_gldm_HighGrayLevelEmphasis
wavelet-HHL_glcm_DifferenceVariance
lbp-3D-k_glcm_Correlation
log-sigma-3-0-mm-3D_glszm_SmallAreaLowGrayLevelEmphasis
wavelet-HHL_glcm_SumSquares
log-sigma-3-0-mm-3D_firstorder_Kurtosis
wavelet-HLL_glszm_ZonePercentage
log-sigma-2-0-mm-3D_glszm_SizeZoneNonUniformity
log-sigma-3-0-mm-3D_gldm_LargeDependenceEmphasis
original_glszm_SizeZoneNonUniformityNormalized
original_gldm_DependenceVariance
original_glszm_LargeAreaEmphasis
wavelet-HHL_glszm_SmallAreaLowGrayLevelEmphasis
wavelet-LHL_glszm_SmallAreaEmphasis
log-sigma-2-0-mm-3D_glszm_SmallAreaHighGrayLevelEmphasis
log-sigma-1-0-mm-3D_glcm_ClusterProminence
original_shape_Maximum2DDiameterRow
wavelet-LLL_firstorder_Entropy
wavelet-LHH_gldm_ShortRunHighGrayLevelEmphasis
gradient_firstorder_Kurtosis
log-sigma-2-0-mm-3D_firstorder_Median
log-sigma-2-0-mm-3D_glcm_Imc1
original_firstorder_Kurtosis

Um eine mögliche Korrelation zwischen den einzelnen Merkmalen zu erkennen wurde der Rangkorrelationskoeffizient nach Spearman für die Werte der Merkmale berechnet. Die Merkmale zeigten alle eine schwache Korrelation (Cohen, 1988).

Die Korrelation wird in Abbildung 3 beispielhaft für die 13 Merkmale höchster Prädiktion gezeigt.

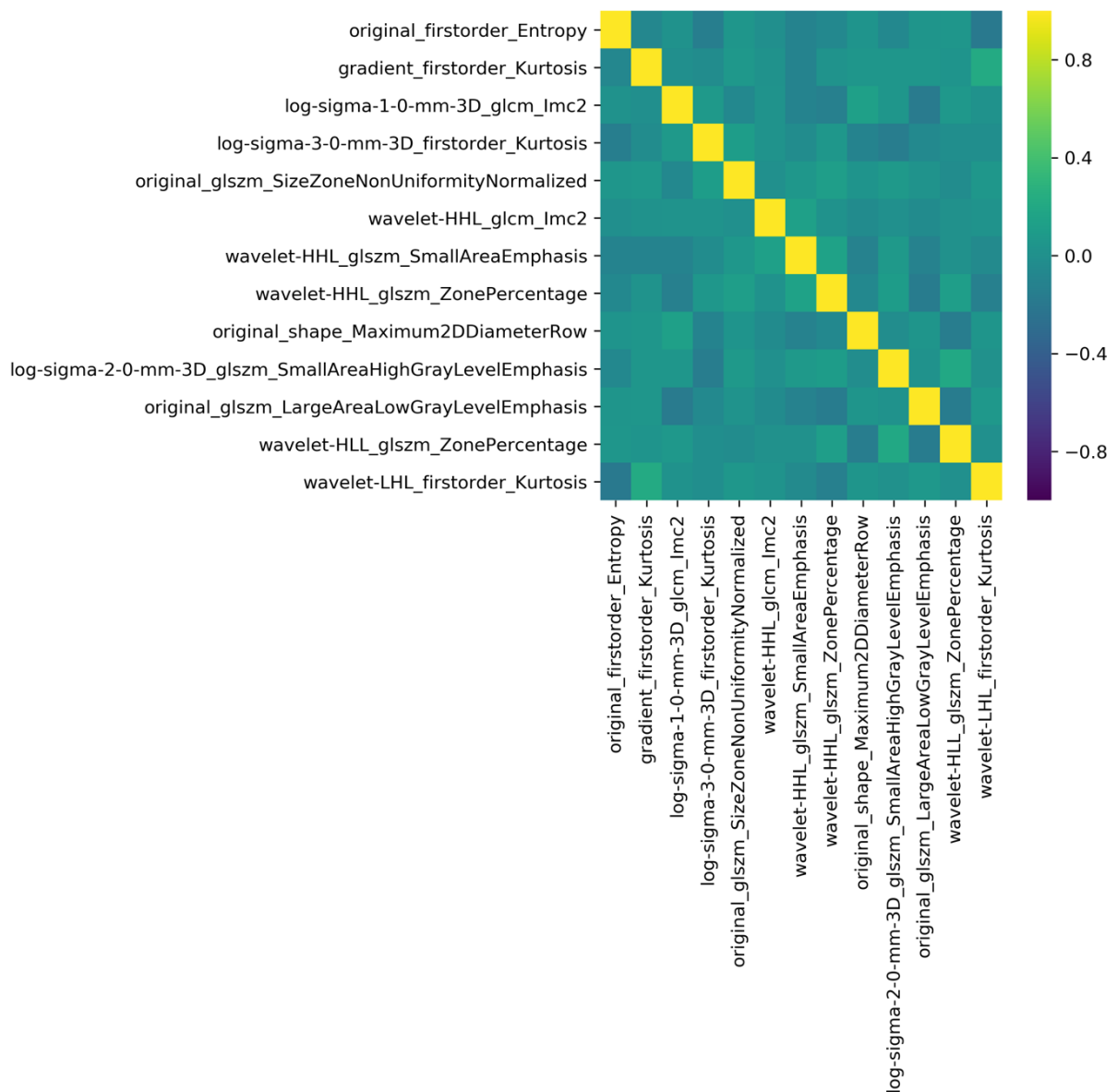


Abbildung 3 Heatmap: Visualisierung der Unabhängigkeit der 13 Merkmale mit höchstem prädiktiven Wert. Quelle: (Kaissis et al., 2019)

4.7.2 Publikation II

Die Präprozessierung der CT-Daten erfolgte nach identischem Vorgehen, jedoch wurde der ICC-Wert auf 0.9 festgelegt. 161 Merkmale wurden selektiert. Die

selektierten Merkmale wurden auf das Tumolvolumen normalisiert (Welch et al., 2019).

4.8 *Machine-learning-Modellierung*

4.8.1 Publikation I

Für die Modellierung der Merkmale und die Durchführung der Überlebensanalysen wurde die Programmierungsumgebung Python (Version 3.7.3) mit folgenden Paketen verwendet: Pandas 0.23.4 (W. McKinney, 2010), Numpy 1.16.2 (Van der Walt & Aivazis, 2011), Scikit Learn 0.20.3 (Pedregosa et al., 2011), XGBoost 0.90 (T. Chen & Guestrin, 2016), Scikit- Survival 0.9 (Pölsterl et al., 2016, 2017) , Matplotlib 3.0.3. (Hunter, 2007), Seaborn 0.9.0, Lifelines 0.21.0, Statsmodels 0.9.0. (Seabold & Perktold, 2010).

Für die Prädiktion der Subtypen wurde der XGBoost-Algorithmus, eine Variante des *Gradient-Boosted-Decision tree* Algorithmus (Breiman, 1998; J. H. Friedman, 2001), genutzt. Die Zielvariable war hierbei der molekulare Subtyp, binär aufgeteilt in QM und Non-QM. Das Modell wurde an den 40 Merkmalen aller Tumoren, beschrieben in Abschnitt 4.7.1, trainiert. Um einer Überanpassung des Modells entgegenzuwirken, wurde die Lernrate auf 0.01 reduziert (Shrinkage-Methode) (J. H. Friedman, 2002) und eine L2-Regularisierung auf die Blattgewichte angewandt. Zur Beurteilung des Modells wurde eine 10-fache stratifizierte Kreuzvalidierung durchgeführt. Die Sensitivität, Spezifität und die Fläche unter der Grenzwertoptimierungskurve (engl. ROC-AUC) wurde für die Testdaten erhoben. Die Wichtigkeit der einzelnen Merkmale wurde anhand der Reduktion im Wert der XGBoost *loss*-Funktion ermittelt. Für die Prädiktion des Gesamtüberlebens und des progressionsfreien Überlebens wurde ein Algorithmus mit stochastischen Gradient-Boosting Regressionsbäumen verwendet. Der Datensatz wurde manuell in einen Trainingsdatensatz von 70% (n=38) und einen Testdatensatz von 30% (n=17) eingeteilt, hierbei wurde auf eine Erhaltung der chronologischen Reihenfolge für beide Datensätze geachtet (*test-set interleaving*). Die Lernrate wurde auf 0.01 gesetzt und die individuellen Bäume wurden an jeweils 90% der Daten trainiert. Als Verlustfunktion wurde das *Cox proportional hazard*-Modell verwendet (Ridgeway, 1999). Die mittlere quadratische Abweichung nach Friedman wurde als Splitstrategie angewandt. Für die Prädiktion des Gesamtüberlebens und des progressionsfreien Überlebens wurde der Konkordanzindex (Penciana & D'Agostino, 2004) zwischen der Prädiktion des Modells für die Testdaten und dem tatsächlichen

Überleben bestimmt. Die 95%-Konfidenzintervalle wurden mittels Bootstrap-Verfahren (Efron, 1977) berechnet. Auf der Basis der Prädiktion der Testdaten (n=17) durch den XGBoost-Algorithmus, wurden die Patienten in zwei Gruppen eingeteilt, einen Hochrisiko-Typ (QM/KRT81+) und einen Niedrigrisiko-Typ (Non-QM/KRT81-). Für die 10 verschiedenen Splits wurden die Kaplan-Meier-Kurven des Gesamtüberlebens und des progressionsfreien Überlebens bestimmt und diese mittels Log-Rank-Test auf einen signifikanten Unterschied getestet. Nach dem gleichen Prinzip wurde das Gesamtüberleben, abhängig von der erhaltenen Chemotherapie, untersucht. Für den Hochrisiko-Typ wurden abhängig vom Split 7 bis 11 Patienten untersucht, für den Niedrigrisiko-Typ wurden 6 bis 10 Patienten untersucht. Zur Kontrolle möglicher Confounder wurde eine multivariate Cox-Regression (Cox, 1972) der klinischen und pathologischen Parameter durchgeführt. Die Verteilungen der Kovariablen zwischen den histopathologischen Subtypen wurde mit dem exakten Test nach Fisher verglichen. Das Signifikanzniveau wurde bei allen statistischen Verfahren auf $p < 0.05$ festgelegt.

4.8.2 Publikation II

Für die Modellierung der CT-Daten wurde die Programmierungsumgebung Python (Version 3.7.6) mit folgenden Paketen verwendet: Pandas 0.23.4, Numpy 1.16.2, Matplotlib 3.0.3, Seaborn 0.9.0, Scikit Learn 0.21.3, Lifelines 0.21.0. Zur Prädiktion des Subtyps wurde ein Random-Forest-Algorithmus (Breiman, 2001) mit folgenden Einstellungen implementiert: 10 Entscheidungsbäume, Wichtigkeit der Merkmale nach Gini. Alle anderen Einstellungen wurden nicht modifiziert. Die Zielvariable war der histopathologische Subtyp, binär in QM und Non-QM aufgeteilt. Das Modell wurde an den 161 Merkmalen, beschrieben in Abschnitt 4.7.2, trainiert. Zur Beurteilung des Modells wurde eine 5-fache Kreuzvalidierung mit einem Trainingsdatensatz von 70% (n=127) und einem Testdatensatz von 30% (n=54) durchgeführt. Die Sensitivität, Spezifität und der ROC-AUC-Wert wurden über die 5 Kreuzvalidierungen gemittelt. Die Wichtigkeit, der für die Prädiktion relevanten Merkmale wurde anhand der Reduktion der *Gini impurity* (Menze et al., 2009) ermittelt, separat für alle fünf Splits

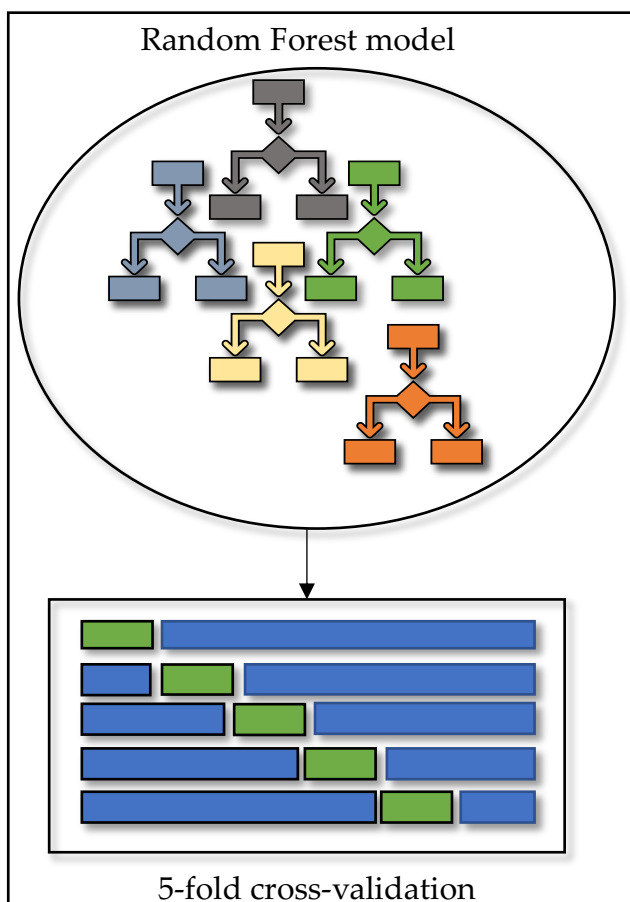


Abbildung 4 Schematische Darstellung eines Entscheidungsbaums und Grundprinzip der Kreuzvalidierung. Quelle: (Kaissis et al., 2020)

bestimmt und gemittelt. Der Algorithmus mit dem besten ROC-AUC-Wert aus den Kreuzvalidierungen wurde im nächsten Schritt auf die nicht klassifizierbaren Tumoren (n=26) angewandt. Der Algorithmus teilte diese in QM-Subtyp (n=12) und Non-QM-Subtyp (n=14) ein. Für die Überlebenszeitanalysen wurden Kaplan-Meier-Kurven erstellt und Log-Rank-Tests durchgeführt. Weiterhin wurde eine multivariate Cox-Regression (Cox, 1972) der klinischen und pathologischen Parameter durchgeführt, um mögliche *Confounder* zu kontrollieren. Die Verteilungen der Kovariablen zwischen den histopathologischen Subtypen wurde mit dem Chi-

Quadrat-Test und dem Zweistichproben-t-Test verglichen. Das Signifikanzniveau wurde bei allen statistischen Verfahren auf $p < 0.05$ festgelegt.

5 Zusammenfassung und Ergebnisse der veröffentlichten Publikationen

5.1 Publikation I

A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response to gemcitabine-based versus FOLFIRINOX chemotherapy.

Kaissis G*, Ziegelmayer S*, Lohöfer F*, et al. PLoS One.

Veröffentlicht: 2. Oktober 2019 DOI:10.1371/journal.pone.0218642

Die retrospektive Beobachtungsstudie befasst sich mit der Fragestellung, ob eine Vorhersage der molekularen Subtypen des Pankreaskarzinoms aus diffusionsgewichteten MRT-Aufnahmen der Tumoren möglich ist. Zusammengefasst, wurden 55 Patienten mit histologisch gesicherten PDAC in die Analyse eingeschlossen. Die molekularen Subtypen wurden mittels immunhistochemischer Färbung von KRT81 und HNF1a bestimmt. Die Tumoren wurden in präoperativen, diffusionsgewichteten MRT-Aufnahmen manuell segmentiert und 1606 Merkmale mit PyRadiomics extrahiert. Basierend auf den extrahierten Merkmalen wurde ein Gradienten-Boosting-Algorithmus (XGBoost) zur Vorhersage der molekularen Subtypen sowie ein stochastischer Gradient-Boosting-Regressionsbaum-Algorithmus zur Vorhersage des progressionsfreien Überlebens und des Gesamtüberlebens modelliert. Die Wichtigkeit der einzelnen Merkmale für die Prädiktion der Subtypen wurde bestimmt. Abschließend wurden das progressionsfreie Überleben, das Gesamtüberleben und das Chemotherapieansprechen in Abhängigkeit des tatsächlichen Subtyps und des radiologischen Phänotyps evaluiert. Der Algorithmus erzielte für die Prädiktion des molekularen Subtyps eine Sensitivität von 0.90 ± 0.07 , eine Spezifität von 0.92 ± 0.11 und eine ROC-AUC von 0.93 ± 0.07 . Die Entropie war das wichtigste Merkmal für die Prädiktion des Subtyps. Der Konkordanz-Index zwischen der Vorhersage des Gradienten-unterstützten Überlebensmodells und dem tatsächlichen Überleben war 0.76 ± 0.05 [95%-Konfidenzintervall 0.66-0.86] für das progressionsfreie Überleben und 0.71 ± 0.06 [95%-Konfidenzintervall 0.60-0.80] für das Gesamtüberleben. Patienten mit histopathologischen QM-Subtyp zeigten ein signifikant ($p < 0.05$) schlechteres Gesamtüberleben als Patienten mit Non-QM-Subtyp (Gesamtüberleben QM: 7.0 Monate vs. Non-QM: 22.6 Monate, Hazard-Ratio 4.03, $p = < 0.001$). Weiterhin hatten Patienten mit Non-QM-Subtyp ein längeres Gesamtüberleben unter Therapie mit FOLFIRINOX als unter Gemcitabin-basierter

Therapie (30.8 Monate vs. 13.4 Monate, Hazard-Ratio = 2.41, $p = 0.027$).
Gegensätzlich zeigten Patienten mit QM-Subtyp ein längeres Gesamtüberleben unter Gemcitabin-basierter Therapie als unter Therapie mit FOLFIRINOX (10.14 Monate vs. 3.8 Monate, Hazard-Ratio = 2.33, $p = 0.037$). Für Patienten mit vorhergesagtem Hochrisiko-Typ (QM/KRT81+) und Niedrigrisiko-Typ (Non-QM/KRT81-) konnte in 8 von 10 Kreuzvalidierungen ein signifikant unterschiedliches Gesamtüberleben und in 7 von 10 Kreuzvalidierungen ein signifikant unterschiedliches progressionsfreies Überleben gezeigt werden. Patienten mit vorhergesagtem Hochrisiko-Typ hatten in 2 von 10 Kreuzvalidierungen ein signifikant längeres Gesamtüberleben unter Gemcitabin-basierter Therapie als unter Therapie mit FOLFIRINOX. Patienten vom Niedrigrisiko-Typ hatten in 6 von 10 Kreuzvalidierungen ein signifikant längeres Gesamtüberleben unter FOLFIRINOX als unter Gemcitabin-basierter Therapie gezeigt.

5.2 Publikation II

Image-Based Molecular Phenotyping of Pancreatic Ductal Adenocarcinoma.

Kaissis G*, Ziegelmayer S*, Lohöfer F*, et al. Journal of Clinical Medicine.

Veröffentlicht: 7.März 2020 DOI:10.3390/jcm9030724

Die retrospektive Studie baut auf die Ergebnisse der ersten Publikation auf und erweitert das Anwendungsfeld auf computertomographische Aufnahmen von Pankreaskarzinompatienten. Die detaillierten Abläufe der Studie sind dem Methodenteil zu entnehmen. Zusammengefasst wurden 207 Patienten mit histopathologisch gesichertem PDAC in die Analyse eingeschlossen. Die Segmentierung der Tumoren erfolgte in präoperativen CT-Aufnahmen. Insgesamt wurden 1474 Radiomics-Merkmale mit Pyradiomics extrahiert. Mittels immunhistochemischer Färbung von KRT81 und HNF1a erfolgte die Einteilung der Tumoren in QM, Non-QM und nicht klassifizierbare Tumoren (KRT81+/HNF1a+). Ein Random-Forest-Algorithmus wurde entwickelt, der, basierend auf den Radiomics-Merkmalen, Tumoren als QM oder Non-QM klassifiziert. Weiterhin wurden die CT-Aufnahmen der nicht klassifizierbaren Tumoren mit dem Algorithmus analysiert. Der Algorithmus erreichte für den Testdatensatz eine Sensitivität von 0.84 ± 0.05 , eine Spezifität von 0.92 ± 0.01 und eine ROC-AUC von 0.93 ± 0.01 . Für die Prädiktion des Subtyps waren Homogenitäts/Heterogenitäts-Merkmale, wie beispielsweise die Entropie von Relevanz. Das mediane Gesamtüberleben für die vorhergesagten Tumoren vom QM-Subtyp und Non-QM-Subtyp war signifikant unterschiedlich. (QM: 16.1 Monate vs. Non-QM: 20.9 Monate, Hazard-Ratio = 1.59, $p = 0.02$).

Der Algorithmus konnte weiterhin die histopathologisch nicht klassifizierbaren Tumoren in einen radiologischen QM-Phänotyp und einen radiologischen Non-QM-Phänotyp einteilen, die ein signifikant unterschiedliches Gesamtüberleben zeigten. (radiologischer QM-Phänotyp: 8.9 Monate vs. radiologischer Non-QM-Phänotyp: 39.8 Monate, Hazard-Ratio 4.33, $p < 0.001$).

Es wurde kein statistisch signifikanter Unterschied in der Verteilung der klinischen *Confounder*-Variablen zwischen den Kohorten gefunden und keiner der Parameter war in der multivariaten Analyse signifikant mit dem Gesamtüberleben assoziiert.

5.3 Schlussfolgerung

Die Ergebnisse dieser Studien zeigen, dass eine prätherapeutische Risikostratifizierung von Pankreaskarzinompatienten im Sinne einer Einteilung der Tumoren in QM und Non-QM anhand präoperativer MRT- und CT-Aufnahmen möglich ist.

6 Diskussion

In der vorliegenden Arbeit wurden Algorithmen des maschinellen Lernens entwickelt, die die nichtinvasive, modalitätsübergreifende Vorhersage des molekularen Subtyps im Pankreaskarzinom aus präoperativen Bildgebungsdaten erlauben. Die Relevanz dieser molekularen Subtypen ist in ihrer klinischen sowie prognostischen Bedeutung begründet. Sie erlauben eine prätherapeutische Patientenstratifikation in definierte Risikogruppen mit unterschiedlichem progressionsfreiem- und Gesamtüberleben sowie unterschiedlichem Ansprechen gegenüber den zurzeit in der klinischen Routine zum Einsatz kommenden Standardchemotherapeutika.

Die Methoden und Ergebnisse dieser Arbeit werden im Folgenden diskutiert.

6.1 Diskussion der Methoden

6.1.1 Studiendesign

Im Rahmen der vorliegenden Arbeit wurden Patienten mit in kurativer Intention resezierten Pankreaskarzinomen untersucht. Die Tumoren wurden histopathologisch gesichert, wodurch sichergestellt werden konnte, dass ausschließlich Adenokarzinome des Pankreas in die Analyse eingeschlossen wurden. Darüber hinaus wurden histopathologische und klinische Parameter als mögliche *Confounder* für die Prädiktion des Subtyps und des Überlebens in die Analyse integriert. Die Studien erfolgten in beiden Fällen an therapienaiven Pankreaskarzinompatienten mit Tumoren in frühen Stadien. Dies ist im Hinblick auf die Erfassung des molekularen Subtyps durch die Algorithmen insofern relevant, als dass eine therapienaive bildmorphologische Repräsentation des Subtyps erlernt werden kann, die frei ist von Selektionsprozessen durch Chemotherapie (Lim et al., 2016).

6.1.2 Immunhistochemische Bestimmung des molekularen Subtyps

Der morpho-molekulare Subtyp der Pankreaskarzinome wurde mittels immunhistochemischer Färbung von KRT81 und HNF1a (Muckenhuber et al., 2018; Noll et al., 2016) bestimmt, die sich an der ursprünglichen transkriptomischen Einteilung der PDACs von Collisson et al. orientiert (Collisson et al., 2011). Diese ursprüngliche Subtypisierung wurde anhand weiterer Ergebnisse aus RNA-Sequenzierung und Genanalyse durch Moffitt et al. 2015 sowie Bailey et al. 2016 modifiziert und erweitert (Bailey et al., 2016; Moffitt et al., 2015). Zwischen den Einteilungen existieren gleichwohl weitgehende Überlappungen, insbesondere im

Hinblick auf die Existenz eines definierten Subtyps, der sich durch signifikant schlechteres Überleben von den anderen abgrenzt (Collisson et al., 2019). Es ist anzunehmen, dass die Positivität für KRT81 das stärkste immunhistochemische Surrogat für diesen Subtyp darstellt (Muckenhuber et al., 2018). Darauf basierend erfolgte in der vorliegenden Arbeit die Einteilung der Subtypen in einen KRT81 positiven Subtyp mit vorwiegend quasi-mesenchymalen Charakteristika (QM) und in einen KRT81 negativen Subtyp, bei dem diese Merkmale nur in geringerem Maße ausgeprägt sind (Non-QM).

Eine Besonderheit stellt in der vorliegenden Arbeit die Untersuchung von doppelt-positiven Tumoren, also Tumoren mit Merkmalen beider erwähneter Subtypen, dar. Diese werden auch als „nicht klassifizierbar“ bezeichnet. Für diese Tumoren kann eine Dualität der Subtypen angenommen werden. Die Koexistenz von mehreren Subtypen in einem Tumor konnte bereits in mehreren Studien demonstriert werden (Chan-Seng-Yue et al., 2020; N. Juiz et al., 2020). Zusätzlich muss angenommen werden, dass eine Transition zwischen quasi-mesenchymalen und eher epithelialen Tumoren regelhaft stattfindet, und insbesondere durch chemotherapeutikabedingte Selektionsprozesse gesteuert wird (Porter et al., 2019). Diese Ergebnisse legen nahe, dass es sich hierbei um einen hochdynamischen Prozess handelt, der durch die Definition eines einzelnen Subtyps nicht hinreichend beschrieben werden kann. Darüber hinaus erfolgt die Beurteilung der immunhistochemischen Färbung in der Routine durch visuelle Quantifizierung der immunhistochemischen Färbungen und ist somit stark untersucherabhängig.

6.1.3 Bildgebung

In der vorliegenden Arbeit wurden präoperative CT-Aufnahmen und diffusionsgewichtete MRT-Aufnahmen analysiert. Die Untersuchungen wurden gemäß klinischer Standardprotokolle durchgeführt, sodass die Algorithmen unter klinischen Bedingungen getestet werden konnten. Dies ermöglicht eine Translation der Ergebnisse sowie die Anwendung der Algorithmen in weiterführenden Studien und legt nahe, dass eine algorithmengestützte Prädiktion des Subtyps nach entsprechender prospektiver Validation Einzug in die klinische Routine halten kann.

Die Untersuchung der Zusammenhänge zwischen Bildmorphologie und Tumorbiologie erfordert Bildgebungsmodalitäten, die morphologische Eigenschaften des

Tumorgewebes sensitiv, spezifisch und reproduzierbar darstellen können. In zahlreichen Tumorentitäten konnte sich hierbei die diffusionsgewichtete Bildgebung als robuster Bestandteil der Diagnostik etablieren (White et al., 2014). So konnte gezeigt werden, dass Bildparameter prädiktiv für die Tumorzellularität, Differenzierung, Nekrose, Aggressivität und Proliferation von Tumoren sein können (H. Wang et al., 2010; Y. Wang et al., 2011). Weiterhin konnte gezeigt werden, dass mit der diffusionsgewichteten MRT auch klinische Endpunkte, wie das Therapieansprechen, vorhergesagt werden können (Perez-Lopez et al., 2017), was die zunehmende klinische Akzeptanz und Anwendung dieser Modalität in der Onkologie begründet (Padhani et al., 2009). Somit besteht ein hohes Interesse, den Nutzen dieser Bildgebungsmodalität für die nicht invasive Tumorcharakterisierung näher zu untersuchen. Für das Pankreaskarzinom konnte bereits gezeigt werden, dass der ADC-Wert prädiktiv für das Gesamtüberleben ist und mit der Zellularität des Tumors korreliert (Heid et al., 2017; Trajkovic-Arsic et al., 2017). Weiterhin konnte in präklinischen Tiermodellen gezeigt werden, dass sich der ADC-Wert zwischen quasi-mesenchymalen und eher epithelial differenzierten Tumorsubregionen unterscheidet (Y. W. Chen et al., 2013).

Bislang existieren jedoch nur wenige Studien, die Methoden der quantitativen Bildanalyse wie Radiomics auf die diffusionsgewichtete Bildgebung des Pankreaskarzinoms anwenden. Dies ist in der hohen technischen Komplexität in der Durchführung von diffusionsgewichteten Pulssequenzen in der abdominalen Bildgebung begründet. Beispielsweise sind Bewegungs- und Atemartefakte im Bereich des Oberbauchs deutlich schwieriger zu kontrollieren als in der neuroradiologischen Bildgebung. Die lange Akquisitionszeit der Sequenzen macht eine atemsynchronisierte Akquisition erforderlich und stellt erhöhte Anforderungen an die Patientenkooperation. Die technische Unterdrückung von Grenzflächenartefakten im Magen und Duodenum sowie von Pulsationsartefakten durch die Aorta abdominalis erfordert robuste und schnelle diffusionsgewichtete Sequenzen, z.B. durch die Anwendung paralleler Bildgebungsverfahren oder reduzierter K-Raumauslesemethoden wie *compressed sensing*, die ein hohes Maß an technischer Expertise für ihre erfolgreiche Implementierung und Anpassung, insbesondere im Hochfeld-MRT-Bereich erfordern (Jethanandani et al., 2018; Kwee et al., 2008).

Im zweiten Teil der Arbeit wurden zur Modellierung CT-Aufnahmen von Pankreaskarzinompatienten genutzt. Die Computertomographie ist durch die kurze Akquisitionszeit, hohe Ortsauflösung, Artefaktarmut sowie die technische Standardisierung für Radiomics-Studien optimal geeignet. Multi-Center-Studien unter Verwertung von CT-Aufnahmen werden durch die flächendeckende Verfügbarkeit und den relativ geringen Preis der Untersuchungen begünstigt. Der prognostische Nutzen von CT-Untersuchungen konnte bereits für verschiedene Tumorentitäten gezeigt werden (Aerts et al., 2014; Huang et al., 2016; Sun et al., 2018). Sie stellt die leitliniengerechte initiale Bildgebung zur Ausbreitungsdiagnostik des Pankreaskarzinoms dar und ist für die Beurteilung der Tumorsektabilität essentiell (Tamm et al., 2012). Auch bei der CT-Untersuchung können jedoch unterschiedliche Akquisitionsparameter sowie Unterschiede zwischen CT-Gerätherstellern die extrahierten quantitativen Bildmerkmale beeinflussen (Berenguer et al., 2018; Ger et al., 2018; MacKin et al., 2018).

Bislang wurden Radiomics-Studien überwiegend an kleinen Datensätzen (bis wenige hundert Patienten) durchgeführt. Wie oben beschrieben konnte festgestellt werden, dass die technische Variabilität zwischen den eingesetzten Scannern sowie die Untersucherabhängigkeit der Tumorsegmentierung zu Ergebnisverzerrungen führten. Folglich wurde bereits in mehreren Publikationen empfohlen worden, die Akquisition der Bilddaten, die Extraktion der Parameter und die Prozessierung zu harmonisieren (Kumar et al., 2012; Leijenaar et al., 2015; Yip & Aerts, 2016). Dieses Vorgehen ist aus mehreren Gründen kritisch zu betrachten. Es ist nicht anzunehmen, dass eine weitreichende Protokollharmonisierung onkologischer Studien –insbesondere im MRT-Bereich– durchsetzbar ist. Ein Ausschluss nicht-konformer Untersuchungen vermindert zusätzlich drastisch die Anzahl der verfügbaren Untersuchungen im Rahmen retrospektiv angelegter Studien. Das Ziel sollte stattdessen sein, durch die Vergrößerung der Kohorten, Diversifizierung der Daten, sowie durch die Anwendung robuster Algorithmen Repräsentationen der Tumorbiologie zu erlernen, die weitgehend agnostisch sind und nicht von nicht-biologischen Parametern abhängen. Es ist ferner anzunehmen, dass solche Algorithmen besser auf neue Kollektive generalisieren. Neuere Empfehlungen, z.B. von Bluemke et al. (Bluemke et al., 2020) ziehen diese Erwägungen zunehmend in Betracht.

6.1.4 Segmentierung

Im Rahmen der vorliegenden Arbeit wurden die Tumolvolumina manuell und unter Berücksichtigung aller vorliegenden Sequenzen und Kontrastmittelphasen segmentiert. Neben der manuellen Segmentierung existieren auch semi-automatische und voll-automatische Segmentierungsmethoden. Semi-automatische Verfahren verlangen weiterhin einen externen Input, wie beispielsweise eine Markierung der zu segmentierenden Struktur. In einer Studie konnte gezeigt werden, dass die semi-automatische Segmentierung von Lungenkarzinomen schneller und exakter sein kann als die manuelle (Rios Velazquez et al., 2012). Ähnliche Ergebnisse wurden bei der semi-automatischen Segmentierung von Glioblastomen beobachtet (X. Zhang et al., 2016). Voll-automatische Modelle sind bereits ebenso zur Organsegmentierung eingesetzt worden (Bilic et al., 2019; Moeskops et al., 2016) und zeigen auch für die Segmentierung von Tumoren vielversprechende Ergebnisse (Harati et al., 2011; Lin et al., 2020). Sowohl das Pankreas als auch Pankreaskarzinome stellen aufgrund des geringen Gewebekontrasts mit teils nicht definierbaren Grenzen, der anatomischen Nähe zu Hohlorganen wie Magen und Dünndarm sowie der pathomorphologischen Alterationen des Pankreas- und umliegenden Fettgewebes durch den Tumor eine große Herausforderung für automatische Segmentierungsalgorithmen dar. Obgleich in anderen Gebieten der Bilderkennung neuronale Netze ausgezeichnete Ergebnisse (Girshick et al., 2016) liefern, konnte für das Pankreas erst unter Verwendung modifizierter neuronaler Netzarchitekturen, beispielsweise unter Verwendung von *Attention-Mechanismen* wie in (Oktay et al., 2018) gezeigt, eine mit dem Menschen vergleichbare Genauigkeit erzielt werden.

Der Vorteil von quantitativen Bildgebungsanalysemethoden wie *Radiomics* gegenüber Biopsien aus Tumorsubregionen liegt in der Erfassung des gesamten Tumolvolumens im Segmentierungsprozess. Wie in Abschnitt 6.1.2 beschrieben ist anzunehmen, dass mehrere Subtypen in einem Tumor koexistieren und dass z. B. durch eine epithelial-mesenchymale Transition Veränderungen der Tumorzusammensetzung stattfinden. Diese Erkenntnis verdeutlicht dass die intratumorale Heterogenität nur suffizient erfasst werden kann, wenn die Gesamtheit des Tumors analysiert wird und, dass jede Art von subregionaler Analyse systematisch Information vernachlässigt (Gerlinger et al., 2012). Dies schließt jedoch nicht aus, dass einzelne Tumorregionen trotzdem einen prädiktiven Wert besitzen können. Für das Pankreaskarzinom erwies sich beispielsweise die Region mit dem niedrigsten ADC-Wert, respektive die Region mit

dem höchsten Zellgehalt, als prädiktiv für das Gesamtüberleben (Heid et al., 2017). Auch für weitere Tumorentitäten, wie beispielsweise das Glioblastom konnte die Wertigkeit einer subregionalen Analyse demonstriert werden (Zhou et al., 2014). Neue Entwicklungen wie die multispektrale Computertomographie oder die hochaufgelöste Magnetresonanztomographie mit reduziertem Sichtfeld (engl. reduced field-of-view) werden zu einer weiteren Erhöhung des Informationsgehalts der akquirierten Datensätze führen, wovon derartige subregionale Analysen profitieren können.

6.1.5 Extraktion und Präprozessierung der Merkmale

Die Extraktion der Merkmale erfolgte mit PyRadiomics (Van Griethuysen et al., 2017). Mit dieser Software wurden aus dem MRT-Datensatz pro Tumor 1606 und aus dem CT-Datensatz 1474 Merkmale extrahiert. Die Berechnung der Merkmale erfolgt nach den Definitionen der Biomarker Standardisierungs-Initiative (ISBI) von Zwanenburg et al., die eine höhere Reproduzierbarkeit der Merkmale gewährleisten soll (Zwanenburg et al., 2017). Durch eine hohe Anzahl an Merkmalen verglichen mit der Anzahl an Patienten kann es zu einer Überanpassung des Modells kommen (Clarke et al., 2008; Dyson, 2004). Daher wurde die Anzahl an extrahierten Merkmalen durch Präprozessierung reduziert. Zur Verbesserung der Reproduzierbarkeit wurden die Tumoren von zwei separaten Auswertern segmentiert, die Übereinstimmung der Merkmale durch Bestimmung des ICC-Werts nach (McGraw & Wong, 1996) verglichen und Merkmale mit niedrigen ICC-Werten entfernt.

Für den CT-Datensatz konnte für einen Teil der Merkmale eine Korrelation mit dem Tumolvolumen festgestellt werden. Vorherige Studien legen nahe, dass das Tumolvolumen ein starker *Confounder* von extrahierten Merkmalen sein kann (Fave et al., 2016), weshalb für den CT-Datensatz die Merkmale nach dem Tumolvolumen normiert wurden. Die hier angewendeten Methoden führen dazu, dass möglichst reproduzierbare und voneinander unabhängige Merkmale für die Analyse verwendet werden.

6.1.6 Algorithmen des maschinellen Lernens

Der Nutzen von Algorithmen der künstlichen Intelligenz wird in sämtlichen Bereichen der Forschung untersucht. Insbesondere im Bereich der Bildanalyse und Bilderkennung konnten bereits vielversprechende Ergebnisse durch die Anwendung von neuronalen Netzwerken veröffentlicht werden, beispielsweise konnte der Verlauf von diabetischen Retinopathien anhand von Funduskopieaufnahmen vorhergesagt

werden und die Genauigkeit in der Diagnostik von Mammakarzinomen erhöht werden (Arcadu et al., 2019; S. M. McKinney et al., 2020). In dieser Arbeit wurde aufgrund der geringen Patientenzahl auf die Implementierung eines neuronalen Netzwerks verzichtet und stattdessen Modelle aus dem Bereich des maschinellen Lernens verwendet. Robuste Netzwerke, die trotz weniger Daten gute Ergebnisse produzieren und gleichzeitig auf externe Daten generalisieren, sind erst in neuerer Zeit verfügbar geworden. Im Bereich der Bildklassifikation ist beispielsweise die Anwendung von vortrainierten Modellen auf die jeweiligen Daten eine bewährte Methode, um auch bei einem kleinen Datensatz gute Ergebnisse zu erzielen (engl. Transfer-learning). Die Annahme, dass klassisches maschinelles Lernen für kleine Datensätze besser geeignet ist, ist somit zunehmend in Frage zu stellen. Dies ist auch durch die Rigidität der extrahierten Merkmale bedingt, welche bei der Verwendung neuronaler Netzwerke nicht zum Einsatz kommen.

In der vorliegenden Arbeit wurden für die Prädiktion des Subtyps und des Überlebens Entscheidungsbaum-basierte Algorithmen benutzt. Ein Vorteil dieser Art von Algorithmen ist, dass sie auch mit nicht skalierten Daten trainiert werden und sowohl lineare als auch nicht lineare Zusammenhänge erfassen können. Für den MRT-Datensatz wurde der XGBoost-Algorithmus (T. Chen & Guestrin, 2016) verwendet, welcher im bildanalytischen Bereich bereits mehrfach zur Prädiktion unterschiedlicher Zielvariablen erfolgreich eingesetzt wurde (Hamerla et al., 2019; Tahmassebi et al., 2019). Das dem Algorithmus zu Grunde liegende Gradient-Boosting führt zu einer verbesserten Prädiktion bei Klassifikationsproblemen (J. Friedman et al., 2000), und die Möglichkeit der Regularisierung verhindert eine Überanpassung des Modells. Für die Prädiktion der Subtypen des CT-Datensatzes wurde ein Random-Forest-Algorithmus (Breiman, 2001) implementiert, für welchen ebenfalls vielversprechende Ergebnisse in Studien gezeigt werden konnten (Jia et al., 2019; Wu et al., 2019). Durch Bagging (*engl. Bootstrap aggregating*) (Breiman, 1996) kommt es bei Random-Forest-Algorithmen selten zu einer Überanpassung des Modells, sodass eine suffiziente Generalisierbarkeit des Modells erreicht werden kann. Zusätzlich besteht die Möglichkeit, den Trainingsprozess zu parallelisieren, was zu einer Verkürzung der Trainingsdauer führt. Die Prädiktion des Überlebens erfolgte mit stochastischen Gradient-Boosting Regressionsbäumen. Ähnlich wie für Klassifikationsprobleme können Regressionbäume nicht lineare Zusammenhänge modellieren. Die Modelle für

die Prädiktion der Subtypen verfügen über eine Funktion, mit welcher die Wichtigkeit der einzelnen Merkmale für die Vorhersage erfasst werden kann. Anhand ihrer Wichtigkeit können einzelne Merkmale auf ihre biologische Grundlage hin näher untersucht und die Komplexität der Prädiktion anhand der Anzahl der relevanten Merkmale erfasst werden.

6.2 Diskussion der Ergebnisse

Fortschritte in der spezifischen Therapie des Pankreaskarzinoms machen eine prätherapeutische Stratifikation in therapierelevante Subtypen sowie ein spezifisches Therapiemonitoring erforderlich. Wie in Abschnitt 3.2.2 beschrieben, existiert eine zunehmende Evidenz, dass für das Pankreaskarzinom Subtypen existieren, die sich in Therapieansprechen, Gesamtüberleben und progressionsfreiem Überleben unterscheiden. Eine prätherapeutische Bestimmung sowie die Erfassung des Subtyps im Therapieverlauf könnten die jeweilige Behandlung der Patienten maßgeblich beeinflussen. Die routinemäßig zur Diagnostik der Tumorausbreitung erhobene Bildgebung ist für eine prätherapeutische Risikostratifizierung optimal geeignet. Für verschiedene Entitäten konnte bereits eine Prädiktion von genetischen Mutationen und weiteren Endpunkten, wie Gesamtüberleben und progressionsfreies Überleben, aus Bilddaten gezeigt werden.

Für das Pankreaskarzinom existieren bislang nur wenige Studien, die den Nutzen von Radiomics untersuchen. Insbesondere die Prädiktion des Subtyps aus Bilddaten konnte zuvor noch nicht gezeigt werden. In der vorliegenden Arbeit wurden die Tumoren basierend auf einer immunhistochemischen Färbung in QM und Non-QM-Subtyp eingeteilt. Die Prädiktion des Subtyps aus den diffusionsgewichteten MRT-Aufnahmen erfolgte mit einer Sensitivität von 0.90 ± 0.07 , Spezifität von 0.92 ± 0.11 und einem ROC-AUC-Wert von 0.93 ± 0.07 . Ebenso war die Subtypprädiktion aus präoperativen CT-Aufnahmen möglich, hierbei konnte eine Sensitivität von 0.84 ± 0.05 , eine Spezifität von 0.92 ± 0.01 und ein ROC-AUC-Wert von 0.93 ± 0.01 erreicht werden. Die Ergebnisse der vorliegenden Arbeit belegen, dass die nichtinvasive Prädiktion therapie- und prognoserelevanter morpho-molekularer Subtypen aus präoperativen Bilddaten möglich ist.

Ferner gelang die Subtypprädiktion modalitätenunabhängig, d.h. sowohl aus CT-Aufnahmen als auch aus MRT-Aufnahmen. Die Definition relevanter und prädiktiver Merkmale wird als essenzieller Bestandteil in der Entwicklung robuster und reproduzierbarer Algorithmen angesehen. Es konnten deutliche Überschneidungen

der prädiktiven Merkmale beobachtet werden. So waren in beiden Modalitäten vor allem heterogenitätsassoziierte Merkmale von hoher Relevanz. Hierzu gehört maßgeblich die Histogrammentropie, deren Wertigkeit bereits in mehreren Studien (Ganeshan et al., 2012; Ng et al., 2013; H. Zhang et al., 2013) untersucht wurde und die als reproduzierbares und stabiles Merkmal gilt (Traverso et al., 2018). Weiterhin gilt die Entropie als wichtige bildmorphologische Repräsentation der tumoralen Heterogenität (Dercle et al., 2017).

Zur Quantifikation der prädiktiven Wertigkeit wurde der trainierte Random-Forest-Algorithmus auf CT-Aufnahmen von 26 nicht klassifizierbaren Tumoren angewandt, wobei sich 12 Tumoren mit einem radiologischen QM-Phänotyp und 14 Tumoren mit einem radiologischen Non-QM-Phänotyp ergaben, die ein signifikant unterschiedliches Gesamtüberleben hatten (radiologischer QM-Phänotyp: 8.9 Monate vs. radiologischer Non-QM-Phänotyp: 39.8 Monate, Hazard-Ratio 4.33, $p < 0.001$). Dieses Ergebnis erlaubt schlusszufolgern, dass der Algorithmus eine bildmorphologische Repräsentation des Subtyps gelernt hat und imstande ist, diese in Tumoren mit mehrdeutiger histopathologischer Zusammensetzung wiederzuerkennen, was das Potential der bildbasierten Risikostratifizierung nochmals betont. Die Erkenntnis, dass mehrere Subtypen in einem Tumor koexistieren (Chan-Seng-Yue et al., 2020) verdeutlicht ebenso, dass eine suffiziente Beurteilung nur durch eine Analyse des gesamten Tumors erreicht werden kann. Die histopathologische Aufarbeitung des gesamten Tumors ist nicht nur zeit- und kostenintensiver als die Bildgebung und Bildverarbeitung, sondern setzt voraus, dass der Tumor bereits reseziert wurde und lässt sich somit nur für resektable Tumoren realisieren. Wie erwähnt können PDACs eine epitheliale-mesenchymale Transition vollziehen (Porter et al., 2019). Die damit verbundene Subtypen-Transformation bedarf einer Analysemethode, die dieses Spektrum dynamisch erfassen kann. Pankreaskarzinompatienten erhalten regelhaft mehrere Verlaufsbildgebungen, die für derartige Analysen optimal verwendet werden können.

Für den MRT-Datensatz erfolgte zusätzlich eine Prädiktion des Gesamtüberlebens und des progressionsfreien Überlebens. Hierbei konnte ein Konkordanz-Index von 0.71 ± 0.06 für das Gesamtüberleben und 0.75 ± 0.06 für das progressionsfreie Überleben zwischen den vorhergesagten und den tatsächlichen Überlebenszeiten erreicht werden. Dieses Ergebnis entspricht Ergebnissen anderer Studien, in denen der prognostische Wert von Radiomics für die Prädiktion des Gesamtüberlebens

anhand präoperativer CT-Aufnahmen untersucht wurde (Khalvati et al., 2019). Weiterhin wurden das Gesamtüberleben, das progressionsfreie Überleben und das Therapieansprechen der Patienten in Abhängigkeit vom histopathologischen Subtyp und dem vorhergesagten radiologischen Phänotyp, erfasst. Patienten mit QM-Subtyp hatten ein signifikant schlechteres Gesamtüberleben als Patienten mit Non-QM-Subtyp auf. Ebenso konnten ein signifikant unterschiedliches Gesamtüberleben und progressionsfreies Überleben zwischen den vorhergesagten Phänotypen festgestellt werden. Weiterhin konnte gezeigt werden, dass Tumoren vom QM-Subtyp ein signifikant längeres Gesamtüberleben unter Gemcitabin-basierter Chemotherapie aufwiesen als unter Therapie mit FOLFIRINOX (10.14 vs. 3.8 Monate, Hazard-Ratio 2.33, $p = 0.037$). Tumoren vom Non-QM-Subtyp zeigten hingegen ein signifikant längeres Gesamtüberleben unter FOLFIRINOX Therapie als unter Gemcitabin-basierter Therapie (30.8 vs. 13.4 Monate, Hazard-Ratio 2.41, $p = 0.027$). Diese Ergebnisse decken sich mit Ergebnissen vorheriger Studien, die dasselbe biologische Verhalten der Subtypen unter Chemotherapie beschreiben (Aung et al., 2018; Collisson et al., 2011). Sie bestärken weiterhin, dass die nichtinvasive Prädiktion des Subtyps anhand bildmorphologischer Merkmale eine präzisionsmedizinische Risiko- und Therapiegruppenstratifikation ermöglicht.

6.3 Limitationen

Bevor quantitative Bildgebungsalgorithmen einer weitreichenden klinischen Translation unterzogen werden können, müssen zahlreiche aktuelle methodologische Einschränkungen gelöst werden. Allen voran ist bei nahezu allen aktuellen Studien die Größe des untersuchten Patientenkollektivs eine Limitation. In der vorliegenden Arbeit konnten aufgrund der hohen Komplexität sowohl der radiologischen Untersuchungstechnik als auch der histopathologischen Analyse nur 55 geeignete MRT-Datensätze identifiziert werden. Prädiktionen der Überlebenszeiten sind somit nur unter Vorbehalt zu interpretieren. Die Prädiktion des Subtyps konnte hingegen durch die Verwendung des CT-Datensatzes validiert werden. Die Größe der Datensätze bedingte ebenso den Einsatz von Methoden des klassischen maschinellen Lernens anstatt der mittlerweile nahezu ausschließlich zum Einsatz kommenden neuronalen Netzwerke. Eine weitere Limitation ist die Verwendung eines einzelnen MRT-Gerätes mit einheitlichem DWI-Protokoll in der Bildakquisition. Für Radiomics-Studien an MRT-Datensätzen wurde eine Abhängigkeit der Merkmale vom jeweiligen Untersuchungsprotokoll postuliert, was eine Reproduzierbarkeit der Merkmale

erschwert (Mayerhoefer et al., 2009). Eine Möglichkeit zur Lösung dieser Problematik wäre die Verwendung mehrerer MRT-Geräte (Bluemke et al., 2020), wie in der vorliegenden Arbeit für den CT-Datensatz erfolgt. Die Vorteile der Datensatzdiversifizierung wurden bereits in einem früheren Abschnitt diskutiert. Weiterhin konnten die Ergebnisse nicht an externen Validierungskohorten untersucht werden. Die einzelnen Modelle wurden zwar mittels Kreuzvalidierung an zurückgehaltenen Testdaten überprüft, jedoch erfolgte die Bildakquisition für diese nach ähnlichen Parametern wie für den Trainingsdatensatz. Im weitesten Sinne kann der CT-Datensatz als externe Validierungskohorte verstanden werden, jedoch sollten die Ergebnisse an optimalerweise prospektiv gesammelten Datensätzen validiert werden. Abschließend wurde diese Arbeit retrospektiv und monoinstitutional durchgeführt. Dies verlangt eine Validation an prospektiven Multi-Center-Studien. Die beschriebenen Limitationen dieser Arbeit sind gängige Probleme von Radiomics-Studien. Dies hat zur Folge, dass zurzeit mehrere Empfehlungen zu einer standardisierten Durchführung derartiger Studien veröffentlicht werden (Shukla-Dave et al., 2019; Zwanenburg et al., 2017).

7 Ausblick

Die Heterogenität des Pankreaskarzinoms ist weiterhin eine große Herausforderung, die einer adäquaten Therapie im Weg steht. Die vorliegende Arbeit ist die erste Studie, in welcher eine Prädiktion molekularer Subtypen des Pankreaskarzinoms anhand von MRT- und CT-Aufnahmen durch Verwendung von Algorithmen des maschinellen Lernens möglich war. Sie demonstriert, dass Bilddaten für eine prätherapeutische Risikostratifizierung von Pankreaskarzinomen verwendet werden können und gibt Anstoß für eine mögliche Implementierung von Algorithmen im klinischen Alltag, was besonders im Hinblick auf die zunehmende Verwendung von neoadjuvanten Therapieregimen in der Behandlung des Pankreaskarzinoms erstrebenswert wäre. Neuronale Netzwerke können für sämtliche Arbeitsschritte, wie beispielsweise Segmentierung und Extraktion der Merkmale, verwendet werden. So könnte für das Pankreaskarzinom eine Rahmenstruktur entwickelt werden, die eine autonome Prädiktion definierter Endpunkte aus CT- und MRT-Daten ermöglicht. Durch Digitalisierung von pathologischen Schnitten und die Ko-registrierung von pathologischen und radiologischen Bildern, könnten weiterführende Untersuchungen in Bezug auf die tumorale Heterogenität erfolgen. Abschließend bleibt zu erwähnen, dass maschinelles Lernen und künstliche Intelligenz unser Verständnis über die Medizin maßgeblich beeinflussen werden. Insbesondere für komplexe Vorgänge, wie die Karzinogenese können Algorithmen wertvolle Informationen generieren.

8 Zusammenfassung

Mit einer Fünf-Jahres-Überlebensrate von neun Prozent ist das Pankreatische Duktale Adenokarzinom (PDAC) die Tumorentität mit der schlechtesten Prognose. Ein Kernproblem in der Behandlung des PDAC ist die ausgeprägte Tumorheterogenität. Entgegen einem generalisierten Vorgehen in der Therapie gibt es zunehmende Evidenz, dass spezifische PDAC-Subtypen existieren, welche ein unterschiedliches Überleben und Therapieansprechen aufweisen. Die Identifikation der Subtypen erfolgte bislang durch eine Analyse der Genprofile oder immunhistochemische Färbungen an Tumoren bzw. Tumorproben. Dies ist jedoch ein invasiver sowie kosten- und zeitintensiver Prozess. Zudem besteht das Risiko der Unterrepräsentation. Daher sind Methoden zur nichtinvasiven Bestimmung des Subtyps nötig. Durch neue Entwicklungen im Gebiet der quantitativen bildbasierten Biomarker wie *Radiomics* und die zunehmende Anwendung von Algorithmen des maschinellen Lernens auf Bilddaten ist es möglich, aus der radiologischen Bildmorphologie eines Tumors dessen biologische Eigenschaften vorherzusagen. Das übergreifende Ziel der Arbeit war daher die Entwicklung von Algorithmen des maschinellen Lernens, welche eine Vorhersage der molekularen Subtypen des Pankreaskarzinoms durch Modellierung quantitativer Merkmale aus verschiedenen Bildgebungsmodalitäten ermöglichen. Hierfür wurden 55 diffusionsgewichtete MRT-Datensätze und 207 CT-Datensätze analysiert. Die Tumoren wurden segmentiert, pathologische und klinische Parameter erhoben und der Subtyp mittels immunhistochemischer Färbung ermittelt. In dieser Arbeit wurden die Tumoren in KRT81 positive (QM) und KRT81 negative (Non-QM) Subtypen eingeteilt. Für die Prädiktion der Subtypen wurden für die Datensätze zwei Algorithmen implementiert, welche an präprozessierten Merkmalen der Tumorsegmentierungen trainiert wurden. Weiterhin wurde für den MRT-Datensatz ein Algorithmus zur Prädiktion des Gesamt- und progressionsfreien Überlebens implementiert. Modalitätsübergreifend war eine Prädiktion des Subtyps mit einer hohen Sensitivität und Spezifität möglich. Zusätzlich wurde der für den CT-Datensatz verwendete Algorithmus auf immunhistochemisch nicht klassifizierbare Tumoren angewandt, wodurch diese in zwei Gruppen mit signifikant unterschiedlichem Überleben eingeteilt werden konnten. Für den MRT-Datensatz konnte ein hoher Konkordanz-Index für die Prädiktion des Gesamt- und progressionsfreien Überlebens erreicht werden. Die Ergebnisse belegen, dass durch die Verwendung von Algorithmen des maschinellen Lernens eine Prädiktion der molekularen Subtypen aus

spezifischen Bildgebungsmodalitäten möglich ist. Weiterhin ist anzunehmen, dass basierend auf den extrahierten Merkmalen ein nichtinvasives radiologisches Korrelat des molekularen Subtyps generiert wurde.

9 Literaturverzeichnis

- Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Cavalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M. M., Leemans, C. R., Dekker, A., Quackenbush, J., Gillies, R. J., & Lambin, P. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, 5(4006).
- Anderson, K., Lutz, C., Van Delft, F. W., Bateman, C. M., Guo, Y., Colman, S. M., Kempinski, H., Moorman, A. V., Titley, I., Swansbury, J., Kearney, L., Enver, T., & Greaves, M. (2011). Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature*, 469(7330), 356–361.
- Arcadu, F., Benmansour, F., Maunz, A., Willis, J., Haskova, Z., & Prunotto, M. (2019). Deep learning algorithm predicts diabetic retinopathy progression in individual patients. *Npj Digital Medicine*, 2(92).
- Aung, K. L., Fischer, S. E., Denroche, R. E., Jang, G. H., Dodd, A., Creighton, S., Southwood, B., Liang, S. Ben, Chadwick, D., Zhang, A., O’Kane, G. M., Albaba, H., Moura, S., Grant, R. C., Miller, J. K., Mbabaali, F., Pasternack, D., Lungu, I. M., Bartlett, J. M. S., Ghai, S., Lemire, M., Holter, S., Connor, A. A., Moffitt, R. A., Yeh, J. J., Timms, L., Krzyzanowski, P. M., Dhani, N., Hedley, D., Notta, F., Wilson, J. M., Moore, M. J., Gallinger, S., & Knox, J. J. (2018). Genomics-driven precision medicine for advanced pancreatic cancer: Early results from the COMPASS trial. *Clinical Cancer Research*, 24(6), 1344–1354.
- Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A. M., Gingras, M. C., Miller, D. K., Christ, A. N., Bruxner, T. J. C., Quinn, M. C., Nourse, C., Murtaugh, L. C., Harliwong, I., Idrisoglu, S., Manning, S., Nourbakhsh, E., Wani, S., Fink, L., Holmes, O., Chin, V., Anderson, M. J., Kazakoff, S., Leonard, C., Newell, F., Waddell, N., Wood, S., Xu, Q., Wilson, P. J., Cloonan, N., Kassahn, K. S., Taylor, D., Quek, K., Robertson, A., Pantano, L., Mincarelli, L., Sanchez, L. N., Evers, L., Wu, J., Pinese, M., Cowley, M. J., Jones, M. D., Colvin, E. K., Nagrial, A. M., Humphrey, E. S., Chantrill, L. A., Mawson, A., Humphris, J., Chou, A., Pajic, M., Scarlett, C. J., Pinho, A. V., Giry-Laterriere, M., Rooman, I., Samra, J. S., Kench, J. G., Lovell, J. A., Merrett, N. D., Toon, C. W., Epari, K., Nguyen, N. Q., Barbour, A., Zeps, N., Moran-Jones, K., Jamieson, N. B., Graham, J. S., Duthie, F., Oien, K., Hair, J., Grützmann, R., Maitra, A., Iacobuzio-Donahue, C. A., Wolfgang, C.

- L., Morgan, R. A., Lawlor, R. T., Corbo, V., Bassi, C., Rusev, B., Capelli, P., Salvia, R., Tortora, G., Mukhopadhyay, D., Petersen, G. M., Munzy, D. M., Fisher, W. E., Karim, S. A., Eshleman, J. R., Hruban, R. H., Pilarsky, C., Morton, J. P., Sansom, O. J., Scarpa, A., Musgrove, E. A., Bailey, U. M. H., Hofmann, O., Sutherland, R. L., Wheeler, D. A., Gill, A. J., Gibbs, R. A., Pearson, J. V., Waddell, N., Biankin, A. V., & Grimmond, S. M. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*, *531*(7592), 47–52.
- Banerjee, J., Moelker, A., Niessen, W. J., & Van Walsum, T. (2013). 3D LBP-based rotationally invariant region description. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7728 LNCS(PART 1)*, 26–37.
- Bartsch, D. K., Gress, T. M., & Langer, P. (2012). Familial pancreatic cancer—current knowledge. *Nature Reviews Gastroenterology and Hepatology*, *9*(8), 445–453.
- Berenguer, R., Del Rosario Pastor-Juan, M., Canales-Vázquez, J., Castro-García, M., Villas, M. V., Legorburo, F. M., & Sabater, S. (2018). Radiomics of CT features may be nonreproducible and redundant: Influence of CT acquisition parameters. *Radiology*, *288*(2), 407–415.
- Bernard, V., Semaan, A., Huang, J., Anthony San Lucas, F., Mulu, F. C., Stephens, B. M., Guerrero, P. A., Huang, Y., Zhao, J., Kamyabi, N., Sen, S., Scheet, P. A., Taniguchi, C. M., Kim, M. P., Tzeng, C. W., Katz, M. H., Singhi, A. D., Maitra, A., & Alvarez, H. A. (2019). Single-cell transcriptomics of pancreatic cancer precursors demonstrates epithelial and microenvironmental heterogeneity as an early event in neoplastic progression. *Clinical Cancer Research*, *25*(7), 2194–2205.
- Bilic, P., Christ, P. F., Vorontsov, E., Chlebus, G., Chen, H., Dou, Q., Fu, C.-W., Han, X., Heng, P.-A., Hesser, J., Kadoury, S., Konopczynski, T., Le, M., Li, C., Li, X., Lipková, J., Lowengrub, J., Meine, H., Moltz, J. H., Pal, C., Piraud, M., Qi, X., Qi, J., Rempfler, M., Roth, K., Schenk, A., Sekuboyina, A., Vorontsov, E., Zhou, P., Hülsemeyer, C., Beetz, M., Ettliger, F., Gruen, F., Kaissis, G., Lohöfer, F., Braren, R., Holch, J., Hofmann, F., Sommer, W., Heinemann, V., Jacobs, C., Mamani, G. E. H., van Ginneken, B., Chartrand, G., Tang, A., Drozdal, M., Ben-Cohen, A., Klang, E., Amitai, M. M., Konen, E., Greenspan, H., Moreau, J., Hostettler, A., Soler, L., Vivanti, R., Szeskin, A., Lev-Cohain, N., Sosna, J., Joskowicz, L., & Menze, B. H. (2019). The Liver Tumor Segmentation Benchmark

- (LiTS). *ArXiv:1901.04056*, 1–43.
- Blomstrand, H., Scheibling, U., Bratthäll, C., Green, H., & Elander, N. O. (2019). Real world evidence on gemcitabine and nab-paclitaxel combination chemotherapy in advanced pancreatic cancer. *BMC Cancer*, *19*(1), 1–9.
- Bluemke, D. A., Moy, L., Bredella, M. A., Ertl-Wagner, B. B., Fowler, K. J., Goh, V. J., Halpern, E. F., Hess, C. P., Schiebler, M. L., & Weiss, C. R. (2020). Assessing radiology research on artificial intelligence: A brief guide for authors, reviewers, and readers-from the Radiology Editorial Board. *Radiology*, *294*(2), 487–489.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140.
- Breiman, L. (1998). Arcing the edge. *Annals of Probability*, *26*, 1683–1702.
- Breiman, L. (2001). RANDOM FORESTS. *Machine Learning*, *45*(1), 5–32.
- Chan-Seng-Yue, M., Kim, J. C., Wilson, G. W., Ng, K., Figueroa, E. F., O’Kane, G. M., Connor, A. A., Denroche, R. E., Grant, R. C., McLeod, J., Wilson, J. M., Jang, G. H., Zhang, A., Dodd, A., Liang, S. Ben, Borgida, A., Chadwick, D., Kalimuthu, S., Lungu, I., Bartlett, J. M. S., Krzyzanowski, P. M., Sandhu, V., Tiriach, H., Froeling, F. E. M., Karasinska, J. M., Topham, J. T., Renouf, D. J., Schaeffer, D. F., Jones, S. J. M., Marra, M. A., Laskin, J., Chetty, R., Stein, L. D., Zogopoulos, G., Haibe-Kains, B., Campbell, P. J., Tuveson, D. A., Knox, J. J., Fischer, S. E., Gallinger, S., & Notta, F. (2020). Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. *Nature Genetics*, *52*(2), 231–240.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Chen, Y. W., Pan, H. Ben, Tseng, H. H., Chu, H. C., Hung, Y. T., Yen, Y. C., & Chou, C. P. (2013). Differentiated epithelial- and mesenchymal-like phenotypes in subcutaneous mouse xenografts using diffusion weighted-magnetic resonance imaging. *International Journal of Molecular Sciences*, *14*(11), 21943–21959.
- Clarke, R., Ransom, H. W., Wang, A., Xuan, J., Liu, M. C., Gehan, E. A., & Wang, Y. (2008). The properties of high-dimensional data spaces: Implications for exploring gene and protein expression data. *Nature Reviews Cancer*, *8*(1), 37–49.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates Publishers.
- Collisson, E. A., Bailey, P., Chang, D. K., & Biankin, A. V. (2019). Molecular subtypes of pancreatic cancer. *Nature Reviews Gastroenterology and Hepatology*, *16*(4),

207–220.

- Collisson, E. A., Sadanandam, A., Olson, P., Gibb, W. J., Truitt, M., Gu, S., Cooc, J., Weinkle, J., Kim, G. E., Jakkula, L., Feiler, H. S., Ko, A. H., Olshen, A. B., Danenberg, K. L., Tempero, M. A., Spellman, P. T., Hanahan, D., & Gray, J. W. (2011). Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nature Medicine*, *17*(4), 500–503.
- Conroy, T., Desseigne, F., Ychou, M., Bouche, O., Guimbaud, R., Becouarn, Y., Adenis, A., Raoul, J. L., Gourgou-Bourgade, S., de la Fouchardiere, C., Bennouna, J., Bachet, J. B., Khemissa-Akouz, F., Pere-Verge, D., Delbaldo, C., Assenat, E., Chauffert, B., Michel, P., Montoto-Grillot, C., Ducreux, M., Groupe Tumeurs Digestives of, U., & Intergroup, P. (2011). FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. *New England Journal of Medicine*, 1817–1825.
- Conroy, T., Hammel, P., Hebbar, M., Ben Abdelghani, M., Wei, A. C., Raoul, J.-L., Chone, L., Francois, E., Artru, P., Biagi, J. J., Lecomte, T., Assenat, E., Faroux, R., Ychou, M., Volet, J., Sauvanet, A., Jouffroy-Zeller, C., RAT, P., Castan, F., & Bachet, J.-B. (2018). Unicancer GI PRODIGE 24/CCTG PA.6 trial: A multicenter international randomized phase III trial of adjuvant mFOLFIRINOX versus gemcitabine (gem) in patients with resected pancreatic ductal adenocarcinomas. *Journal of Clinical Oncology*, *36*.
- Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N., & Tsirigos, A. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature Medicine*, *24*(10), 1559–1567.
- Cox, D. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society*, *34*(2), 187–220.
- Dercle, L., Ammari, S., Bateson, M., Durand, P. B., Haspinger, E., Massard, C., Jaudet, C., Varga, A., Deutsch, E., Soria, J.-C., & Ferte, C. (2017). Limits of radiomic-based entropy as a surrogate of tumor heterogeneity: ROI-area, acquisition protocol and tissue site exert substantial influence. *Scientific Reports*, *7*(1), 7952.
- Diehn, M., Nardini, C., Wang, D. S., McGovern, S., Jayaraman, M., Liang, Y., Aldape, K., Cha, S., & Kuo, M. D. (2008). Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(13), 5213–5218.

- Dirkse, A., Golebiewska, A., Buder, T., Nazarov, P. V., Muller, A., Poovathingal, S., Brons, N. H. C., Leite, S., Sauvageot, N., Sarkisjan, D., Seyfrid, M., Fritah, S., Stieber, D., Michelucci, A., Hertel, F., Herold-Mende, C., Azuaje, F., Skupin, A., Bjerkvig, R., Deutsch, A., Voss-Böhme, A., & Niclou, S. P. (2019). Stem cell-associated heterogeneity in Glioblastoma results from intrinsic tumor plasticity shaped by the microenvironment. *Nature Communications*, *10*(1), 1–16.
- Dyson, F. (2004). A meeting with Enrico Fermi How one intuitive physicist rescued a team from fruitless research. *Nature*, *427*, 8540.
- Efron, B. (1977). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, *7*.
- Faltas, B. M., Prandi, D., Tagawa, S. T., Molina, A. M., Nanus, D. M., Sternberg, C., Rosenberg, J., Mosquera, J. M., Robinson, B., Elemento, O., Sboner, A., Beltran, H., Demichelis, F., & Rubin, M. A. (2016). Clonal evolution of chemotherapy-resistant urothelial carcinoma. *Nature Genetics*, *48*(12), 1490–1499.
- Fave, X., Zhang, L., Yang, J., Mackin, D., Balter, P., Gomez, D., Followill, D., Jones, A. K., Stingo, F., & Court, L. E. (2016). Impact of image preprocessing on the volume dependence and prognostic potential of radiomics features in non-small cell lung cancer. *Translational Cancer Research*, *5*(4), 349–363.
- Feig, C., Gopinathan, A., Neesse, A., Chan, D. S., Cook, N., & Tuveson, D. A. (2012). The pancreas cancer microenvironment. *Clinical Cancer Research*, *18*(16), 4266–4276.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232.
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics and Data Analysis*, *38*(4), 367–378.
- Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting. *The Annals of Statistics*, *28*(2), 337–407.
- Ganeshan, B., Skogen, K., Pressney, I., Coutroubis, D., & Miles, K. (2012). Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: Preliminary evidence of an association with tumour metabolism, stage, and survival. *Clinical Radiology*, *67*(2), 157–164.
- Gatenby, R. A., & Gillies, R. J. (2008). A microenvironmental model of carcinogenesis. *Nature Reviews Cancer*, *8*(1), 56–61.
- Gauvain, K. M., McKinstry, R. C., Mukherjee, P., Perry, A., Neil, J. J., Kaufman, B. A.,

- & Hayashi, R. J. (2001). Evaluating pediatric brain tumor cellularity with diffusion-tensor imaging. *American Journal of Roentgenology*, 177(2), 449–454.
- Ger, R. B., Zhou, S., Chi, P. C. M., Lee, H. J., Layman, R. R., Jones, A. K., Goff, D. L., Fuller, C. D., Howell, R. M., Li, H., Stafford, R. J., Court, L. E., & Mackin, D. S. (2018). Comprehensive Investigation on Controlling for CT Imaging Variabilities in Radiomics Studies. *Scientific Reports*, 8(1), 1–14.
- Gerlinger, M., Rowan, A. J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., Varela, I., Phillimore, B., Begum, S., McDonald, N. Q., Butler, A., Jones, D., Raine, K., Latimer, C., Santos, C. R., Nohadani, M., Eklund, A. C., Spencer-Dene, B., Clark, G., Pickering, L., Stamp, G., Gore, M., Szallasi, Z., Downward, J., Futreal, P. A., & Swanton, C. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New England Journal of Medicine*, 366(10), 883–892.
- Gibbons, C., Richards, S., Valderas, J. M., & Campbell, J. (2017). Supervised machine learning algorithms can classify open-text feedback of doctor performance with human-level accuracy. *Journal of Medical Internet Research*, 19(3), 1–12.
- Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images Are More than Pictures, They Are Data. *Radiology*, 278(2), 563–577.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2016). Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(1), 142–158.
- Golan, T., Hammel, P., Reni, M., Van Cutsem, E., Macarulla, T., Hall, M. J., Park, J. O., Hochhauser, D., Arnold, D., Oh, D. Y., Reinacher-Schick, A., Tortora, G., Algül, H., O'Reilly, E. M., McGuinness, D., Cui, K. Y., Schlienger, K., Locker, G. Y., & Kindler, H. L. (2019). Maintenance olaparib for germline BRCA-mutated metastatic pancreatic cancer. *New England Journal of Medicine*, 381(4), 317–327.
- Goldman, L. W. (2007). Principles of CT and CT technology. *Journal of Nuclear Medicine Technology*, 35(3), 115–128.
- Greaves, M., & Maley, C. C. (2012). Clonal evolution in cancer. *Nature*, 481(7381), 306–313.
- Hall, B. R., Cannon, A., Atri, P., Wichman, C. S., Smith, L. M., Ganti, A. K., Are, C., Sasson, A. R., Kumar, S., & Batra, S. K. (2018). Advanced pancreatic cancer: A meta-analysis of clinical trials over thirty years. *Oncotarget*, 9(27), 19396–19405.
- Hamerla, G., Meyer, H. J., Schob, S., Ginat, D. T., Altman, A., Lim, T., Gühr, G. A.,

- Horvath-Rizea, D., Hoffmann, K. T., & Surov, A. (2019). Comparison of machine learning classifiers for differentiation of grade 1 from higher gradings in meningioma: A multicenter radiomics study. *Magnetic Resonance Imaging*, *63*, 244–249.
- Haralick, R. M., Dinstein, I., & Shanmugam, K. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, *SMC-3*(6), 610–621.
- Harati, V., Khayati, R., & Farzan, A. (2011). Fully automated tumor segmentation based on improved fuzzy connectedness algorithm in brain MR images. *Computers in Biology and Medicine*, *41*(7), 483–492.
- Heid, I., Steiger, K., Trajkovic-Arsic, M., Settles, M., Eßwein, M. R., Erkan, M., Kleeff, J., Jäger, C., Friess, H., Haller, B., Steingötter, A., Schmid, R. M., Schwaiger, M., Rummeny, E. J., Esposito, I., Siveke, J. T., & Braren, R. F. (2017). Co-clinical assessment of tumor cellularity in pancreatic cancer. *Clinical Cancer Research*, *23*(6), 1461–1470.
- Hounsfield, G. N. (1980). Computed medical imaging. *Science*, *210*(4465), 22–28.
- Huang, Y. Q., Liang, C. H., He, L., Tian, J., Liang, C. S., Chen, X., Ma, Z. L., & Liu, Z. Y. (2016). Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *Journal of Clinical Oncology*, *34*(18), 2157–2164.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, *9*(3), 99–104.
- Jethanandani, A., Lin, T. A., Volpe, S., Elhalawani, H., Mohamed, A. S. R., Yang, P., & Fuller, C. D. (2018). Exploring applications of radiomics in magnetic resonance imaging of head and neck cancer: A systematic review. *Frontiers in Oncology*, *8*, 131.
- Jia, T. Y., Xiong, J. F., Li, X. Y., Yu, W., Xu, Z. Y., Cai, X. W., Ma, J. C., Ren, Y. C., Larsson, R., Zhang, J., Zhao, J., & Fu, X. L. (2019). Identifying EGFR mutations in lung adenocarcinoma by noninvasive imaging using radiomics features and random forest modeling. *European Radiology*, *29*(9), 4742–4750.
- Juiz, N. A., Iovanna, J., & Duseti, N. (2019). Pancreatic cancer heterogeneity can be explained beyond the genome. *Frontiers in Oncology*, *9*, 1–8.
- Juiz, N., Elkaoutari, A., Bigonnet, M., Gayet, O., Roques, J., Nicolle, R., Iovanna, J., & Duseti, N. (2020). Basal-like and Classical cells coexistence in pancreatic cancer

- revealed by single cell analysis. *BioRxiv*, June, 2020.01.07.897454.
- Kaissis, G., Ziegelmayer, S., Lohöfer, F. K., Harder, F. N., Jungmann, F., Sasse, D., Muckenhuber, A., Yen, H.-Y., Steiger, K., Siveke, J., Friess, H., Schmid, R., Weichert, W., Makowski, M. R., & Braren, R. F. (2020). Image-Based Molecular Phenotyping of Pancreatic Ductal Adenocarcinoma. *Journal of Clinical Medicine*, 9(3), 724.
- Kaissis, G., Ziegelmayer, S., Lohöfer, F., Steiger, K., Algül, H., Muckenhuber, A., Yen, H. Y., Rummeny, E., Friess, H., Schmid, R., Weichert, W., Siveke, J. T., & Braren, R. (2019). A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response to gemcitabine-based versus FOLFIRINOX chemotherapy. *PLoS ONE*, 14(10).
- Kandel, P., Wallace, M. B., Stauffer, J., Bolan, C., Raimondo, M., Woodward, T. A., Gomez, V., Ritter, A. W., Asbun, H., & Mody, K. (2018). Survival of Patients with Oligometastatic Pancreatic Ductal Adenocarcinoma Treated with Combined Modality Treatment Including Surgical Resection: A Pilot Study. *Journal of Pancreatic Cancer*, 4(1), 88–94.
- Kelsen, D. P., Portenoy, R., Thaler, H., Tao, Y., & Brennan, M. (1997). Pain as a predictor of outcome in patients with operable pancreatic carcinoma. *Surgery*, 122(1), 53–59.
- Kenner, B. J., Chari, S. T., Maitra, A., Srivastava, S., Cleeter, D. F., Go, V. L. W., Rothschild, L. J., & Goldberg, A. E. (2016). Early Detection of Pancreatic Cancer - A Defined Future Using Lessons from Other Cancers: A White Paper. *Pancreas*, 45(8), 1073–1079.
- Khalvati, F., Zhang, Y., Baig, S., Lobo-Mueller, E. M., Karanicolas, P., Gallinger, S., & Haider, M. A. (2019). Prognostic Value of CT Radiomic Features in Resectable Pancreatic Ductal Adenocarcinoma. *Scientific Reports*, 9(1), 1–9.
- Kickingeder, P., Bonekamp, D., Nowosielski, M., Kratz, A., Sill, M., Burth, S., Wick, A., Eidel, O., Schlemmer, H. P., Radbruch, A., Debus, J., Herold-Mende, C., Unterberg, A., Jones, D., & Pfister, S. (2016). Radiogenomics of glioblastoma: Machine learning-based classification of molecular characteristics by using multiparametric and multiregional MR imaging features. *Radiology*, 281(3), 907–918.
- Koh, D. M., & Collins, D. J. (2007). Diffusion-weighted MRI in the body: Applications and challenges in oncology. *American Journal of Roentgenology*, 188(6), 1622–

1635.

- Kovac, M., Navas, C., Horswell, S., Salm, M., Bardella, C., Rowan, A., Stares, M., Castro-Giner, F., Fisher, R., De Bruin, E. C., Kovacova, M., Gorman, M., Makino, S., Williams, J., Jaeger, E., Jones, A., Howarth, K., Larkin, J., Pickering, L., Gore, M., Nicol, D. L., Hazell, S., Stamp, G., O'Brien, T., Challacombe, B., Matthews, N., Phillimore, B., Begum, S., Rabinowitz, A., Varela, I., Chandra, A., Horsfield, C., Polson, A., Tran, M., Bhatt, R., Terracciano, L., Eppenberger-Castori, S., Protheroe, A., Maher, E., El Bahrawy, M., Fleming, S., Ratcliffe, P., Heinemann, K., Swanton, C., & Tomlinson, I. (2015). Recurrent chromosomal gains and heterogeneous driver mutations characterise papillary renal cancer evolution. *Nature Communications*, 6.
- Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S. A., Schabath, M. B., Forster, K., Aerts, H. J. W. L., Dekker, A., Fenstermacher, D., Goldgof, D. B., Hall, L. O., Lambin, P., Balagurunathan, Y., Gatenby, R. A., & Gillies, R. J. (2012). Radiomics: The process and the challenges. *Magnetic Resonance Imaging*, 30(9), 1234–1248.
- Kunzmann, V., Martens, U. M., Alguel, H., Siveke, J. T., Goekkurt, E., Pelzer, U., Siegler, G. M., Hennes, E., Waldschmidt, D., Jakobs, R., Ferenczy, P., Keller, R., Boeck, S. H., Kullmann, F., Kapp, M., & Germer, C. T. (2018). Secondary resectability in locally advanced pancreatic cancer (LAPC) after nab-paclitaxel/gemcitabine- versus FOLFIRINOX-based induction chemotherapy: Interim results of a randomized phase II AIO trial (NEOLAP). *Journal of Clinical Oncology*, 36, 348.
- Kwee, T. C., Takahara, T., Koh, D. M., Nievelstein, R. A. J., & Luijten, P. R. (2008). Comparison and reproducibility of ADC measurements in breathhold, respiratory triggered, and free-breathing diffusion-weighted MR imaging of the liver. *Journal of Magnetic Resonance Imaging*, 28(5), 1141–1148.
- Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., Van Stiphout, R. G. P. M., Granton, P., Zegers, C. M. L., Gillies, R., Boellard, R., Dekker, A., & Aerts, H. J. W. L. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer*, 48(4), 441–446.
- Le Bihan, D., Breton, E., Lallemand, D., Aubin, M. L., Vignaud, J., & Laval-Jeantet, M. (1988). Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology*, 168(2), 497–505.

- Leijenaar, R. T. H., Nalbantov, G., Carvalho, S., Van Elmpt, W. J. C., Troost, E. G. C., Boellaard, R., Aerts, H. J. W. L., Gillies, R. J., & Lambin, P. (2015). The effect of SUV discretization in quantitative FDG-PET Radiomics: The need for standardized methodology in tumor texture analysis. *Scientific Reports*, 5, 1–10.
- Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, A. (2013). S3-Leitlinie zum exokrinen Pankreaskarzinom. *AWMF-Registernummer: 032/0100L*, 1–136.
- Lim, S. K., Lee, M. H., Park, I. H., You, J. Y., Nam, B. H., Kim, B. N., Ro, J., Lee, K. S., Jung, S. Y., Kwon, Y. M., & Lee, E. S. (2016). Impact of molecular subtype conversion of breast cancers after neoadjuvant chemotherapy on clinical outcome. *Cancer Research and Treatment*, 48(1), 133–141.
- Lin, Y. C., Lin, C. H., Lu, H. Y., Chiang, H. J., Wang, H. K., Huang, Y. T., Ng, S. H., Hong, J. H., Yen, T. C., Lai, C. H., & Lin, G. (2020). Deep learning for fully automated tumor segmentation and extraction of magnetic resonance radiomics features in cervical cancer. *European Radiology*, 30(3), 1297–1305.
- Luz, L. P., Al-Haddad, M. A., Sey, M. S. L., & Dewitt, J. M. (2014). Applications of endoscopic ultrasound in pancreatic cancer. *World Journal of Gastroenterology*, 20(24), 7808–7818.
- MacKin, D., Ger, R., Dodge, C., Fave, X., Chi, P. C., Zhang, L., Yang, J., Bache, S., Dodge, C., Jones, A. K., & Court, L. (2018). Effect of tube current on computed tomography radiomic features. *Scientific Reports*, 8(1), 1–10.
- Maley, C. C., Galipeau, P. C., Finley, J. C., Wongsurawat, V. J., Li, X., Sanchez, C. A., Paulson, T. G., Blount, P. L., Risques, R. A., Rabinovitch, P. S., & Reid, B. J. (2006). Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nature Genetics*, 38(4), 468–473.
- Martincich, L., Deantoni, V., Bertotto, I., Redana, S., Kubatzki, F., Sarotto, I., Rossi, V., Liotti, M., Ponzzone, R., Aglietta, M., Regge, D., & Montemurro, F. (2012). Correlations between diffusion-weighted imaging and breast cancer biomarkers. *European Radiology*, 22(7), 1519–1528.
- Mayerhoefer, M. E., Szomolanyi, P., Jirak, D., Materka, A., & Trattnig, S. (2009). Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: An application-oriented study. *Medical Physics*, 36(4), 1236–1243.
- McGraw, K. O., & Wong, S. P. (1996). "Forming inferences about some intraclass

- correlations coefficients”: Correction. *Psychological Methods*, 1(4), 390–390.
- McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafiyan, H., Back, T., Chesus, M., Corrado, G. C., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., Ledsam, J. R., Melnick, D., Mostofi, H., Peng, L., Reicher, J. J., Romera-Paredes, B., Sidebottom, R., Suleyman, M., Tse, D., Young, K. C., De Fauw, J., & Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89–94.
- McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, 51–56.
- Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., & Hamprecht, F. A. (2009). A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics*, 10, 1–16.
- Mitchell, T. (1997). *Machine Learning*. McGraw-Hill Science.
- Moeskops, P., Viergever, M. A., Mendrik, A. M., De Vries, L. S., Benders, M. J. N. L., & Isgum, I. (2016). Automatic Segmentation of MR Brain Images with a Convolutional Neural Network. *IEEE Transactions on Medical Imaging*, 35(5), 1252–1261.
- Moffitt, R. A., Marayati, R., Flate, E. L., Volmar, K. E., Loeza, S. G. H., Hoadley, K. A., Rashid, N. U., Williams, L. A., Eaton, S. C., Chung, A. H., Smyla, J. K., Anderson, J. M., Kim, H. J., Bentrem, D. J., Talamonti, M. S., Iacobuzio-donahue, C. A., Hollingsworth, M. A., & Yeh, J. J. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nature Genetics*, 47(10), 1168–1178.
- Muckenhuber, A., Berger, A. K., Schlitter, A. M., Steiger, K., Konukiewicz, B., Trumpp, A., Eils, R., Werner, J., Friess, H., Esposito, I., Klöppel, G., Ceyhan, G. O., Jesinghaus, M., Denkert, C., Bahra, M., Stenzinger, A., Sprick, M. R., Jäger, D., Springfeld, C., & Weichert, W. (2018). Pancreatic ductal adenocarcinoma subtyping using the biomarkers hepatocyte nuclear factor-1A and cytokeratin-81 correlates with outcome and treatment response. *Clinical Cancer Research*, 24(2), 351–359.
- Nagarajan, R., Margolis, D., Raman, S., Sheng, K., King, C., Reiter, R., & Thomas, M. A. (2012). Correlation of gleason scores with diffusion-weighted imaging findings

- of prostate cancer. *Advances in Urology*, 374805.
- Neil, J. J. (1997). Measurement of water motion (apparent diffusion) in biological systems. *Concepts in Magnetic Resonance*, 9(6), 385–401.
- Ng, F., Kozarski, R., Ganeshan, B., & Goh, V. (2013). Assessment of tumor heterogeneity by CT texture analysis: Can the largest cross-sectional area be used as an alternative to whole tumor analysis? *European Journal of Radiology*, 82(2), 342–348.
- Noda, Y., Goshima, S., Kawada, H., Kawai, N., Miyoshi, T., Matsuo, M., & Bae, K. (2018). Modified National Comprehensive Cancer Network Criteria for Assessing Resectability of Pancreatic Ductal Adenocarcinoma. *American Journal of Roentgenology*, 1–8.
- Noll, E. M., Eisen, C., Stenzinger, A., Espinet, E., Muckenhuber, A., Klein, C., Vogel, V., Klaus, B., Nadler, W., Rösli, C., Lutz, C., Kulke, M., Engelhardt, J., Zickgraf, F. M., Espinosa, O., Schlesner, M., Jiang, X., Kopp-Schneider, A., Neuhaus, P., Bahra, M., Sinn, B. V., Eils, R., Giese, N. A., Hackert, T., Strobel, O., Werner, J., Büchler, M. W., Weichert, W., Trumpp, A., & Sprick, M. R. (2016). CYP3A5 mediates basal and acquired therapy resistance in different subtypes of pancreatic ductal adenocarcinoma. *Nature Medicine*, 22(3), 278–287.
- Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, 194(4260), 23–28.
- Oettle, H., Neuhaus, P., Hochhaus, A., Hartmann, J. T., Gellert, K., Ridwelski, K., Niedergethmann, M., Zülke, C., Fahlke, J., Arning, M. B., Sinn, M., Hinke, A., & Riess, H. (2013). Adjuvant chemotherapy with gemcitabine and long-term outcomes among patients with resected pancreatic cancer: The CONKO-001 randomized trial. *JAMA*, 310(14), 1473–1481.
- Ojala, T., Pietikäinen, M., & Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. *Proceedings - International Conference on Pattern Recognition*, 3, 582–585.
- Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29(1), 51–59.
- Oktay, O., Schlemper, J., Folgoc, L. Le, Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., & Rueckert, D. (2018). Attention U-Net: Learning Where to Look for the Pancreas. *ArXiv:1804.03999*.

- Ong, M. S., Magrabi, F., & Coiera, E. (2012). Automated identification of extreme-risk events in clinical incident reports. *Journal of the American Medical Informatics Association*, *19*, 110–118.
- Özdemir, B. C., Pentcheva-Hoang, T., Carstens, J. L., Zheng, X., Wu, C. C., Simpson, T. R., Laklai, H., Sugimoto, H., Kahlert, C., Novitskiy, S. V., DeJesus-Acosta, A., Sharma, P., Heidari, P., Mahmood, U., Chin, L., Moses, H. L., Weaver, V. M., Maitra, A., Allison, J. P., LeBleu, V. S., & Kalluri, R. (2014). Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. *Cancer Cell*, *25*(6), 719–734.
- Padhani, A. R., Liu, G., Mu-Koh, D., Chenevert, T. L., Thoeny, H. C., Takahara, T., Dzik-Jurasz, A., Ross, B. D., Van Cauteren, M., Collins, D., Hammoud, D. A., Rustin, G. J. S., Taouli, B., & Choyke, P. L. (2009). Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations. *Neoplasia*, *11*(2), 102–125.
- Parkin, D. M., Boyd, L., & Walker, L. C. (2011). The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *British Journal of Cancer*, *105*, 77–81.
- Parsons, D. W., Jones, S., Zhang, X., Lin, J. C. H., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., Hartigan, J., Smith, D. R., Strausberg, R. L., Marie, S. K. N., Shinjo, S. M. O., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E., & Kinzler, K. W. (2008). An integrated genomic analysis of human glioblastoma multiforme. *Science*, *321*(5897), 1807–1812.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.
- Penciana, M. J., & D’Agostino, R. B. (2004). Overall C as a measure of discrimination in survival analysis: Model specific population value and confidence interval estimation. *Statistics in Medicine*, *23*(13), 2109–2123.
- Perez-Lopez, R., Mateo, J., Mossop, H., Blackledge, M. D., Collins, D. J., Rata, M., Morgan, V. A., Macdonald, A., Sandhu, S., Lorente, D., Rescigno, P., Zafeiriou,

- Z., Bianchini, D., Porta, N., Hall, E., Leach, M. O., de Bono, J. S., Koh, D.-M., & Tunariu, N. (2017). Diffusion-weighted Imaging as a Treatment Response Biomarker for Evaluating Bone Metastases in Prostate Cancer: A Pilot Study. *Radiology*, *283*(1), 168–177.
- Polan, D. F., Brady, S. L., & Kaufman, R. A. (2016). Tissue segmentation of computed tomography images using a Random Forest algorithm: A feasibility study. *Physics in Medicine and Biology*, *61*(17), 6553–6569.
- Pölsterl, S., Gupta, P., Wang, L., Conjeti, S., Katouzian, A., & Navab, N. (2017). Heterogeneous ensembles for predicting survival of metastatic, castrate-resistant prostate cancer patients. *F1000Research*, *5*(May), 1–29.
- Pölsterl, S., Navab, N., & Katouzian, A. (2016). An Efficient Training Algorithm for Kernel Survival Support Vector Machines. *ArXiv:1611.07054*, *1*.
- Porter, R. L., Magnus, N. K. C., Thapar, V., Morris, R., Szabolcs, A., Neyaz, A., Kulkarni, A. S., Tai, E., Chougule, A., Hillis, A., Golczer, G., Guo, H., Yamada, T., Kurokawa, T., Yashaswini, C., Ligorio, M., Vo, K. D., Nieman, L., Liss, A. S., Deshpande, V., Lawrence, M. S., Maheswaran, S., Castillo, C. F. Del, Hong, T. S., Ryan, D. P., O'Dwyer, P. J., Drebin, J. A., Ferrone, C. R., Haber, D. A., & Ting, D. T. (2019). Epithelial to mesenchymal plasticity and differential response to therapies in pancreatic ductal adenocarcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(52).
- Provenzano, P. P., Cuevas, C., Chang, A. E., Goel, V. K., Von Hoff, D. D., & Hingorani, S. R. (2012). Enzymatic Targeting of the Stroma Ablates Physical Barriers to Treatment of Pancreatic Ductal Adenocarcinoma. *Cancer Cell*, *21*(3), 418–429.
- Puleo, F., Nicolle, R., Blum, Y., Cros, J., Marisa, L., Demetter, P., Quertinmont, E., Svrcek, M., Elarouci, N., Iovanna, J., Franchimont, D., Verset, L., Galdon, M. G., Devière, J., de Reyniès, A., Laurent-Puig, P., Van Laethem, J. L., Bachet, J. B., & Maréchal, R. (2018). Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. *Gastroenterology*, *155*(6).
- Randen, T., & Husøy, J. H. (1999). Filtering for texture classification: A comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *21*(4), 291–310.
- Raphael, B. J., Hruban, R. H., Aguirre, A. J., Moffitt, R. A., Yeh, J. J., Stewart, C., Robertson, A. G., Cherniack, A. D., Gupta, M., Getz, G., Gabriel, S. B., Meyerson, M., Cibulskis, C., Fei, S. S., Hinoue, T., Shen, H., Laird, P. W., Ling, S., Lu, Y.,

Mills, G. B., Akbani, R., Loher, P., Londin, E. R., Rigoutsos, I., Telonis, A. G., Gibb, E. A., Goldenberg, A., Mezlini, A. M., Hoadley, K. A., Collisson, E., Lander, E., Murray, B. A., Hess, J., Rosenberg, M., Bergelson, L., Zhang, H., Cho, J., Tiao, G., Kim, J., Livitz, D., Leshchiner, I., Reardon, B., Van Allen, E., Kamburov, A., Beroukhim, R., Saksena, G., Schumacher, S. E., Noble, M. S., Heiman, D. I., Gehlenborg, N., Kim, J., Lawrence, M. S., Adsay, V., Petersen, G., Klimstra, D., Bardeesy, N., Leiserson, M. D. M., Bowlby, R., Kasaian, K., Birol, I., Mungall, K. L., Sadeghi, S., Weinstein, J. N., Spellman, P. T., Liu, Y., Amundadottir, L. T., Tepper, J., Singhi, A. D., Dhir, R., Paul, D., Smyrk, T., Zhang, L., Kim, P., Bowen, J., Frick, J., Gastier-Foster, J. M., Gerken, M., Lau, K., Leraas, K. M., Lichtenberg, T. M., Ramirez, N. C., Renkel, J., Sherman, M., Wise, L., Yena, P., Zmuda, E., Shih, J., Ally, A., Balasundaram, M., Carlsen, R., Chu, A., Chuah, E., Clarke, A., Dhalla, N., Holt, R. A., Jones, S. J. M., Lee, D., Ma, Y., Marra, M. A., Mayo, M., Moore, R. A., Mungall, A. J., Schein, J. E., Sipahimalani, P., Tam, A., Thiessen, N., Tse, K., Wong, T., Brooks, D., Auman, J. T., Balu, S., Bodenheimer, T., Hayes, D. N., Hoyle, A. P., Jefferys, S. R., Jones, C. D., Meng, S., Mieczkowski, P. A., Mose, L. E., Perou, C. M., Perou, A. H., Roach, J., Shi, Y., Simons, J. V., Skelly, T., Soloway, M. G., Tan, D., Veluvolu, U., Parker, J. S., Wilkerson, M. D., Korkut, A., Senbabaoglu, Y., Burch, P., McWilliams, R., Chaffee, K., Oberg, A., Zhang, W., Gingras, M. C., Wheeler, D. A., Xi, L., Albert, M., Bartlett, J., Sekhon, H., Stephen, Y., Howard, Z., Judy, M., Breggia, A., Shroff, R. T., Chudamani, S., Liu, J., Lolla, L., Naresh, R., Pihl, T., Sun, Q., Wan, Y., Wu, Y., Jennifer, S., Roggin, K., Becker, K. F., Behera, M., Bennett, J., Boice, L., Burks, E., Carlotti Junior, C. G., Chabot, J., Pretti da Cunha Tirapelli, D., Sebastião dos Santos, J., Dubina, M., Eschbacher, J., Huang, M., Huelsenbeck-Dill, L., Jenkins, R., Karpov, A., Kemp, R., Lyadov, V., Maithel, S., Manikhas, G., Montgomery, E., Noushmehr, H., Osunkoya, A., Owonikoko, T., Paklina, O., Potapova, O., Ramalingam, S., Rathmell, W. K., Rieger-Christ, K., Saller, C., Setdikova, G., Shabunin, A., Sica, G., Su, T., Sullivan, T., Swanson, P., Tarvin, K., Tavobilov, M., Thorne, L. B., Urbanski, S., Voronina, O., Wang, T., Crain, D., Curley, E., Gardner, J., Mallery, D., Morris, S., Paulauskis, J., Penny, R., Shelton, C., Shelton, T., Janssen, K. P., Bathe, O., Bahary, N., Slotta-Huspenina, J., Johns, A., Hibshoosh, H., Hwang, R. F., Sepulveda, A., Radenbaugh, A., Baylin, S. B., Berrios, M., Bootwalla, M. S., Holbrook, A., Lai, P. H., Maglinte, D. T., Mahurkar, S., Triche, T. J., Van Den Berg,

- D. J., Weisenberger, D. J., Chin, L., Kucherlapati, R., Kucherlapati, M., Pantazi, A., Park, P., Saksena, G., Voet, D., Lin, P., Frazer, S., Defreitas, T., Meier, S., Chin, L., Kwon, S. Y., Kim, Y. H., Park, S. J., Han, S. S., Kim, S. H., Kim, H., Furth, E., Tempero, M., Sander, C., Biankin, A., Chang, D., Bailey, P., Gill, A., Kench, J., Grimmond, S., Johns, A., Initiative, C. G., Postier, R., Zuna, R., Sicotte, H., Demchok, J. A., Ferguson, M. L., Hutter, C. M., Mills Shaw, K. R., Sheth, M., Sofia, H. J., Tarnuzzer, R., Wang, Z., Yang, L., Zhang, J., Felau, I., & Zenklusen, J. C. (2017). Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell*, 32(2), 185-203.e13.
- Rawla, P., Sunkara, T., & Gaduputi, V. (2019). Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. *World Journal of Oncology*, 10(1), 10–27.
- Ridgeway, G. (1999). The state of boosting. *Computing Science and Statistics*, 31, 172–181.
- Rios Velazquez, E., Aerts, H. J. W. L., Gu, Y., Goldgof, D. B., De Ruyscher, D., Dekker, A., Korn, R., Gillies, R. J., & Lambin, P. (2012). A semiautomatic CT-based ensemble segmentation of lung tumors: Comparison with oncologists' delineations and with the surgical specimen. *Radiotherapy and Oncology*, 105(2), 167–173.
- Roberts, N. J., Norris, A. L., Petersen, G. M., Bondy, M. L., Brand, R., Gallinger, S., Kurtz, R. C., Olson, S. H., Rustgi, A. K., Schwartz, A. G., Stoffel, E., Syngal, S., Zogopoulos, G., Ali, S. Z., Axilbund, J., Chaffee, K. G., Chen, Y. C., Cote, M. L., Childs, E. J., Douville, C., Goes, F. S., Herman, J. M., Iacobuzio-Donahue, C., Kramer, M., Makohon-Moore, A., McCombie, R. W., Wyatt McMahon, K., Niknafs, N., Parla, J., Pirooznia, M., Potash, J. B., Rhim, A. D., Smith, A. L., Wang, Y., Wolfgang, C. L., Wood, L. D., Zandi, P. P., Goggins, M., Karchin, R., Eshleman, J. R., Papadopoulos, N., Kinzler, K. W., Vogelstein, B., Hruban, R. H., & Klein, A. P. (2016). Whole genome sequencing defines the genetic heterogeneity of familial pancreatic cancer. *Cancer Discovery*, 6(2), 166–175.
- Rutman, A. M., & Kuo, M. D. (2009). Radiogenomics: Creating a link between molecular diagnostics and diagnostic imaging. *European Journal of Radiology*, 70(2), 232–241.
- Samuel, A. L. (2000). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 44(1), 261–269.

- Scheufele, F., Hartmann, D., & Friess, H. (2019). Treatment of pancreatic cancer—neoadjuvant treatment in borderline resectable/locally advanced pancreatic cancer. *Translational Gastroenterology and Hepatology*, 4.
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*, 57.
- Seeram, E. (2009). *Computed Tomography: Physical Principles, Clinical Applications, and Quality Control*. Saunders/Elsevier.
- Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Van Riel, S. J., Wille, M. M. W., Naqibullah, M., Sanchez, C. I., & Van Ginneken, B. (2016). Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. *IEEE Transactions on Medical Imaging*, 35(5), 1160–1169.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Shukla-Dave, A., Obuchowski, N. A., Chenevert, T. L., Jambawalikar, S., Schwartz, L. H., Malyarenko, D., Huang, W., Noworolski, S. M., Young, R. J., Shiroishi, M. S., Kim, H., Coolens, C., Laue, H., Chung, C., Rosen, M., Boss, M., & Jackson, E. F. (2019). Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *Journal of Magnetic Resonance Imaging*, 49(7), e101–e121.
- Sjöblom, T., Jones, S., Wood, L. D., Parsons, D. W., Lin, J., Barber, T. D., Mandelker, D., Leary, R. J., Ptak, J., Silliman, N., Szabo, S., Buckhaults, P., Farrell, C., Meeh, P., Markowitz, S. D., Willis, J., Dawson, D., Willson, J. K. V., Gazdar, A. F., Hartigan, J., Wu, L., Liu, C., Parmigiani, G., Park, B. H., Bachman, K. E., Papadopoulos, N., Vogelstein, B., Kinzler, K. W., & Velculescu, V. E. (2006). The consensus coding sequences of human breast and colorectal cancers. *Science*, 314(5797), 268–274.
- Stenzinger, A., Endris, V., Klauschen, F., Sinn, B., Lorenz, K., Warth, A., Goeppert, B., Ehemann, V., Muckenhuber, A., Kamphues, C., Bahra, M., Neuhaus, P., & Weichert, W. (2013). High SIRT1 expression is a negative prognosticator in pancreatic ductal adenocarcinoma. *BMC Cancer*, 13, 1–12.
- Sugahara, T., Korogi, Y., Kochi, M., Ikushima, I., Shigematu, Y., Hirai, T., Okuda, T., Liang, L., Ge, Y., Komohara, Y., Ushio, Y., & Takahashi, M. (1999). Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity

- in gliomas. *Journal of Magnetic Resonance Imaging*, 9(1), 53–60.
- Sun, R., Limkin, E. J., Vakalopoulou, M., Dercle, L., Champiat, S., Han, S. R., Verlingue, L., Brandao, D., Lancia, A., Ammari, S., Hollebecque, A., Scoazec, J. Y., Marabelle, A., Massard, C., Soria, J. C., Robert, C., Paragios, N., Deutsch, E., & Féré, C. (2018). A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *The Lancet Oncology*, 19(9), 1180–1191.
- Tahmassebi, A., Wengert, G. J., Helbich, T. H., Bago-Horvath, Z., Alaei, S., Bartsch, R., Dubsky, P., Baltzer, P., Clauser, P., Kapetas, P., Morris, E. A., Meyer-Baese, A., & Pinker, K. (2019). Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Investigative Radiology*, 54(2), 110–117.
- Tamm, E. P., Balachandran, A., Bhosale, P. R., Katz, M. H., Fleming, J. B., Lee, J. H., & Varadhachary, G. R. (2012). Imaging of Pancreatic Adenocarcinoma: Update on Staging/Resectability. *Radiologic Clinics of North America*, 50(3), 407–428.
- Tempero, M. A., Malafa, M. P., Al-Hawary, M., Asbun, H., Bain, A., Behrman, S. W., Benson, A. B., Binder, E., Cardin, D. B., Cha, C., Chiorean, E. G., Chung, V., Czito, B., Dillhoff, M., Dotan, E., Ferrone, C. R., Hardacre, J., Hawkins, W. G., Herman, J., Ko, A. H., Komanduri, S., Koong, A., LoConte, N., Lowy, A. M., Moravek, C., Nakakura, E. K., O'Reilly, E. M., Obando, J., Reddy, S., Scaife, C., Thayer, S., Weekes, C. D., Wolff, R. A., Wolpin, B. M., Burns, J., & Darlow, S. (2017). Pancreatic Adenocarcinoma, Version 2.2017, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network*, 15(8), 1028–1061.
- Torphy, R. J., Wang, Z., True-Yasaki, A., Volmar, K. E., Rashid, N., Yeh, B., Johansen, J. S., Hollingsworth, M. A., Yeh, J. J., & Collisson, E. A. (2018). Stromal Content Is Correlated With Tissue Site, Contrast Retention, and Survival in Pancreatic Adenocarcinoma. *JCO Precision Oncology*, 2, 1–12.
- Trajkovic-Arsic, M., Heid, I., Steiger, K., Gupta, A., Fingerle, A., Wörner, C., Teichmann, N., Sengkwawoh-Lueong, S., Wenzel, P., Beer, A. J., Esposito, I., Braren, R., & Siveke, J. T. (2017). Apparent Diffusion Coefficient (ADC) predicts therapy response in pancreatic ductal adenocarcinoma. *Scientific Reports*, 7(1), 1–9.

- Traverso, A., Wee, L., Dekker, A., & Gillies, R. (2018). Repeatability and Reproducibility of Radiomic Features: A Systematic Review. *International Journal of Radiation Oncology Biology Physics*, 102(4), 1143–1158.
- Ueno, H., Kosuge, T., Matsuyama, Y., Yamamoto, J., Nakao, A., Egawa, S., Doi, R., Monden, M., Hatori, T., Tanaka, M., Shimada, M., & Kanemitsu, K. (2009). A randomised phase III trial comparing gemcitabine with surgery-only in patients with resected pancreatic cancer: Japanese Study Group of Adjuvant Therapy for Pancreatic Cancer. *British Journal of Cancer*, 101(6), 908–915.
- Vallièrès, M., Freeman, C. R., Skamene, S. R., & El Naqa, I. (2015). A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Physics in Medicine and Biology*, 60(14), 5471–5496.
- Van der Walt, S., & Aivazis, M. (2011). The NumPy Array: A Structure for Efficient Numerical Computation, Computing in Science & Engineering. *Computing in Science and Engineering*, 13(2), 22–30.
- Van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillion-Robin, J. C., Pieper, S., & Aerts, H. J. W. L. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, 77(21), 104–107.
- Von Hoff, D. D., Ervin, T., Arena, F. P., Chiorean, E. G., Infante, J., Moore, M., Seay, T., Tjulandin, S. A., Ma, W. W., Saleh, M. N., Harris, M., Reni, M., Dowden, S., Laheru, D., Bahary, N., Ramanathan, R. K., Tabernero, J., Hidalgo, M., Goldstein, D., Van Cutsem, E., Wei, X., Iglesias, J., & Renschler, M. F. (2013). Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *New England Journal of Medicine*, 369(18), 1691–1703.
- Waghray, M., Yalamanchili, M., Magliano, M. P. Di, & Simeone, D. M. (2013). Deciphering the role of stroma in pancreatic cancer. *Current Opinion in Gastroenterology*, 29(5), 537–543.
- Wang, H., Cheng, L., Zhang, X., Wang, D., Guo, A., Gao, Y., & Ye, H. (2010). Renal cell carcinoma: Diffusion-weighted MR imaging for subtype differentiation at 3.0 T. *Radiology*, 257(1), 135–143.
- Wang, Y., Chen, Z. E., Yaghmai, V., Nikolaidis, P., McCarthy, R. J., Merrick, L., & Miller, F. H. (2011). Diffusion-weighted MR imaging in pancreatic endocrine tumors correlated with histopathologic characteristics. *Journal of Magnetic*

Resonance Imaging, 33(5), 1071–1079.

- Welch, M. L., McIntosh, C., Haibe-Kains, B., Milosevic, M. F., Wee, L., Dekker, A., Huang, S. H., Purdie, T. G., O'Sullivan, B., Aerts, H. J. W. L., & Jaffray, D. A. (2019). Vulnerabilities of radiomic signature development: The need for safeguards. *Radiotherapy and Oncology*, 130, 2–9.
- White, N. S., McDonald, C. R., Farid, N., Kuperman, J., Karow, D., Schenker-Ahmed, N. M., Bartsch, H., Rakow-Penner, R., Holland, D., Shabaik, A., Bjørnerud, A., Hope, T., Hattangadi-Gluth, J., Liss, M., Parsons, J. K., Chen, C. C., Raman, S., Margolis, D., Reiter, R. E., Marks, L., Kesari, S., Mundt, A. J., Kaine, C. J., Carter, B. S., Bradley, W. G., & Dale, A. M. (2014). Diffusion-weighted imaging in cancer: Physical foundations and applications of restriction spectrum imaging. *Cancer Research*, 74(17), 4638–4652.
- Witkiewicz, A. K., McMillan, E. A., Balaji, U., Baek, G. H., Lin, W. C., Mansour, J., Mollaee, M., Wagner, K. U., Koduru, P., Yopp, A., Choti, M. A., Yeo, C. J., McCue, P., White, M. A., & Knudsen, E. S. (2015). Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nature Communications*, 6, 1–11.
- Wu, J., Liu, A., Cui, J., Chen, A., Song, Q., & Xie, L. (2019). Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on precontrast magnetic resonance images. *BMC Medical Imaging*, 19(1), 1–11.
- Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R. H., Eshleman, J. R., Nowak, M. A., Velculescu, V. E., Kinzler, K. W., Vogelstein, B., & Iacobuzio-Donahue, C. A. (2010). Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature*, 467(7319), 1114–1117.
- Ying, H., Dey, P., Yao, W., Kimmelman, A. C., Draetta, G. F., Maitra, A., & Depinho, R. A. (2016). Genetics and biology of pancreatic ductal adenocarcinoma. *Genes and Development*, 30(4), 355–385.
- Yip, S. S. F., & Aerts, H. J. W. L. (2016). Applications and limitations of radiomics. *Physics in Medicine and Biology*, 61(13), 150–166.
- Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 31(3), 1116–1128.
- Zentrum für Krebsregisterdaten im Robert Koch-Institut. (2016). Bericht zum Krebsgeschehen in Deutschland 2016. *Bericht Zum Krebsgeschehen in*

Deutschland 2016. Zentrum Für Krebsregisterdaten Im Robert Koch-Institut (Hrsg). Berlin, 274.

- Zhang, H., Graham, C. M., Elci, O., Griswold, M. E., Zhang, X., Khan, M. A., Pitman, K., Caudell, J. J., Hamilton, R. D., Ganeshan, B., & Smith, A. D. (2013). Locally advanced squamous cell carcinoma of the head and neck: CT texture and histogram analysis allow independent prediction of overall survival in patients treated with induction chemotherapy. *Radiology*, 269(3), 801–809.
- Zhang, X., Li, X., Li, H., & Feng, Y. (2016). A semi-automatic brain tumor segmentation algorithm. *Proceedings - IEEE International Conference on Multimedia and Expo*.
- Zhou, M., Hall, L., Goldgof, D., Russo, R., Balagurunathan, Y., Gillies, R., & Gatenby, R. (2014). Radiologically Defined Ecological Dynamics and Clinical Outcomes in Glioblastoma Multiforme: Preliminary Results. *Translational Oncology*, 7(1), 5–13.
- Zwanenburg, A., Leger, S., Vallières, M., & Löck, S. (2017). The image biomarker standardisation initiative. *ArXiv:1612.07003v7*.

10 Abbildungsverzeichnis

Abbildung 1: Flussdiagramm über die Rekrutierung der Patienten der MRT- und CT-Kohorte	13
Abbildung 2: Fotografien repräsentativer immunhistochemisch gefärbter Tumorproben der jeweiligen Subtypen	18
Abbildung 3: Heatmap zur Visualisierung der Korrelation von 13 exemplarischen Merkmalen	23
Abbildung 4: Schematische Darstellung eines Entscheidungsbaums und Grundprinzip der Kreuzvalidierung	26

11 Tabellenverzeichnis

Tabelle 1: Verteilung der klinischen, histologischen Parameter und der Mittelwerte der Überlebenszeiten der MRT-Kohorte	15
Tabelle 2: Verteilung der klinischen, histologischen Parameter und der Mittelwerte der Überlebenszeiten der CT-Kohorte	16

12 Danksagung

Die abschließenden Worte meiner Promotionsarbeit sind den Personen gewidmet, die dieses Vorhaben ermöglicht und begleitet haben.

Neben den vielen prägenden Persönlichkeiten, die ich im Rahmen meiner Promotionsarbeit kennen lernen durfte, möchte ich insbesondere meinem Mentor Dr. Georgios Kaissis danken. Seine fachliche Kompetenz, seine Hilfsbereitschaft und sein Sinn für Struktur und Ordnung sind wenige von vielen Charakterzügen, die ihn zu einem ausgezeichneten Wissenschaftler und einem wahren Freund machen.





Weiterhin möchte ich PD Dr. Rickmer Braren danken, der die Rahmenstruktur für dieses Projekt legte und der seiner Position als Betreuer im professionellen, wie auch persönlichen Sinn vollumfänglich gerecht wurde.

Meiner Familie möchte ich danken, ohne die mein Werdegang nicht möglich gewesen wäre und die in jeder Lebenslage an mich geglaubt haben.


Schlussendlich möchte ich meiner Verlobten Nadja danken, die mich auch in schwierigsten Zeiten unterstützt hat und die mich durch ihre liebevolle Art zu einem besseren Menschen macht.

RESEARCH ARTICLE

A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response to gemcitabine-based versus FOLFIRINOX chemotherapy

Georgios Kaissis¹ , Sebastian Ziegelmayer¹ , Fabian Lohöfer¹ , Katja Steiger², Hana Algül³, Alexander Muckenhuber² , Hsi-Yu Yen², Ernst Rummeny¹, Helmut Friess⁴, Roland Schmid³, Wilko Weichert², Jens T. Siveke^{5,6}, Rickmer Braren^{1*}

1 Department of Diagnostic and Interventional Radiology, School of Medicine, Technical University of Munich, Munich, Germany, **2** Department of Pathology, School of Medicine, Technical University of Munich, Munich, Germany, **3** Department of Internal Medicine II, School of Medicine, Technical University of Munich, Munich, Germany, **4** Department of Surgery, School of Medicine, Technical University of Munich, Munich, Germany, **5** Division of Solid Tumor Translational Oncology, West German Cancer Center, University Hospital Essen, Essen, Germany, **6** German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Heidelberg, Germany

 These authors contributed equally to this work.

* rbraren@tum.de



 OPEN ACCESS

Citation: Kaissis G, Ziegelmayer S, Lohöfer F, Steiger K, Algül H, Muckenhuber A, et al. (2019) A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response to gemcitabine-based versus FOLFIRINOX chemotherapy. *PLoS ONE* 14(10): e0218642. <https://doi.org/10.1371/journal.pone.0218642>

Editor: Francisco X. Real, Centro Nacional de Investigaciones Oncológicas, SPAIN

Received: June 3, 2019

Accepted: September 19, 2019

Published: October 2, 2019

Copyright: © 2019 Kaissis et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data can be found at <https://zenodo.org/record/3337876#.XS2dCi2Q2WY>.

Funding: This work was supported by funding of the German Research Foundation (DFG) within the SFB-Initiative 824 (collaborative research center), "Imaging for Selection, Monitoring and Individualization of Cancer Therapies" to RB (SFB824, project C6) and WW (project Z2). The

Abstract

Purpose

Development of a supervised machine-learning model capable of predicting clinically relevant molecular subtypes of pancreatic ductal adenocarcinoma (PDAC) from diffusion-weighted-imaging-derived radiomic features.

Methods

The retrospective observational study assessed 55 surgical PDAC patients. Molecular subtypes were defined by immunohistochemical staining of KRT81. Tumors were manually segmented and 1606 radiomic features were extracted with *PyRadiomics*. A gradient-boosted-tree algorithm was trained on 70% of the patients (N = 28) and tested on 30% (N = 17) to predict KRT81+ vs. KRT81- tumor subtypes. A gradient-boosted survival regression model was fit to the disease-free and overall survival data. Chemotherapy response and survival were assessed stratified by subtype and radiomic signature. Radiomic feature importance was ranked.

Results

The mean±STDEV sensitivity, specificity and ROC-AUC were 0.90±0.07, 0.92±0.11, and 0.93±0.07, respectively. The mean±STDEV concordance indices between the disease-free and overall survival predicted by the model based on the radiomic parameters and actual

fundings had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

patient survival were 0.76 ± 0.05 and 0.71 ± 0.06 , respectively. Patients with a KRT81+ subtype experienced significantly diminished median overall survival compared to KRT81- patients (7.0 vs. 22.6 months, HR 4.03, log-rank-test $P = < 0.001$) and a significantly improved response to gemcitabine-based chemotherapy over FOLFIRINOX (10.14 vs. 3.8 months median overall survival, HR 2.33, $P = 0.037$) compared to KRT81- patients, who responded significantly better to FOLFIRINOX over gemcitabine-based treatment (30.8 vs. 13.4 months median overall survival, HR 2.41, $P = 0.027$). Entropy was ranked as the most important radiomic feature.

Conclusions

The machine-learning based analysis of radiomic features enables the prediction of subtypes of PDAC, which are highly relevant for disease-free and overall patient survival and response to chemotherapy.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) carries the worst prognosis of all tumor entities. Complete resection, often combined with an adjuvant chemotherapy regimen, remains the only curative therapy option in PDAC. In the metastatic setting, gemcitabine/nab-paclitaxel or FOLFIRINOX-based chemotherapy have been the mainstay in the treatment of PDAC [1–3]. However, although both intensified treatment protocols increased response rates up to approximately 30%, a substantial number of patients does not respond or acquires resistance in a considerably short time. Pre-clinical and clinical evidence suggests differential response of specific PDAC subtypes to these treatments. Among these, a particularly aggressive subtype, termed quasi-mesenchymal, basal-like or cytokeratin 81 positive (KRT81+) [4,5] has been investigated and found to be more sensitive to gemcitabine treatment *in vitro* [6] and less sensitive to FOLFIRINOX in a prospective clinical trial [7]. Thus, pre-therapeutic identification of specific subtypes in pancreatic cancer is urgently required to guide individual treatment decision.

So far, molecular profiling has relied on tissue biopsies, which are prone to undersampling, not least because of this entity's morphological heterogeneity, which manifests as a heterogenic mix of tumor cell clusters, stroma and non-tumoral cell infiltrates. In addition, molecular subtyping requires high tissue quality and is both costly and time consuming, thus at current not introduced in routine patient care.

Non-invasive diffusion weighted-magnetic resonance imaging (DW-MRI, DWI), is an imaging technique which is part of the routine diagnostic work-up in many centers. It measures the random motion of water molecules and can thus quantify tissue microstructure and heterogeneity with high sensitivity [8]. Radiomics, i.e. the computer-based analysis of non-perceptual image features, provides a novel tool for the evaluation of DWI beyond traditional descriptive radiology. Recent work has shown its potential in e.g. the differentiation of tumor grading or the prediction of therapy response and survival in various tumor entities including PDAC [9,10].

In the current study we developed a machine learning algorithm capable of predicting clinically relevant histopathological PDAC subtypes from pre-operative DW-MRI derived ADC maps, evaluated tumor subtype-stratified overall survival for different chemotherapy regimens and assessed the clinical utility of this radiomic algorithm in the prediction of patient survival and chemotherapy response.

Materials and methods

Study design

The study was designed as a retrospective observational cohort study matched on histopathological tumor subtype.

Data collection, processing and analysis were approved by the institutional ethics committee (Ethics Commission of the Faculty of Medicine of the Technical University of Munich, protocol number 180/17). The requirement for consent was waived. All procedures were carried out in accordance to pertinent laws and regulations.

The STROBE checklist and inclusion flowchart can be found in [S1 File](#). In brief, we considered 102 consecutive patients with final histopathological diagnosis of PDAC of the head or body for inclusion in the study. Patients without a final diagnosis of PDAC, with *unclassifiable* tumor subtype, who had undergone prior therapy (chemotherapy, resection prior to enrollment), died within the first 6 weeks of follow-up (to limit bias from postoperative complications), did not undergo the full imaging protocol or did not have technically sufficient imaging available (due to e.g. motion artifacts or stent placement), were excluded. A total of 55 patients who underwent surgical resection in curative intention were included in the study using histopathological subtype as the matching criterion. 27 patients with a KRT81+ subtype and 28 patients with a KRT81- subtype [5] were included. The follow-up interval began on the 1st of January, 2010 and ended on the 31st of December 2016. All patients died within the follow-up interval thus observed (uncensored) endpoint data is available for all patients. For 21 patients, follow-up data and histopathological data was sourced from the “PR2” cohort described in [5]. For all other patients, clinical follow-up was handled by the departments of surgery and internal medicine, clinical data was sourced from the hospital’s clinical system and histopathological data was generated during the study. Radiomic data for all patients was generated during data analysis. All analyses were performed on pseudonymized datasets by separate individuals (G.K. and S.Z.) from January to May 2019.

Clinical data

The following clinical data was collected: age at diagnosis, sex, pTNM, R, G, tumor volume (from the final histopathological report), ECOG-status, adjuvant chemotherapy (gemcitabine-based vs. no chemotherapy), palliative chemotherapy (gemcitabine-based vs. FOLFIRINOX) and lymph-node ratio (LNR). Disease-free survival was defined as the time from diagnosis to tumor recurrence or occurrence of metastatic disease and overall survival as the time from diagnosis to disease-related death.

Imaging data

Patients underwent magnetic resonance imaging (MRI) at 1.5T (Siemens Magnetom Avanto, release VB17). The protocol included the following sequences: axial and coronal T2-weighted spin echo (SE) images at 5mm; axial T1w gradient echo (GE) images at 5mm before contrast media injection and during the arterial, pancreatic parenchymal, portal-venous, systemic venous and delayed phases (as determined by testing bolus injection); axial unidirectional diffusion-weighted imaging at b-values of 0, 50, 300 and 600 with echo-planar imaging (EPI) read-out and ADC map calculation. ADC map reconstructions were 5.5x5.5x5 mm (xyz) to a 192x192 voxel matrix. Furthermore, single-shot T2w magnetic resonance cholangiopancreatography (MRCP) was performed and reconstructed as a radial maximum intensity projection (MIP) series. The imaging protocol, and the technical software and hardware specifications of the MRI machine remained unaltered during the data acquisition period.

Image segmentation

The datasets were exported in pseudonymized form to a segmentation workstation running ITK-SNAP v. 3.8.0 (beta). Segmentation was performed under radiological reporting room conditions by consensus reading of two experienced observers (G.K. and S.Z.). After a period of two weeks, datasets were shuffled by a third person (F.L.) and segmented again by the same observers. Segmentations were then quality-controlled by an abdominal radiologist with >10 years of experience in pancreatic MRI (R.B) and the best segmentations retained. Segmentation was performed manually in the $b = 600$ images and transferred to the ADC maps. All other sequences were available to observers for anatomical correlation.

Biostatistical and machine learning modeling

For assessing bias due to clinical confounders, overall survival time was evaluated by a multivariate *Cox proportional hazards* model. The distributions of covariates were compared between groups with different histopathological subtype using *Fisher's exact test*.

Biostatistical modeling was performed using the Python (v.3.7.3) packages *Lifelines* and *Scikit-Survival* [11]. Kaplan-Meier-Plots were drawn in GraphPad Prism (v.8). For all inferential statistical procedures, a P-value of <0.05 was considered statistically significant.

Image postprocessing, feature extraction, feature preprocessing, feature engineering and machine learning modeling are described in [S1 File](#). In brief, radiomic features were derived using *PyRadiomics* (v. 2.1) [12] yielding a total of 1606 features, of which 40 were retained after exclusion of features with low-variance or repeated segmentation instability. Unless otherwise noted, a randomized, 10-fold shuffle-splitting cross-validation strategy was used with 70% ($N = 38$) of the cohort used for training and 30% ($N = 17$) for testing.

For the prediction of tumor subtype, a supervised *Gradient Boosted Decision Tree* model (*XGBoost* [13]), instantiated as a binary classifier within the Python library *scikit-learn* was fit with histopathological subtype as a binary label to the radiomic features and tested for predictive sensitivity, specificity and ROC-AUC. Significance testing for model evaluation metrics was carried out using permutation testing [14]. The threshold probability for classification was the default value of .50. Feature importance was assessed by the inbuilt feature importance classifier (using the "gain" parameter).

For survival modeling, two strategies were followed: For assessing the agreement between model survival predictions and actual patient survival, the cohort was split into an interleaved training/testing set of 70%/30% and a stochastic gradient boosted survival regression model was trained using a subsampling tree building approach. The concordance indices between the resulting proportional hazards survival model and the actual patient survival of the unseen holdout set were calculated separately for disease-free and overall survival. 95% confidence intervals were calculated by bootstrap resampling. For evaluating the capacity of the model to separate between patients with a high and low survival risk, disease-free and overall survival were assessed for patients for whom the algorithm predicted a KRT81+ subtype (designated as KRT81+ or *high risk* signature) vs. patients predicted KRT81- (*low risk* signature). The Kaplan-Meier/ log-rank methods were used to compare the survival functions in the cross-validation folds.

Finally, for the assessment of chemotherapy sensitivity, disease-free and overall survival were evaluated stratified by radiomic signature (KRT81+/*high risk* vs. KRT81-/*low risk*) and chemotherapy regimen under the assumption of a differential response of the subtypes identified by the signature to the applied chemotherapy regimen, using the Kaplan-Meier/log-rank methods in the cross-validation folds.

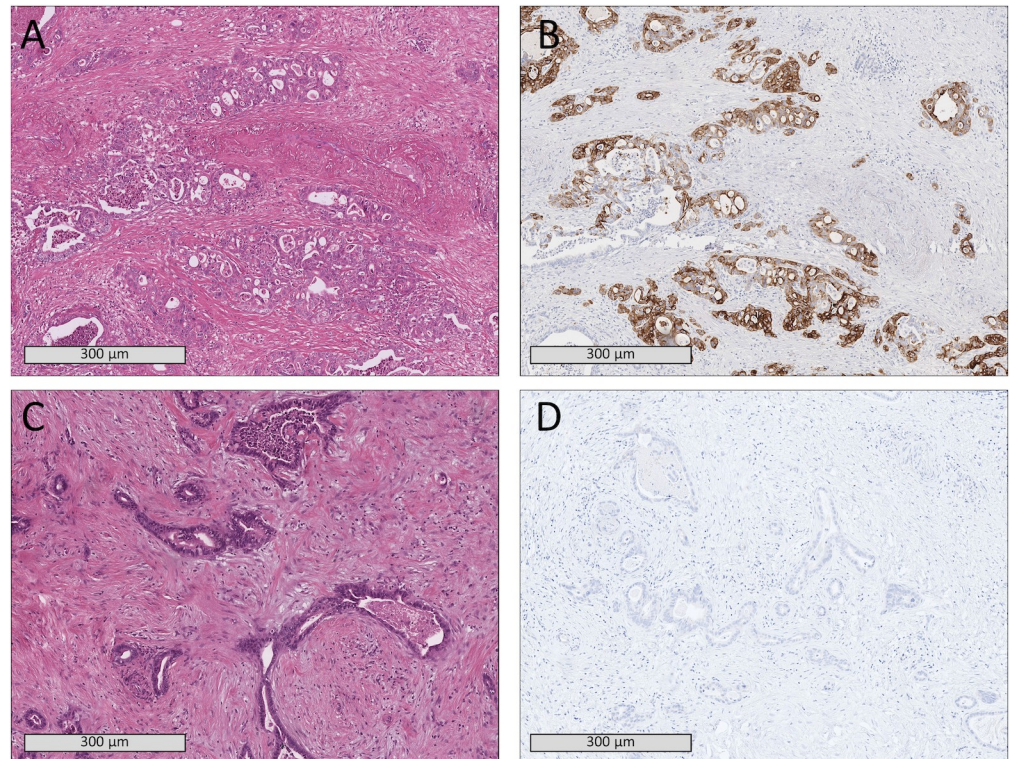


Fig 1. Histopathological samples of two patients showing comparable tissue morphology in H&E staining (A,C) but a KRT81+ subtype (B) in one patient and KRT81- subtype (D) in the other patient.

<https://doi.org/10.1371/journal.pone.0218642.g001>

Histopathological workup

Histopathological staining and immunohistochemical workup were performed by application of surrogate markers to determine the molecular subtype of PDAC based on the previously established immunohistochemical protocol described in [5]. In brief, 2 μ m sections were stained for HNF1a and KRT81+ and tumors were categorized into either one of two classes: KRT81+/[HNF1a-] or KRT81-/[HNF1a+] (Fig 1). Tumors positive or negative for both markers were excluded, the former due to recently reported suspicions of contamination with acinar cells [15], the latter due to unclassifiability.

Results

The molecular subtype of PDAC was significantly associated with overall survival. Patients with a KRT81+ subtype experienced significantly diminished overall survival (7.0 [1.93 to 29.0] vs. 22.6 [2.63 to 96.97] months median survival, HR 4.03 log-rank-test $P = <0.001$, Fig 2, Table 1). No other covariate was significantly associated with overall survival in this cohort and the baseline distribution of clinical covariates did not differ significantly between the two patient subcohorts (Table 2).

The machine learning algorithm achieved a mean \pm STDEV sensitivity, specificity and ROC-AUC of 0.90 \pm 0.07, 0.92 \pm 0.11, and 0.93 \pm 0.07, respectively; all $P = 0.01$ (Fig 3).

The feature importance evaluation of the algorithm yielded 13 radiomic parameters with an importance greater than zero for the classification process. Among these, entropy, a radiomic feature derived from the histogram of the original image and signifying the degree of heterogeneity in the tumor region [16], was classified as the most important feature by a large margin.

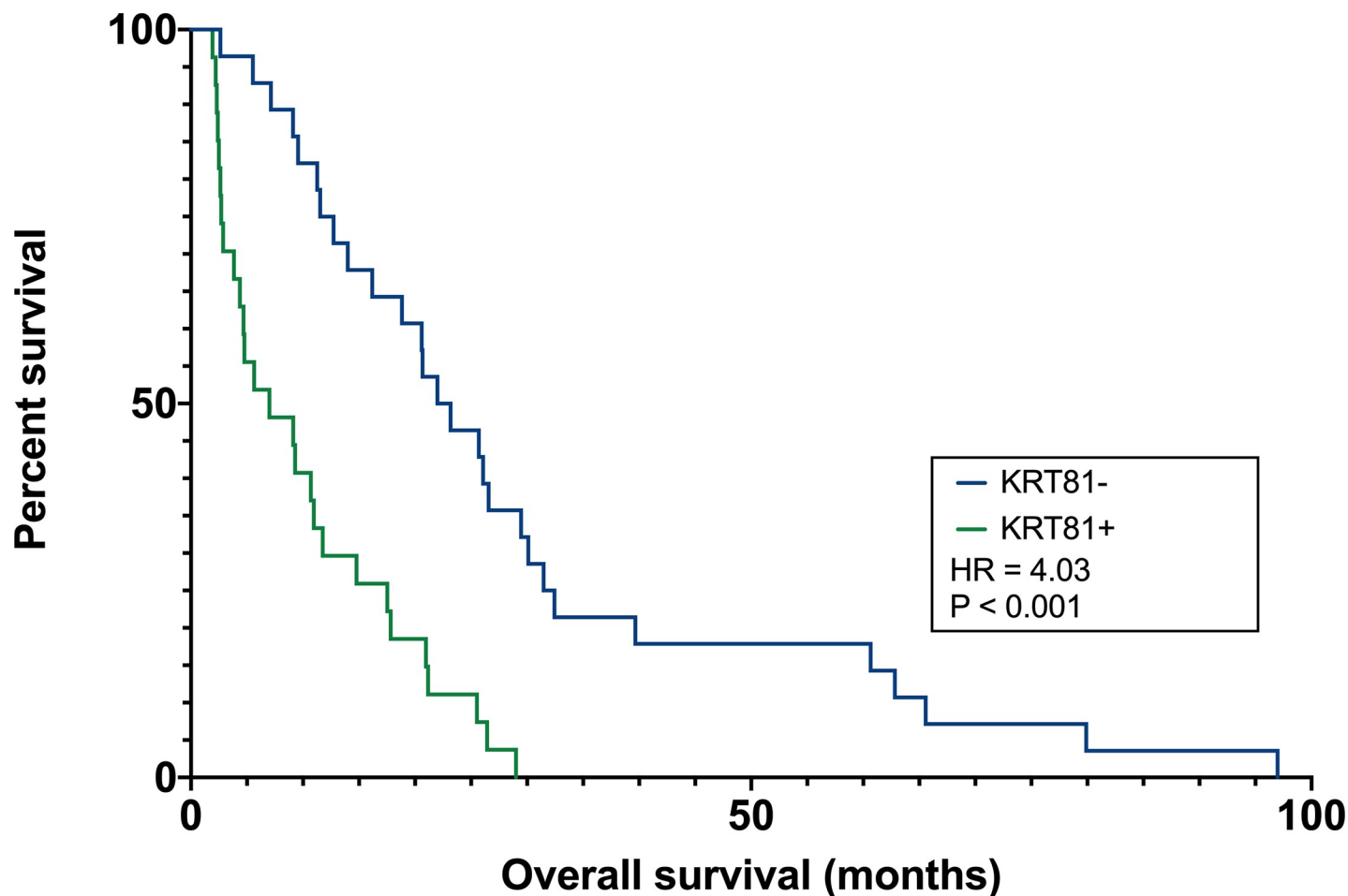


Fig 2. Patients with a KRT81+ subtype experienced significantly diminished overall survival.

<https://doi.org/10.1371/journal.pone.0218642.g002>

Table 1. Multivariate Cox proportional hazards analysis results of clinical parameters.

Parameter	coef	exp(coef)	p	lower 0.95	upper 0.95
Subtype (KRT 81+ vs. KRT 81-)	1.44	4.03	<0.001	0.76	2.12
pN (0 vs. 1)	1.20	3.32	0.20	-0.63	3.03
Age	0.02	1.02	0.30	-0.01	0.05
pM (0 vs. 1)	0.50	1.65	0.30	-0.45	1.44
Palliative CTX (Gem mono vs. FOLFIRINOX)	-0.39	0.68	0.32	-1.16	0.38
pT	0.22	1.25	0.36	-0.25	0.70
Tumor Volume	-0.01	0.99	0.38	-0.04	0.01
Grading (2 vs. 3)	0.21	1.24	0.49	-0.40	0.83
Adjuvant CTX (Gem-based vs. None)	-0.45	0.64	0.53	-1.85	0.95
LNR	-2.69	0.07	0.65	-14.23	8.85
R (0 vs. 1)	-0.14	0.87	0.69	-0.84	0.56
Sex (F vs. M)	0.08	1.08	0.81	-0.57	0.73
ECOG (0 vs. 1)	-0.06	0.94	0.88	-0.81	0.70

<https://doi.org/10.1371/journal.pone.0218642.t001>

Table 2. Distribution of clinical parameters between the cohorts with KRT81+ and KRT81- tumor subtypes alongside crosstabulation results.

Parameter	KRT 81+ Subcohort (27)	STDEV	KRT 81- Subcohort (28)	STDEV	P
Age	67	11.7	65	10.5	.52
Adjuvant CTX	Gem-based: 25, Did not receive: 2		Gem-based: 26, Did not receive: 2		-
Palliative CTX	Gem-based: 14, FOLFIRINOX: 13		Gem-based: 16, FOLFIRINOX: 12		.78
Experienced Event	Yes: 27		Yes: 28		-
G	2: 16, 3: 11		2: 15, 3: 13		.79
pM	0: 22, 1: 5		0: 23, 1: 5		1.0
pN	0: 6, 1: 21		0: 8, 1: 20		.76
pT	1: 3, 2: 2, 3: 22		1: 3, 2: 3, 3: 22		-
R	0: 20, 1: 7		0: 21, 1: 7		1.0
Sex	Female: 12, Male: 15		Female: 13, Male: 15		1.0
ECOG	0: 11, 1: 16		0: 13, 1: 15		.79
Tumor Volume (ml)	16.4	15.6	15.0	14.0	.72
Lymph Node Ratio	0.12	0.07	0.10	0.07	.29

P: Fisher's exact test P

<https://doi.org/10.1371/journal.pone.0218642.t002>

All features alongside their importance metrics can be found in Table 3 and further information about the radiomic parameters can be found in S1 File.

To test the association of the radiomic parameters with patient survival, a survival regression model including all radiomic features was developed. The concordance index between model predictions and actual survival on the entire dataset was 0.76 ± 0.05 [95% CI 0.66–0.86] for disease-free and 0.71 ± 0.06 [95% CI 0.60–0.80] for overall survival.

Patients with a KRT81+ (*high risk*) radiomic signature experienced diminished disease-free and overall survival. Over the 10 cross-validation folds, the median hazard ratio for patients with a *high risk* vs. a *low risk* radiomic signature was 3.08 [range 0.46 to 3.37] for disease-free and 3.04 [range 0.27 to 3.36] for overall survival. The radiomics-derived stratification led to a

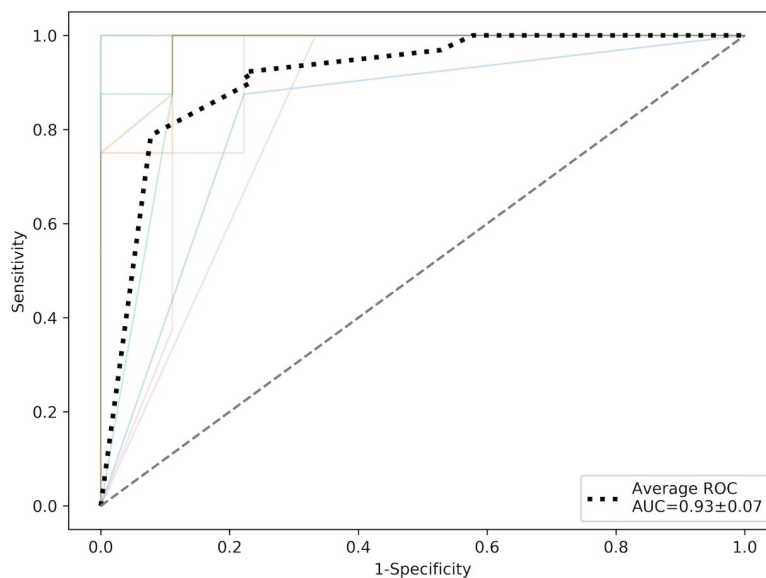


Fig 3. ROC curves (colored) and average ROC-curve (black dotted) over 10 random stratified shuffle-splits of the dataset.

<https://doi.org/10.1371/journal.pone.0218642.g003>

Table 3. Radiomic features alongside their importance as ranked by the algorithm.

Feature	Importance
original_firstorder_Entropy	0.73
gradient_firstorder_Kurtosis	0.10
log-sigma-1-0-mm-3D_glcmm_Imc2	0.09
log-sigma-3-0-mm-3D_firstorder_Kurtosis	0.05
original_glszm_SizeZoneNonUniformityNormalized	0.005
wavelet-HHL_glcmm_Imc2	0.005
wavelet-HHL_glszm_SmallAreaEmphasis	0.004
wavelet-HHL_glszm_ZonePercentage	0.003
original_shape_Maximum2DDiameterRow	0.003
log-sigma-2-0-mm-3D_glszm_SmallAreaHighGrayLevelEmphasis	0.002
original_glszm_LargeAreaLowGrayLevelEmphasis	0.001
wavelet-HLL_glszm_ZonePercentage	0.001
wavelet-LHL_firstorder_Kurtosis	0.0005

<https://doi.org/10.1371/journal.pone.0218642.t003>

statistically significant difference in survival functions (*log-rank-test* $p < 0.05$) in 7 out of 10 cross-validation folds for disease-free and in 8 out of 10 folds for overall survival. Figs 4 and 5 show exemplary survival curves. All survival curves can be found in the [S1 File](#).

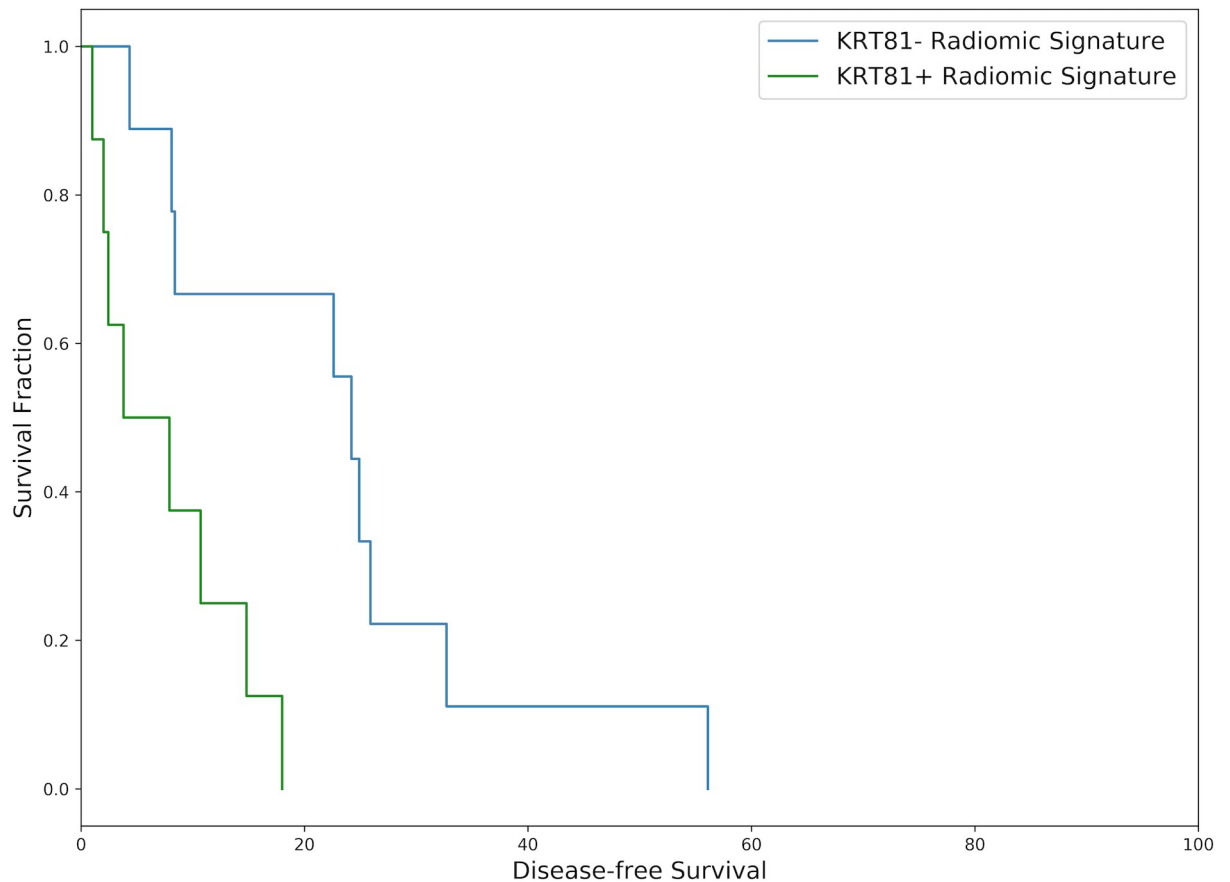


Fig 4. Exemplary Kaplan-Meier disease-free survival curve drawn from the 10 cross-validation folds representing the median hazard ratio. Patients with a *high risk* radiomic phenotype experienced significantly diminished survival (7.90 vs. 24.20 months median DFS, *log-rank-test* $p = 0.004$, HR = 3.17, N = 17).

<https://doi.org/10.1371/journal.pone.0218642.g004>

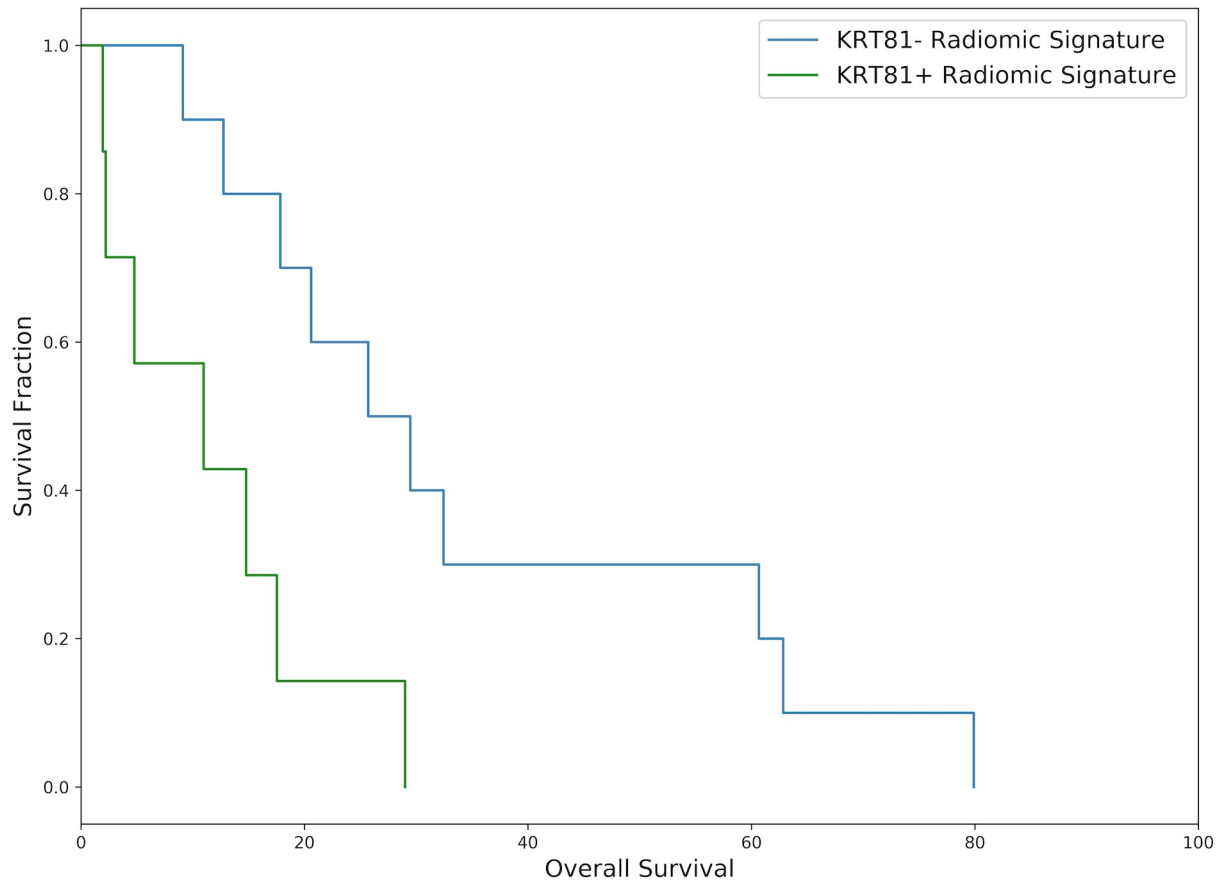


Fig 5. Exemplary Kaplan-Meier overall survival curve drawn from the 10 cross-validation folds representing the median hazard ratio. Patients with a *high risk* radiomic phenotype experienced significantly diminished survival (10.97 vs. 25.70 months median OS, *log-rank-test* $p = 0.006$, HR = 3.03, N = 17).

<https://doi.org/10.1371/journal.pone.0218642.g005>

For evaluating the potential of the radiomic algorithm to inform the choice of palliative chemotherapy regimen, overall survival was evaluated stratified by radiomic signature (KRT81+/high risk vs. KRT81-/low risk) and assessed separately by selected chemotherapy regimen. Patients with a KRT81+ radiomic signature experienced prolonged overall survival under palliative gemcitabine therapy compared to FOLFIRINOX (median HR 1.13 [range 0.03 to 2.57]) but statistical significance was only observed for *log-rank-tests* of 2 out of 10 cross-validation folds. Inversely, patients with a KRT81- radiomic signature experienced improved overall survival under palliative FOLFIRINOX compared to gemcitabine (median HR 2.89 [range 0.99 to 3.34] with statistical significance observed in 6 out of 10 cross-validation folds. Figs 6 and 7 show exemplary survival curves and all survival curves can be found in the S1 File.

Lastly, overall survival was evaluated separately for histopathological subtypes stratified by chemotherapy regimen. Patients with a KRT81+ histopathological subtype who received gemcitabine-based palliative chemotherapy experienced significantly improved survival compared to patients with KRT81+ tumors who received FOLFIRINOX (10.14 vs. 3.8 months median survival, HR 2.33, $P = 0.037$, Fig 8). Conversely, KRT81- subtype patients experienced significantly improved survival under FOLFIRINOX chemotherapy compared to gemcitabine-based regimens (30.8 vs. 13.4 months median survival, HR 2.41, $P = 0.027$, Fig 9).

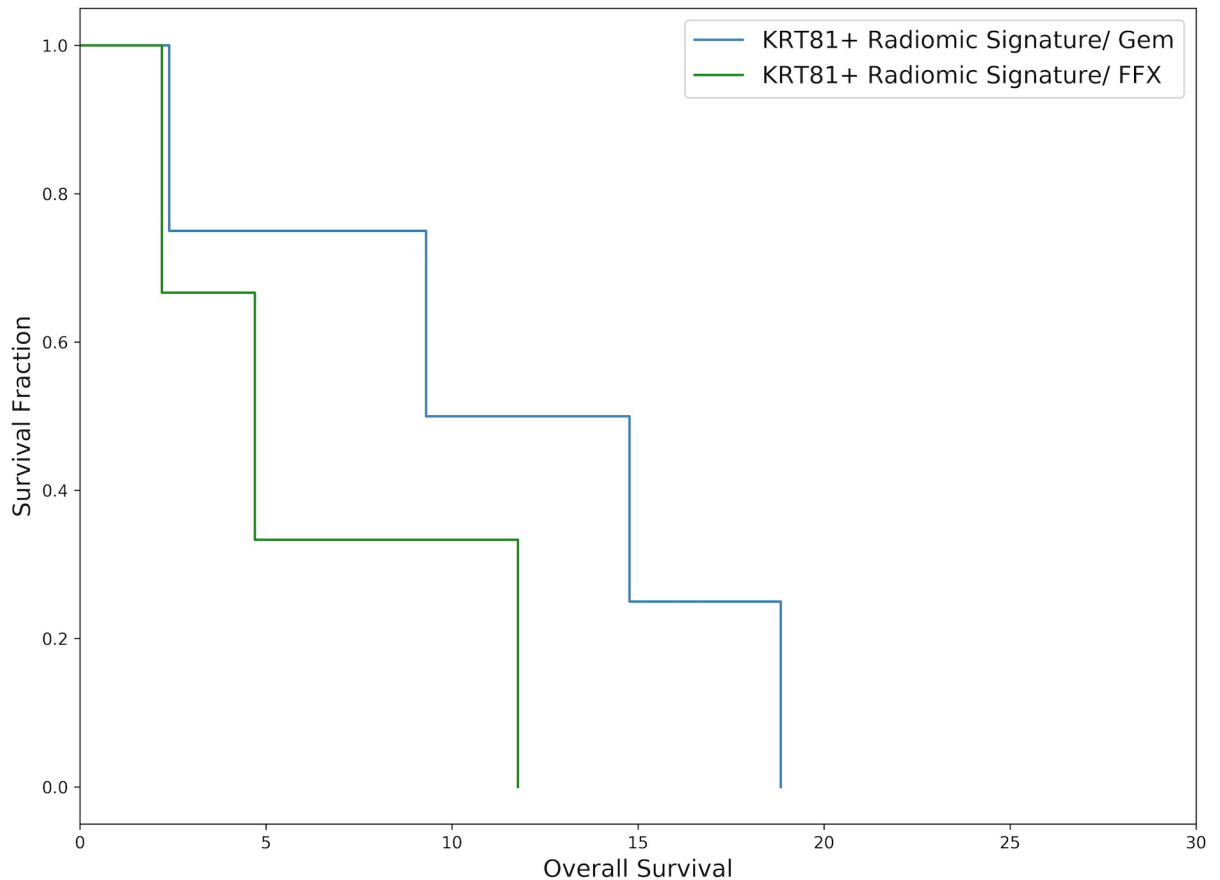


Fig 6. Exemplary Kaplan-Meier plot of patients with a predicted KRT81+ / high risk radiomic signature drawn from the 10 cross-validation folds and representing the median hazard ratio. Patients who received gemcitabine experienced improved survival, although no statistical significance is observed in this case (median survival 4.70 vs 9.30 months, $p = 0.23$, HR = 1.03, N = 7).

<https://doi.org/10.1371/journal.pone.0218642.g006>

Discussion

In this exploratory study, we demonstrate that radiomic analysis of ADC maps paired with machine-learning modeling can discriminate with high sensitivity and specificity between two groups of histomorphologically defined molecular subtypes of pancreatic ductal adenocarcinoma (PDAC), associated with significantly different disease-free and overall survival and predictive of distinct responses to commonly employed chemotherapeutic regimens in the palliative setting. Although our findings should be interpreted mindful of the small cohort size, we provide evidence for the utility of radiomics and machine learning for the non-invasive therapy stratification and survival risk-assessment of pancreatic cancer patients.

The potential of non-invasive imaging-derived biomarkers (from non-perceptual image features or source data) has been demonstrated in several studies with the prediction of tumor genetics and patient outcome [17–19]. However, their widespread application beyond proof-of-principle studies requires the identification of stable and reproducible parameters, embedded within a standardized and quality-controlled workflow [20–23].

Among the parameters tested for classification in our study, Entropy was ranked the most important by the algorithm. Entropy and Entropy-related features, which express disorder and heterogeneity of the image and -by extension- are hypothesized to mirror tumoral

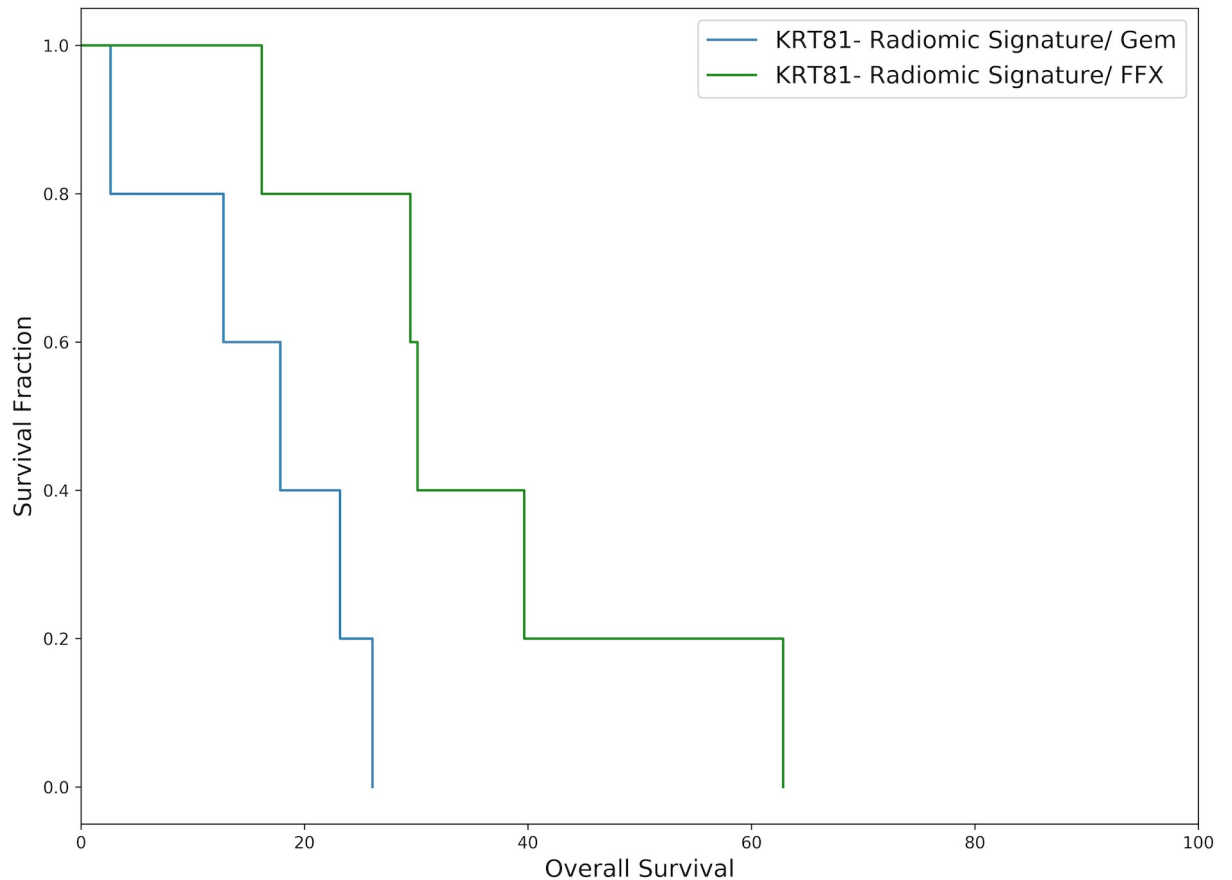


Fig 7. Exemplary Kaplan-Meier plot of patients with a predicted KRT81-/*low risk* radiomic signature drawn from the 10 cross-validation folds and representing the median hazard ratio. Patients who received FOLFIRINOX experienced significantly improved survival, (median survival 17.83 vs 30.10 months, $p = 0.01$, HR = 2.57, N = 10).

<https://doi.org/10.1371/journal.pone.0218642.g007>

heterogeneity itself, have been demonstrated in (meta-)analyses of several different tumor entities and across imaging modalities as reliable and repeatable quantitative parameters [24–26].

Considering the sampling errors immanent in this histopathologically heterogeneous tumor entity, the complexity of mutational events (e.g. variable amounts of mutational *Kras* [27] and the likelihood of ongoing transitional processes, Entropy as a continuous variable can be imagined as a non-invasive measure of the KRT81+ partition of the tumor region. To test this hypothesis would require an integrated whole-tumor analysis, including high resolution, data-rich imaging, histopathology and molecular profiling [28].

The rapid evolution of new therapeutic options in the treatment of PDAC requires the development of markers for a reliable pre-therapeutic patient stratification and -in light of the above-mentioned tumoral plasticity, therapy monitoring. Conroy et al. demonstrated significantly improved survival rates of FOLFIRINOX over Gemcitabine monotherapy in the palliative setting [3]. However, the COMPASS trial [7] demonstrated differential response of the basal-like versus non-basal-like PDAC subtypes to FOLFIRINOX treatment, which is well in accordance with our study results. If further validated in prospective trials, these findings could have tremendous implications in patient stratification and subtype-guided therapy selection. In addition, targeted therapies such as Olaparib, are highly effective yet even more specific for a certain molecular profile [29] and many new targeted, stroma- and immune-based treatment strategies are being explored. This increasing complexity requires robust and cost-

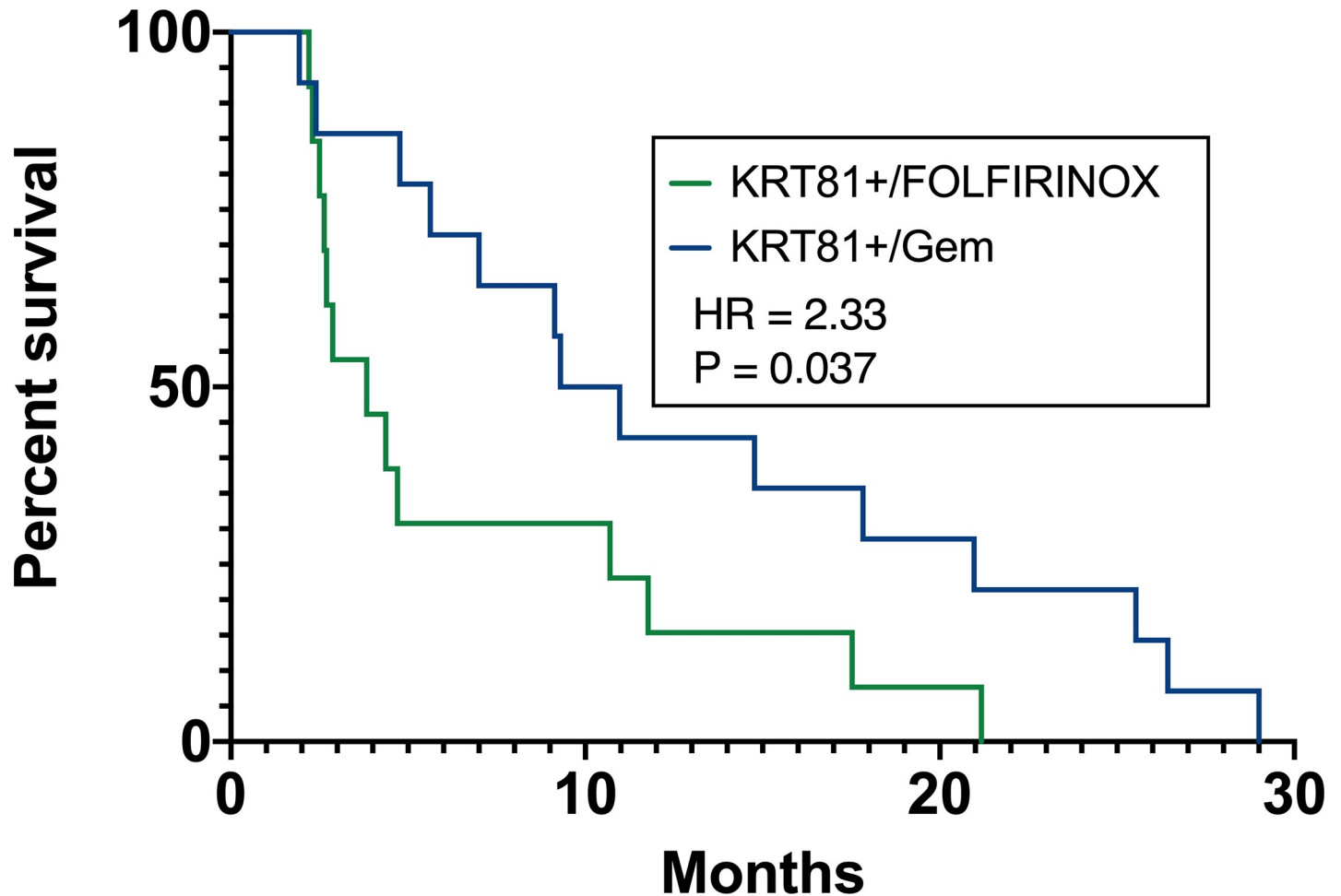


Fig 8. Patients with KRT81+ subtype experience longer overall survival under palliative gemcitabine chemotherapy.

<https://doi.org/10.1371/journal.pone.0218642.g008>

efficient tools for clinically relevant patient stratification to best leverage current knowledge and advance the field. Informed decision based on molecular profiling (microdissection and genome sequencing) as applied in the COMPASS trial faces serious limitations (i.e. sampling error, high cost) and is therefore currently not feasible in routine patient care. Quantitative noninvasive imaging, and especially radiation and contrast-free quantitative modalities such as DWI may serve this purpose and are thus excellent candidates for exploration in a prospective trial design.

Limitations of this study are the small cohort size, precluding statistical significance in the survival predictions and necessitating a cross-validation approach, and the lack of an external testing cohort as well as the retrospective, single-center nature of the investigation. Such issues are still common in the imaging field and compounded by the lack of standardization in sequence acquisition between institutions and of overarching registers or study centers permitting patient pooling. Recently, initiatives have arisen to combat some of these issues by harmonization of MRI protocols [30] and the standardization of imaging markers [31].

In conclusion, our study is an exploratory venture into the field of quantitative imaging analysis and radiology/pathology-correlation in PDAC. We encourage the validation of our findings in a larger cohort and in a prospective trial design.

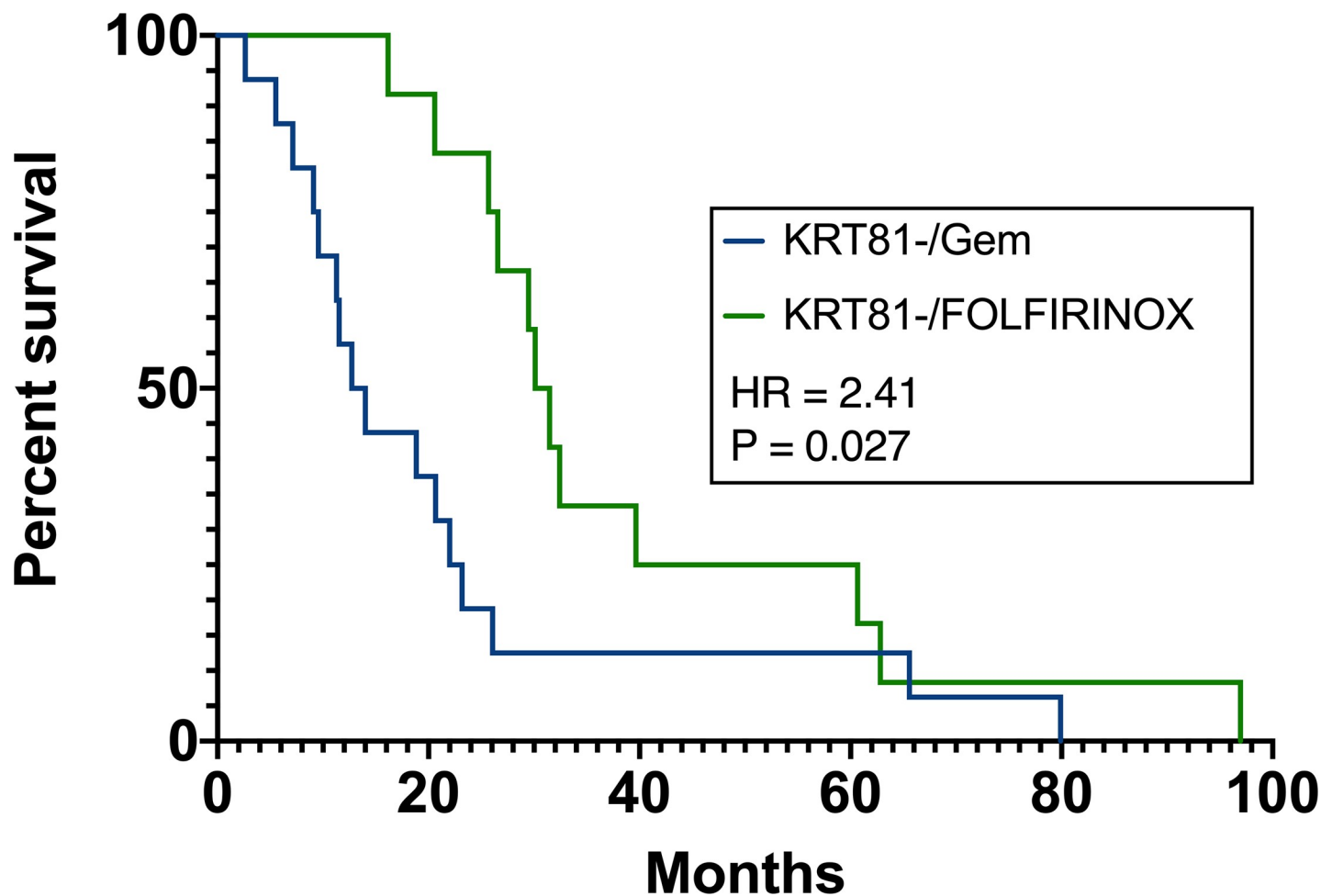


Fig 9. Patients with KRT81- subtype experience improved survival under palliative FOLFIRINOX chemotherapy.

<https://doi.org/10.1371/journal.pone.0218642.g009>

Supporting information

S1 File.
(PDF)

Acknowledgments

The authors wish to thank Irina Heid for the ongoing support.

Author Contributions

Conceptualization: Georgios Kaissis, Jens T. Siveke, Rickmer Braren.

Data curation: Georgios Kaissis, Rickmer Braren.

Formal analysis: Georgios Kaissis, Sebastian Ziegelmayer, Fabian Lohöfer, Katja Steiger, Alexander Muckenhuber, Hsi-Yu Yen.

Funding acquisition: Wilko Weichert, Rickmer Braren.

Investigation: Georgios Kaissis, Sebastian Ziegelmayr, Alexander Muckenhuber, Hsi-Yu Yen, Jens T. Siveke.

Methodology: Georgios Kaissis, Fabian Lohöfer, Katja Steiger, Alexander Muckenhuber, Hsi-Yu Yen, Rickmer Braren.

Project administration: Hana Algül, Ernst Rummeny, Helmut Friess, Roland Schmid, Wilko Weichert, Rickmer Braren.

Software: Georgios Kaissis, Sebastian Ziegelmayr.

Supervision: Fabian Lohöfer, Hana Algül, Ernst Rummeny, Helmut Friess, Roland Schmid, Wilko Weichert, Jens T. Siveke, Rickmer Braren.

Validation: Georgios Kaissis, Jens T. Siveke, Rickmer Braren.

Visualization: Georgios Kaissis, Katja Steiger, Hsi-Yu Yen.

Writing – original draft: Georgios Kaissis, Sebastian Ziegelmayr, Fabian Lohöfer, Jens T. Siveke, Rickmer Braren.

Writing – review & editing: Georgios Kaissis, Sebastian Ziegelmayr, Fabian Lohöfer, Jens T. Siveke, Rickmer Braren.

References

1. Von Hoff DD, Ervin TJ, Arena FP, Chiorean EG, Infante JR, Moore MJ, et al. Results of a randomized phase III trial (MPACT) of weekly nab-paclitaxel plus gemcitabine versus gemcitabine alone for patients with metastatic adenocarcinoma of the pancreas with PET and CA19-9 correlates. *J Clin Oncol*. 2013; 31(15_suppl):4005.
2. Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante J, Moore M, et al. Increased Survival in Pancreatic Cancer with nab-Paclitaxel plus Gemcitabine. *N Engl J Med*. 2013 Oct; 369(18):1691–703. <https://doi.org/10.1056/NEJMoa1304369> PMID: 24131140
3. Conroy T, Desseigne F, Ychou M, Bouche O, Guimbaud R, Becouarn Y, et al. FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. *N Engl J Med* [Internet]. 2011 May 12; 364(19):1817–25. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/21561347> <https://doi.org/10.1056/NEJMoa1011923> PMID: 21561347
4. Collisson EA, Bailey P, Chang DK, Biankin A V. Molecular subtypes of pancreatic cancer. *Nat Rev Gastroenterol Hepatol* [Internet]. 2019 Apr 4; 16(4):207–20. Available from: <https://doi.org/10.1038/s41575-019-0109-y> PMID: 30718832
5. Muckenhuber A, Berger AK, Schlitter AM, Steiger K, Konukiewitz B, Trumpp A, et al. Pancreatic Ductal Adenocarcinoma Subtyping Using the Biomarkers Hepatocyte Nuclear Factor-1A and Cytokeratin-81 Correlates with Outcome and Treatment Response. *Clin Cancer Res* [Internet]. 2018; 24(2):351–9. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/29101303> <https://doi.org/10.1158/1078-0432.CCR-17-2180> PMID: 29101303
6. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med* [Internet]. 2011; 17(4):500–3. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/21460848> <https://doi.org/10.1038/nm.2344> PMID: 21460848
7. Aung KL, Fischer SE, Denroche RE, Jang G-H, Dodd A, Creighton S, et al. Genomics-Driven Precision Medicine for Advanced Pancreatic Cancer: Early Results from the COMPASS Trial. *Clin Cancer Res* [Internet]. 2018 Mar 15; 24(6):1344–54. Available from: <https://doi.org/10.1158/1078-0432.CCR-17-2994> PMID: 29288237
8. Heid I, Steiger K, Trajkovic-Arsic M, Settles M, Esswein MR, Erkan M, et al. Co-clinical Assessment of Tumor Cellularity in Pancreatic Cancer. *Clin Cancer Res* [Internet]. 2017; 23(6):1461–70. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/27663591> <https://doi.org/10.1158/1078-0432.CCR-15-2432> PMID: 27663591
9. Trajkovic-Arsic M, Heid I, Steiger K, Gupta A, Fingerle A, Wörner C, et al. Apparent Diffusion Coefficient (ADC) predicts therapy response in pancreatic ductal adenocarcinoma. *Sci Rep* [Internet]. 2017 Dec 6; 7(1):17038. Available from: <http://www.nature.com/articles/s41598-017-16826-z> <https://doi.org/10.1038/s41598-017-16826-z> PMID: 29213099

10. Kaissis G, Ziegelmayer S, Lohöfer F, Algül H, Eiber M, Weichert W, et al. A prospectively validated machine learning model for the prediction of survival and tumor subtype in pancreatic ductal adenocarcinoma. *bioRxiv*. 2019 Jan;643809.
11. Pölsterl S, Gupta P, Wang L, Conjeti S, Katouzian A, Navab N. Heterogeneous ensembles for predicting survival of metastatic, castrate-resistant prostate cancer patients. *F1000Research* [Internet]. 2017 Jul 6; 5:2676. Available from: <https://f1000research.com/articles/5-2676/v3>
12. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* [Internet]. 2017 Nov 1; 77(21): e104–7. Available from: <https://doi.org/10.1158/0008-5472.CAN-17-0339> PMID: 29092951
13. Chen T, Guestrin C. XGBoost. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD '16*. New York, New York, USA: ACM Press; 2016. p. 785–94.
14. Ojala M, Garriga GC. Permutation Tests for Studying Classifier Performance. In: *2009 Ninth IEEE International Conference on Data Mining*. IEEE; 2009. p. 908–13.
15. Puleo F, Nicolle R, Blum Y, Cros J, Marisa L, Demetter P, et al. Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. *Gastroenterology* [Internet]. 2018 Dec; 155(6):1999–2013.e3. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0016508518349199> <https://doi.org/10.1053/j.gastro.2018.08.033> PMID: 30165049
16. PyRadiomics. Feature Documentation. Available from: <https://pyradiomics.readthedocs.io/en/latest/features.html>
17. Hanania AN, Bantis LE, Feng Z, Wang H, Tamm EP, Katz MH, et al. Quantitative imaging to evaluate malignant potential of IPMNs. *Oncotarget* [Internet]. 2016 Dec 27; 7(52):85776–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27588410> <https://doi.org/10.18632/oncotarget.11769> PMID: 27588410
18. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RTH, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol* [Internet]. 2015; 114(3):345–50. Available from: <https://doi.org/10.1016/j.radonc.2015.02.015> PMID: 25746350
19. Lao J, Chen Y, Li ZC, Li Q, Zhang J, Liu J, et al. A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme. *Sci Rep* [Internet]. 2017; 7(1):1–8. Available from: <https://doi.org/10.1038/s41598-016-0028-x>
20. Caramella C, Allorant A, Orlhac F, Bidault F, Asselain B, Ammari S, et al. Can we trust the calculation of texture indices of CT images? A phantom study. *Med Phys* [Internet]. 2018 Feb 14; Available from: <http://doi.wiley.com/10.1002/mp.12809>
21. Verma V, Simone CB, Krishnan S, Lin SH, Yang J, Hahn SM. The Rise of Radiomics and Implications for Oncologic Management. *J Natl Cancer Inst*. 2017;
22. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: The process and the challenges. *Magn Reson Imaging* [Internet]. 2012; 30(9):1234–48. Available from: <https://doi.org/10.1016/j.mri.2012.06.010> PMID: 22898692
23. Zhao B, Tan Y, Tsai WY, Qi J, Xie C, Lu L, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep* [Internet]. 2016; 6:1–7. Available from: <https://doi.org/10.1038/s41598-016-0001-8>
24. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and Reproducibility of Radiomic Features: A Systematic Review. *Int J Radiat Oncol* [Internet]. 2018 Jun; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0360301618309052>
25. Khalvati F, Zhang Y, Baig S, Lobo-Mueller EM, Karanicolas P, Gallinger S, et al. Prognostic Value of CT Radiomic Features in Resectable Pancreatic Ductal Adenocarcinoma. *Sci Rep* [Internet]. 2019 Dec 1; 9(1):5449. Available from: <http://www.nature.com/articles/s41598-019-41728-7> <https://doi.org/10.1038/s41598-019-41728-7> PMID: 30931954
26. Eilaghi A, Baig S, Zhang Y, Zhang J, Karanicolas P, Gallinger S, et al. CT texture features are associated with overall survival in pancreatic ductal adenocarcinoma—a quantitative analysis. *BMC Med Imaging*. 2017; 17(1):1–7. <https://doi.org/10.1186/s12880-016-0171-7>
27. Mueller S, Engleitner T, Maresch R, Zukowska M, Lange S, Kaltenbacher T, et al. Evolutionary routes and KRAS dosage define pancreatic cancer phenotypes. *Nature* [Internet]. 2018 Feb 24; 554(7690):62–8. Available from: <http://www.nature.com/articles/nature25459> <https://doi.org/10.1038/nature25459> PMID: 29364867
28. Bailey DL, Pichler BJ, Gückel B, Antoch G, Barthel H, Bhujwalla ZM, et al. Combined PET/MRI: Global Warming—Summary Report of the 6th International Workshop on PET/MRI, March 27–29, 2017, Tübingen, Germany. *Mol Imaging Biol* [Internet]. 2018 Feb 2; 20(1):4–20. Available from: <https://doi.org/10.1007/s11307-017-1123-5> PMID: 28971346

29. Kaufman B, Shapira-Frommer R, Schmutzler RK, Audeh MW, Friedlander M, Balmaña J, et al. Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. *J Clin Oncol*. 2015 Jan; 33(3):244–50. <https://doi.org/10.1200/JCO.2014.56.2728> PMID: 25366685
30. Bach M, Röhke M, Henzler T, Kreft M, Amler B SH. Standardized and quality assured prostate diffusion MRI. *Eur Congr Radiol* [Internet]. 2019; Available from: https://www.radiagnostix.de/fileadmin/radiagnostix/PDF/Artikel/Poster_ECR2019_C-2163_002_.pdf
31. Zwanenburg A, Leger S, Vallières M, Löck S, Initiative for the IBS. Image biomarker standardisation initiative. *CoRR* [Internet]. 2016;abs/1612.0. Available from: <http://arxiv.org/abs/1612.07003>



Article

Image-Based Molecular Phenotyping of Pancreatic Ductal Adenocarcinoma

Georgios A. Kaissis ^{1,2,†}, Sebastian Ziegelmayr ^{1,†}, Fabian K. Lohöfer ^{1,†}, Felix N. Harder ¹, Friederike Jungmann ¹, Daniel Sasse ¹, Alexander Muckenhuber ³, Hsi-Yu Yen ³, Katja Steiger ³, Jens Siveke ^{4,5}, Helmut Friess ⁶, Roland Schmid ⁷, Wilko Weichert ³, Marcus R. Makowski ¹ and Rickmer F. Braren ^{1,*}

¹ Technical University of Munich, School of Medicine, Department of Diagnostic and Interventional Radiology, 81675 Munich, Germany; g.kaissis@tum.de (G.A.K.); ga89rog@mytum.de (S.Z.); fabian.lohoefer@tum.de (F.K.L.); felix.harder@tum.de (F.N.H.); friederike.jungmann@tum.de (F.J.); daniel.sasse@tum.de (D.S.); marcus.makowski@tum.de (M.R.M.)

² Imperial College of Science, Technology and Medicine, Faculty of Engineering, Department of Computing, SW7 2AZ London, UK

³ Technical University of Munich, School of Medicine, Institute for Pathology, 81675 Munich, Germany; Alexander.Muckenhuber@tum.de (A.M.); hsi-yu.yen@tum.de (H.-Y.Y.); katja.steiger@tum.de (K.S.); wilko.weichert@tum.de (W.W.)

⁴ Institute of Developmental Cancer Therapeutics, West German Cancer Center, University Hospital Essen, 45147 Essen, Germany; Jens.Siveke@uk-essen.de

⁵ Division of Solid Tumor Translational Oncology, German Cancer Consortium (DKTK, parter site Essen, Germany) and German Cancer Research Center, DKFZ, 69120 Heidelberg, Germany

⁶ Technical University of Munich, School of Medicine, Surgical Clinic and Policlinic, 81675 Munich, Germany; helmut.friess@tum.de

⁷ Technical University of Munich, School of Medicine, Department of Internal Medicine II, 81675 Munich, Germany; direktion.med2@mri.tum.de

* Correspondence: rbraren@tum.de; Tel.: +49-89-4140-5627

† Authors contributed equally.

Received: 17 February 2020; Accepted: 4 March 2020; Published: 7 March 2020



Abstract: To bridge the translational gap between recent discoveries of distinct molecular phenotypes of pancreatic cancer and tangible improvements in patient outcome, there is an urgent need to develop strategies and tools informing and improving the clinical decision process. Radiomics and machine learning approaches can offer non-invasive whole tumor analytics for clinical imaging data-based classification. The retrospective study assessed baseline computed tomography (CT) from 207 patients with proven pancreatic ductal adenocarcinoma (PDAC). Following expert level manual annotation, Pyradiomics was used for the extraction of 1474 radiomic features. The molecular tumor subtype was defined by immunohistochemical staining for KRT81 and HNF1a as quasi-mesenchymal (QM) vs. non-quasi-mesenchymal (non-QM). A Random Forest machine learning algorithm was developed to predict the molecular subtype from the radiomic features. The algorithm was then applied to an independent cohort of histopathologically unclassifiable tumors with distinct clinical outcomes. The classification algorithm achieved a sensitivity, specificity and ROC-AUC (area under the receiver operating characteristic curve) of 0.84 ± 0.05 , 0.92 ± 0.01 and 0.93 ± 0.01 , respectively. The median overall survival for predicted QM and non-QM tumors was 16.1 and 20.9 months, respectively, log-rank-test $p = 0.02$, hazard ratio (HR) 1.59. The application of the algorithm to histopathologically unclassifiable tumors revealed two groups with significantly different survival (8.9 and 39.8 months, log-rank-test $p < 0.001$, HR 4.33). The machine learning-based analysis of preoperative (CT) imaging allows the prediction of molecular PDAC subtypes highly relevant for patient survival, allowing advanced pre-operative patient stratification for precision medicine applications.

Keywords: radiomics; pancreatic cancer; molecular subtypes

1. Introduction

In pancreatic ductal adenocarcinoma (PDAC), several lines of evidence suggest the existence of distinct subtypes with prognostic and predictive relevance. For example, Breast Cancer Gene (BRCA) 1/2 mutations have been identified in a subgroup of PDAC, and these patients exhibit an improved therapy response to the poly (ADP-ribose) polymerase (PARP) inhibitor Olaparib [1]. The majority of PDAC can be classified into two distinct subtypes based on transcriptome profiling and immunohistochemical staining of cytokeratin-81 (KRT81) and hepatocyte nuclear factor-1A (HNF1a) [2–4]: a so-called non-quasi-mesenchymal (non-QM), i.e., classical, epithelial KRT81-/HNF1a-subtype exhibiting slightly improved survival and therapy response, notably to the FOLFIRINOX regimen, while not responding as well to gemcitabine-based treatment. The so-called quasi-mesenchymal (QM), basal-like or KRT81+/HNF1a- subtype has a dismal overall survival and resistance towards virtually all currently employed therapy regimens. However, the QM subtype does exhibit a superior response to gemcitabine in comparison to the non-QM subtype [5]. These findings support differential treatment of patients based on individual molecular tumor make-up. Hitherto this has proven infeasible in clinical routine because of individual tumor heterogeneity and the immanent sampling errors of biopsy and lack of clinically suitable and sufficiently robust transcriptomic assays.

Recent developments in machine learning (ML)-based medical image analysis such as Radiomics provide encouraging examples of molecular phenotyping from imaging data. For instance, the non-invasive genetic profiling of lung carcinoma has been recently demonstrated [6], and imaging biomarkers have recently been shown to outperform the risk metrics defined in the current WHO classification of gliomas [7].

We recently reported on machine-learning approaches for the prediction of molecular subtypes and survival risk in PDAC patients from pre-operative magnetic resonance imaging (MRI) [8,9]. We noted that limited availability of MR imaging data, overall reduced image quality and the less-quantitative and unstandardized nature of MRI pose barriers to algorithm development and generalization.

To enhance clinical translation, we extend our previous results to computed tomography (CT) by training and validating an algorithm capable of discriminating between the QM and the non-QM subtypes of PDAC with high performance based on pre-operative CT imaging in a therapy-naïve surgical cohort of PDAC patients. We next applied this algorithm to histopathologically so-called unclassifiable, KRT81+/HNF1a+ tumors resulting in significant separation of overall survival time, suggesting identification of clinically distinct subgroups not identified by traditional histology.

2. Experimental Section

The study was designed as a retrospective cohort study. The STROBE checklist [10] and patient recruitment flowchart are included in the Supplementary Materials. Institutional review board approval was obtained for the study, waiving the requirement for individual informed consent. All analyses were carried out in accordance with pertinent laws and regulations and in conformity with the Helsinki Declaration. The study was approved by the Ethics Committee of the Technical University of Munich, School of Medicine (Protocol Number 180/17S; date of approval: 9 May 2017).

Participants were screened for eligibility based on a search of the hospital picture archiving system (PACS) for portal-venous-phase CT scans (70 s post injection of iodinated contrast media) including the pancreatic region from October 2006 to March 2019. A total of 237 candidates were confirmed eligible based on histologically ascertained pancreatic adenocarcinoma. Of these, 30 were excluded from the analysis due to insufficient technical quality of the CT scan (including motion artifacts and significant beam hardening due to nearby foreign materials), pre-existing malignant disease or any previous therapy, including chemotherapy, or loss to follow-up earlier than 2 weeks post-operatively.

The 207 resulting patients were separated into two subcohorts: cohort A, including the patients whose immunohistochemical assessment resulted in an unequivocal classification as QM or non-QM ($n = 181$), and cohort B, including the patients with histopathologically unclassifiable double positive (KRT81+/HNF1a+) tumors ($n = 26$), as described below.

The patients underwent computed tomography on the following CT scanner models: Siemens Somatom Definition ($n = 87$, 64-row, Siemens Healthineers, Erlangen, Germany), Philips iCT ($n = 79$, 256-row, Philips Healthcare, Best, The Netherlands), Philips IQON Spectral CT ($n = 41$, 64-row, Philips Healthcare, Best, The Netherlands). Clinical data collection and follow-up were handled by the departments of surgery and gastrointestinal oncology at our institution and ended on the 31st of March, 2019. The clinical variables collected were: Age, Sex, pTNM according to UICC 6th edition (pT: tumor stage, pN: nodal status, M: metastasis), R (resection margin), G (tumor grading), first-line adjuvant chemotherapy regimen, baseline CA19-9 (carbohydrate antigen 19-9), baseline CEA (carcinoembryonic antigen), tumor location (head/ body vs. tail) and overall survival.

Tumors were segmented independently under standardized conditions by two experts with 3- and 5-year experience in abdominal radiology, quality-controlled or corrected by a third expert with 8 years of experience in abdominal radiology and pancreatic imaging. After a period of two weeks, 20 randomly selected datasets from the three groups were sampled, randomly shuffled and presented to the same observers for re-segmentation. Segmentation was performed using the segmentation tool ITK-SNAP [11]. Radiomic features were extracted using PyRadiomics [12] using the settings detailed in the Supplementary Materials and normalized to the (0,1) interval. In total, 1474 radiomic features were extracted. Of these, features with missing values, all-null values, zero variance, features unstable to between-observer segmentation or to segmentation and re-segmentation (based on an intra-class correlation coefficient [13] below 0.9) were eliminated from the analysis. The remaining 161 features were normalized by tumor volume (calculated by PyRadiomics as mesh volume) as suggested by [14]. A Random Forest machine-learning algorithm [15] was used to model the features using the settings detailed in the Supplementary Materials with target labels of QM (KRT81+/HNF1a-) or non-QM (KRT81-/HNF1a- or KRT81-/HNF1a+). To alleviate class imbalance, per-sample weighting inversely proportional to the class population was applied. The classification performance with respect to the labels was assessed using sensitivity, specificity and ROC-AUC (area under the receiver operating characteristic curve) metrics using five-fold shuffle-split cross-validation with a test sample fraction of 0.3. Feature importance was assessed by reduction in Gini impurity [16] for each of the five folds and the average feature importance and standard deviation are reported. The algorithm achieving the highest ROC-AUC in cross-validation was applied to the cohort of unclassifiable PDAC, and the resulting predicted labels used as inputs for successive survival modelling. A technical evaluation of the study according to the recently published RSNA criteria [17] can be found in the Supplementary Materials.

To address bias associated with clinical covariates, cross-tabulations and multivariate Cox proportional hazards testing were performed. Univariate overall survival was modelled using the Kaplan Meier method including any censorship. The chi-squared-test was used for cross-tabulations, Students t-test for continuous variables and the log-rank-test for survival comparisons. A two-sided significance level of $p < 0.05$ was chosen.

Histopathological staining and immunohistochemical workup were performed by application of surrogate markers to determine the molecular subtype of PDAC based on the previously established immunohistochemical protocol described in [3]. Briefly, 2 μm sections were stained for KRT81 and HNF1a, and tumors were categorized into either one of three classes based on a cut-off value of 30% for tumor cell positivity/negativity: KRT81+/HNF1a- tumors were designated QM, KRT81-/HNF1a- and KRT81-/HNF1a+ tumors were grouped as non-QM. KRT81+/HNF1a+ tumors were designated double positive, i.e., unclassifiable. Classification was performed separately by two expert pathologists with 8 and 12 years of experience and quality-controlled or corrected by a third pathologist with 18 years of experience. Exemplary micro-photographs of the immunohistochemical stains alongside representative CT images can be found in Figure 1.

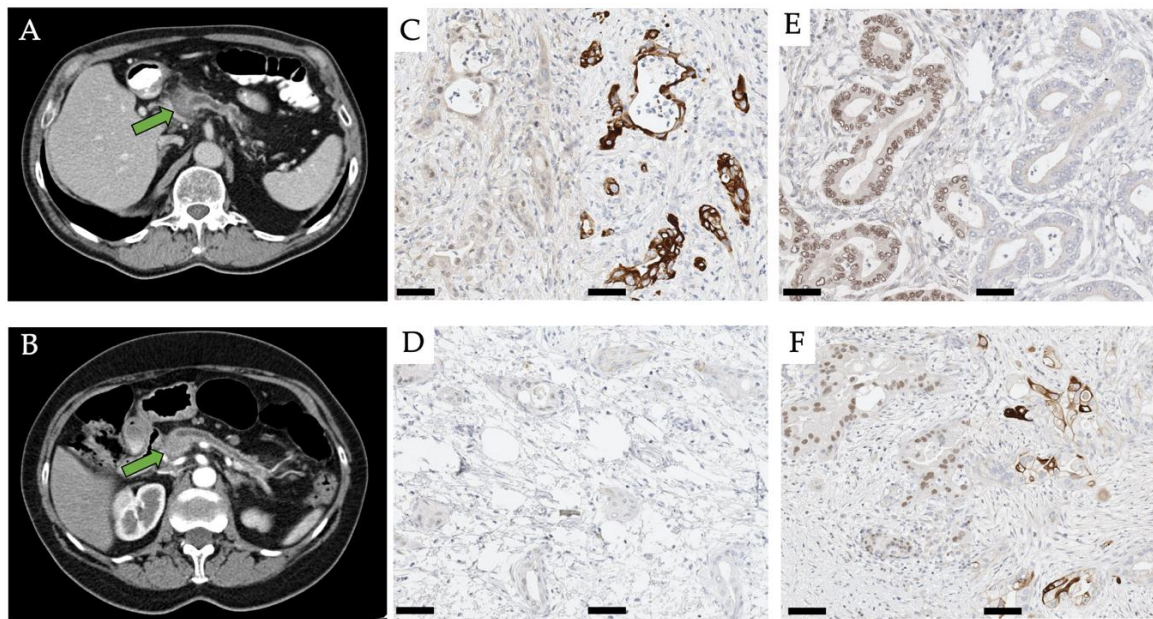


Figure 1. CT images of a patient with a QM (A) and a non-QM (B) PDAC in the pancreatic head (arrow). Window level 36 Hounsfield-Unit width 350 Hounsfield-Unit in both cases. Micro-photographs of representative immunohistochemical specimens of a HNF1a-/KRT81+ (QM) tumor (C), a HNF1a-/KRT81- (non-QM) tumor (D), a HNF1a+/KRT81- (non-QM) tumor (E) and a HNF1a+/KRT81+ (unclassifiable) tumor (F). Scale bar 50 μ m. HNF/KRT immunostainings left/right in each subfigure, respectively.

3. Results

In total, 207 patients were included in the study, 181 in the training/cross-validation cohort A and 26 in the testing cohort B. A schematic representation of the cohorts can be found in Figure 2.

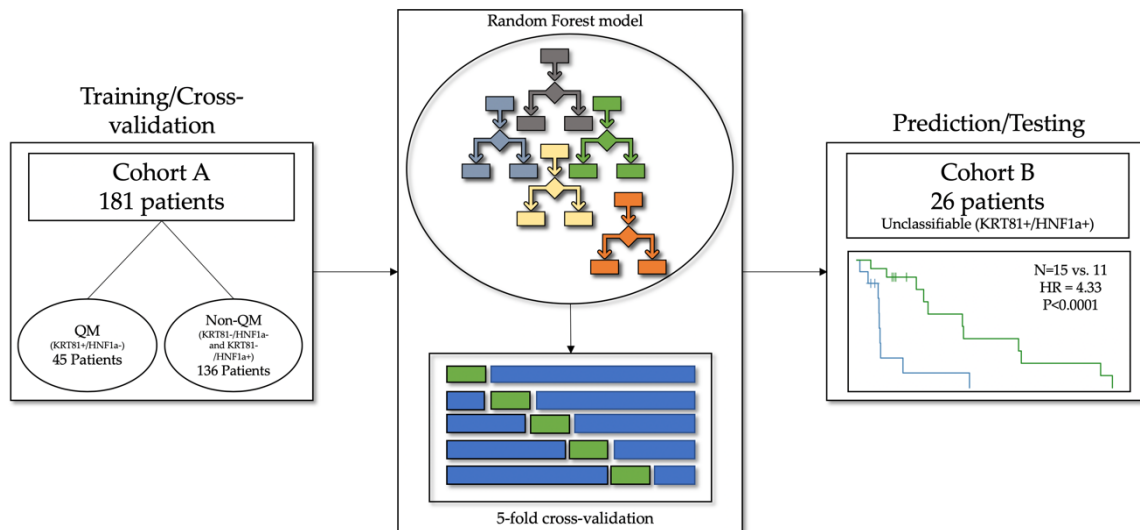


Figure 2. In total, 207 patients were included in the study. Among them, 181 patients in cohort A with confirmed QM and non-QM tumors served as the training and cross-validation data, and 26 patients in cohort B with unclassifiable tumors were used for model testing.

To assess relevant clinical covariates and confounders, clinical parameter evaluation and cross-tabulations were performed to assess the associated parameter distributions. These are found in Table 1.

Table 1. Clinical parameters and cross-tabulation results for the QM, non-QM and unclassifiable cohorts. Abbreviations: pT: tumor T-stage, pN: nodal status, M: metastasis, G: histopathological grading, R: resection margins (All UICC 6th ed.), CA19-9 and CEA: Carbohydrate Antigen 19-9 and Carcinoembryonic Antigen, N.A.: Not available. Statistical tests used: 1: Chi-Squared-Test, 2: one-way ANOVA, 3: Log-Rank-Test. N.S.: Not significant at the two-sided level of $p < 0.05$.

	Variable	QM (n = 45)	Non-QM (n = 136)	Unclassifiable (n = 26)	p-Value
Sex	Male	25 (55%)	75 (55%)	14 (54%)	0.99 ¹
	Female	20 (45%)	61 (45%)	12 (46%)	
Age (years)	Mean	68	67	72	0.84 ²
	Range	42–87	45–90	53–90	
pT	T1	1 (2%)	3 (2%)	1 (4%)	0.97 ¹
	T2	8 (18%)	19 (14%)	4 (15%)	
	T3	31 (69%)	103 (76%)	18 (69%)	
	T4	5 (11%)	11 (8%)	3 (12%)	
pN	N0	13 (29%)	31 (23%)	5 (19%)	0.60 ¹
	N1	32 (71%)	105 (77%)	21 (81%)	
M	M0	39 (87%)	125 (91%)	25 (96%)	0.36 ¹
	M1	6 (13%)	11 (8%)	1 (4%)	
G	G1	1 (2%)	6 (4%)	4 (15%)	0.11 ¹
	G2	21 (47%)	60 (45%)	13 (50%)	
	G3	23 (51%)	70 (51%)	9 (35%)	
R	R0	20 (44%)	68 (50%)	15 (58%)	0.56 ¹
	R1	25 (56%)	68 (50%)	11 (42%)	
CA19-9	Normal	5 (10%)	22 (16%)	2 (8%)	0.11 ¹
	Elevated	20 (45%)	43 (32%)	5 (19%)	
	N.A.	20 (45%)	71 (52%)	19 (73%)	
CEA	Normal	12 (27%)	38 (27%)	3 (11%)	0.08 ¹
	Elevated	8 (18%)	12 (9%)	1 (4%)	
	N.A.	25 (55%)	86 (64%)	22 (85%)	
First-Line Chemotherapy	Gemcitabine	16 (36%)	68 (50%)	12 (46%)	0.16 ¹
	FOLFIRINOX	3 (7%)	2 (2%)	2 (8%)	
	Did not receive	26 (57%)	66 (48%)	12 (46%)	
Median Overall Survival (months)		9.5	16.5	14.6	- QM vs. Non-QM: 0.03 ³ - Others N.S. ³
Censored	No	31 (69%)	97 (71%)	20 (77%)	0.77 ¹
	Yes	14 (31%)	39 (29%)	6 (23%)	
Tumor Location	Head/Body	44 (98%)	133 (98%)	25 (96%)	0.87 ¹
	Tail	1 (2%)	3 (2%)	1 (4%)	

No statistically significant difference was observed in the distribution of clinical confounding variables between the cohorts and none of the parameters were significantly associated with overall survival in multivariate survival modelling (Supplementary Figures S1 and S2).

In total, 1474 features were extracted from the CTs, of which 161 remained after feature engineering. The Random Forest algorithm was trained using the five-fold cross-validation approach detailed above on cohort A, and achieved a sensitivity of 0.84 ± 0.05 , a specificity of 0.92 ± 0.01 and a ROC-AUC of 0.93 ± 0.01 for the classification of QM vs. non-QM tumors across the five folds. The ROC curves for cohort A and the average ROC curve are shown in Figure 3.

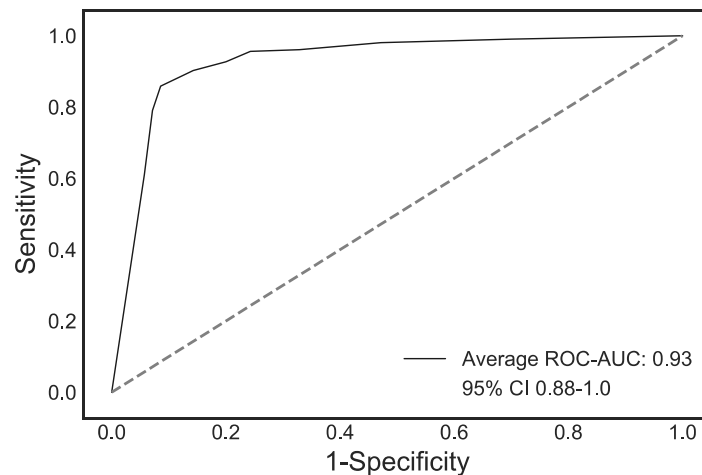


Figure 3. Average ROC curve (black) and 95% confidence interval (CI) (shaded blue) for the five cross-validation folds.

Feature importance assessment was performed using the *Gini* impurity index, measuring the quality of the split in each node of the random forest trees. The importance averages were averaged across folds, and the 20 features with the highest importance are presented in Supplementary Table S1 sorted in descending order. Among these features, several represent the spectrum of image homogeneity/heterogeneity, notably Entropy-/Energy-, Uniformity/Non-Uniformity and Correlation-/Variance-related features.

The trained Random Forest algorithm was applied to the previously unseen data of cohort B, consisting of histopathologically unclassifiable “double positive” tumors. The algorithm predicted a label of QM or non-QM for 12 and 14 out of 26 patients, respectively. Kaplan-Meier survival analysis of these cases resulted in a highly significant separation of the two predicted cohorts, with a median survival of 8.9 months for the predicted QM cases and of 39.8 months for the predicted non-QM cases (log-rank-test $p < 0.0001$, HR 4.33, 95% CI 1.41–13.32, Figure 4).

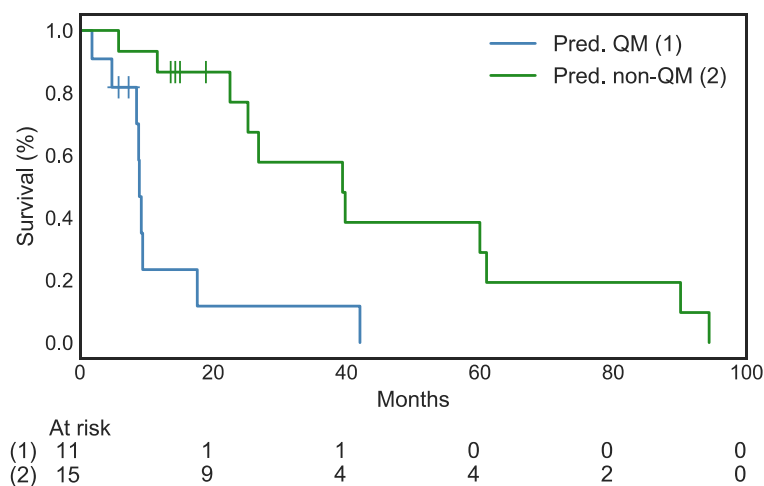


Figure 4. Highly significant separation of overall survival in the groups with predicted QM vs. non-QM tumors. HR: 4.33, 95% CI 1.14–13.32, log-rank test $p < 0.0001$. Vertical ticks indicate censorship.

4. Discussion

Here, we present a machine-learning algorithm capable of distinguishing between image-derived phenotypes representative of immunohistochemically defined molecular subtypes of PDAC.

The application of this algorithm to histopathologically unclassifiable tumors identifies two patient groups with significantly different overall survival. We therefore hypothesize that the algorithm is capable of re-identifying the dominant features of the QM and non-QM molecular subtypes in CT images and that radiomics-based phenotyping may thus offer a clinically usable classification advantageous over histopathology in the notoriously heterogenous entity of PDAC [18–20]. This notion is reinforced by the fact that histopathological samples are by default a significant underrepresentation of the tumor, since they are derived from a small sub-section of the tissue [21], and regions of differing molecular subtype are likely to coexist within the same tumor [22]. The exact nature of the so-called unclassifiable subtype might thus be presumed to represent a tumor simultaneously exposing a dual phenotype or a transitional phenotype on a continuum between full HNF1a and KRT81 positivity, mirrored in the intermediate survival outcome, both in our own and in previous studies [3]. A more fine-grained analysis would benefit from a global, i.e., whole tumor quantitative analysis of KRT81 and HNF1a expression, which is currently unfeasible due to the necessity of whole tumor work-up. Furthermore, such classification would result in small subgroup sizes, rendering the machine-learning analysis impossible in the current study.

The potential benefit of a radiomic workflow consequently lies in volumetric whole-tumor assessment, providing an opportunity to establish a clinically relevant phenotyping system and to better inform precision therapy regimens, and this concept of correlating quantitative morphometric evaluation with molecular phenotypes has been recently demonstrated for PDAC [23].

We previously presented a machine learning-algorithm capable of subtype characterization and—by extension—patient survival, based on pre-operative diffusion weighted MRI [8,9]. Extending this work, our current findings successfully transfer this methodology to routine CT acquisitions. The benefits of CT include broad availability, fewer motion artifacts and high standardization. We included CT images from several vendors, both to maximize sample size and to prompt vendor-invariance of the algorithm [17].

The process of Radiomics is dependent on source data standardization, pre-processing and can suffer from limited reproducibility, especially across modalities [24]. However, in our previous work and in other authors' work, heterogeneity-related features have been shown to be both prognostic and among the most reproducible [25], and it is reassuring to observe features belonging to this group, such as Entropy and Variance, re-surfacing in the current analysis. We therefore hypothesize that heterogeneity is a distinguishing feature of QM tumors and propose further investigation and development of specific heterogeneity imaging biomarkers.

Our study suffers from several limitations: the dataset in our study is unlikely to allow broad generalization due to its limited size, class imbalance and retrospective mono-institutional nature, necessitating cross-validation, which may encourage overfitting and yield overly optimistic classification metrics. This lack of multi-institutional prospective validation and datasets of sufficient statistical power is a notable issue with many radiomic studies [26]. Furthermore, any definition of PDAC subtypes may suffer from label noise, and it cannot be conclusively resolved at this point, how robustly the applied histopathological methodology can represent the transcriptome-based molecular phenotype [22]. Nevertheless, the highly significant separation of survival in the unclassifiable tumor subgroup observed in our study seems to support the binary classification of PDAC subtypes as recently proposed [27].

5. Conclusions

In conclusion, our study represents an iterative evolution of previously developed methods of radiomic phenotyping of PDAC and should be expanded and validated in larger retrospective and prospective study settings.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2077-0383/9/3/724/s1>: Biostatistics analyses, details on the radiomic extraction process, the machine learning modelling and the radiomic parameters, alongside technical evaluation of the study and the STROBE checklist.

Author Contributions: Conceptualization, G.A.K. and R.F.B.; Data curation, G.A.K., F.K.L., F.N.H., F.J. and D.S.; Formal analysis, G.A.K., S.Z., A.M., H.-Y.Y. and K.S.; Funding acquisition, G.A.K. and J.S.; Investigation, G.A.K., S.Z., A.M., H.-Y.Y., K.S. and R.F.B.; Methodology, G.A.K., F.K.L., F.J. and D.S.; Project administration, F.N.H., F.J., W.W. and M.R.M.; Resources, G.A.K., H.F., R.S., W.W., M.R.M. and R.F.B.; Software, G.A.K., S.Z. and F.J.; Supervision, H.F., R.S. and M.R.M.; Validation, A.M., K.S., W.W. and R.F.B.; Visualization, S.Z.; Writing—original draft, G.A.K., S.Z. and R.F.B.; Writing—review and editing, G.A.K., J.S., W.W. and R.F.B. All authors have read and agreed to the published version of the manuscript.

Funding: Rickmer Braren received funding from the German Research Foundation (DFG), SPP 2177/1, “Radiomics”, the German Research Foundation, SFB 824, Project C6 and from the German Cancer Consortium (DKTK), Joint Funding UPGRADE programme. Georgios Kaissis received funding from the Technical University Faculty of Medicine Clinician Scientist Programme. Jens Siveke received funding from the German Cancer Consortium (DKTK), the German Cancer Aid (grant no. 70112505; PIPAC consortium) and the German Research Foundation (DFG) through grant SI1549/3-1 (Clinical Research Unity KFO337) and the Collaborative Research Center SFB 824, project C4. Wilko Weichert received funding from the German Research Foundation, SFB 824, Project Z6.

Acknowledgments: Authors wish to thank Arzu Eckert for excellent technical support.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish this paper.

References

- Golan, T.; Hammel, P.; Reni, M.; Van Cutsem, E.; Macarulla, T.; Hall, M.J.; Park, J.-O.; Hochhauser, D.; Arnold, D.; Oh, D.-Y.; et al. Maintenance olaparib for germline brca-mutated metastatic pancreatic cancer. *N. Engl. J. Med.* **2019**, *381*, 317–327. [[CrossRef](#)] [[PubMed](#)]
- Collisson, E.A.; Bailey, P.; Chang, D.K.; Biankin, A.V. Molecular subtypes of pancreatic cancer. *Nat. Rev. Gastroenterol. Hepatol.* **2019**, *16*, 207–220. [[CrossRef](#)] [[PubMed](#)]
- Muckenhuber, A.; Berger, A.K.; Schlitter, A.M.; Steiger, K.; Konukiewitz, B.; Trumpp, A.; Eils, R.; Werner, J.; Friess, H.; Esposito, I.; et al. Pancreatic ductal adenocarcinoma subtyping using the biomarkers hepatocyte nuclear factor-1A and cytokeratin-81 correlates with outcome and treatment response. *Clin. Cancer Res.* **2018**, *24*, 351–359. [[CrossRef](#)] [[PubMed](#)]
- Cancer genome atlas research network. Electronic address aadhe, cancer genome atlas research N. integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell* **2017**, *32*, 185–203.e13. [[CrossRef](#)] [[PubMed](#)]
- Aung, K.L.; Fischer, S.E.; Denroche, R.E.; Jang, G.H.; Dodd, A.; Creighton, S.; Southwood, B.; Liang, S.B.; Chadwick, D.; Zhang, A.; et al. Genomics-Driven Precision Medicine for Advanced Pancreatic Cancer: Early Results from the COMPASS Trial. *Clin. Cancer Res.* **2018**, *24*, 1344–1354. [[CrossRef](#)] [[PubMed](#)]
- Zhou, M.; Leung, A.; Echegaray, S.; Gentles, A.; Shrager, J.B.; Jensen, K.C.; Berry, G.J.; Plevritis, S.K.; Rubin, D.L.; Napel, S. Non-small cell lung cancer radiogenomics map identifies relationships between molecular and imaging phenotypes with prognostic implications. *Radiology* **2018**, *286*, 307–315. [[CrossRef](#)]
- Rathore, S.; Akbari, H.; Rozycki, M.; Abdullah, K.G.; Nasrallah, M.P.; Binder, Z.A.; Davuluri, R.V.; Lustig, R.A.; Dahmane, N.; Bilello, M.; et al. Radiomic MRI signature reveals three distinct subtypes of glioblastoma with different clinical and molecular characteristics, offering prognostic value beyond IDH1. *Sci. Rep.* **2018**, *8*, 5087. [[CrossRef](#)]
- Kaissis, G.; Ziegelmayer, S.; Lohöfer, F.; Algül, H.; Eiber, M.; Weichert, W.; Schmid, R.; Friess, H.; Rummeny, E.; Ankerst, D.; et al. A machine learning model for the prediction of survival and tumor subtype in pancreatic ductal adenocarcinoma from preoperative diffusion-weighted imaging. *Eur. Radiol. Exp.* **2019**, *3*, 41. [[CrossRef](#)]
- Kaissis, G.; Ziegelmayer, S.; Lohöfer, F.; Steiger, K.; Algül, H.; Muckenhuber, A.; Yen, H.-Y.; Rummeny, E.; Friess, H.; Schmid, R.; et al. A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response to gemcitabine-based versus folfirinix chemotherapy. *PLoS ONE* **2019**, *14*, e0218642. [[CrossRef](#)]
- Von Elm, E.; Altman, D.G.; Egger, M.; Pocock, S.J.; Gøtzsche, P.C.; Vandenbroucke, J.P.; Initiative, S. The strengthening the reporting of observational studies in epidemiology (strobe) statement: Guidelines for reporting observational studies. *Int. J. Surg.* **2014**, *12*, 1495–1499. [[CrossRef](#)]

11. Yushkevich, P.A.; Piven, J.; Hazlett, H.C.; Smith, R.G.; Ho, S.; Gee, J.C.; Gerig, G. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* **2006**, *31*, 1116–1128. [[CrossRef](#)]
12. Van Griethuysen, J.J.M.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.H.; Fillion-Robin, J.-C.; Pieper, S.; Aerts, H.J.W.L. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)]
13. McGraw, K.O.; Wong, S.P. Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* **1996**, *1*, 30. [[CrossRef](#)]
14. Welch, M.L.; McIntosh, C.; Haibe-Kains, B.; Milosevic, M.F.; Wee, L.; Dekker, A.; Huang, S.H.; Purdie, T.G.; O’Sullivan, B.; Aerts, H.J.W.L.; et al. Vulnerabilities of radiomic signature development: The need for safeguards. *Radiother. Oncol.* **2019**, *130*, 2–9. [[CrossRef](#)]
15. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
16. Menze, B.H.; Kelm, B.M.; Masuch, R.; Himmelreich, U.; Bachert, P.; Petrich, W.; Hamprecht, F.A. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinform.* **2009**, *10*, 213. [[CrossRef](#)]
17. Bluemke, D.A.; Moy, L.; Bredella, M.A.; Ertl-Wagner, B.B.; Fowler, K.J.; Goh, V.J.; Halpern, E.F.; Hess, C.P.; Schiebler, M.L.; Weiss, C.R. Assessing radiology research on artificial intelligence: A brief guide for authors, reviewers, and readers—From the radiology editorial board. *Radiology* **2019**, *294*, 3. [[CrossRef](#)]
18. Peng, J.; Sun, B.F.; Chen, C.Y.; Zhou, J.Y.; Chen, Y.S.; Chen, H.; Liu, L.; Huang, D.; Jiang, J.; Cui, G.S.; et al. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res.* **2019**, *29*, 725–738. [[CrossRef](#)]
19. Orth, M.; Metzger, P.; Gerum, S.; Mayerle, J.; Schneider, G.; Belka, C.; Schnurr, M.; Lauber, K. Pancreatic ductal adenocarcinoma: Biological hallmarks, current status, and future perspectives of combined modality treatment approaches. *Radiat. Oncol.* **2019**, *14*, 141. [[CrossRef](#)]
20. Juiz, N.A.; Iovanna, J.; Dusetti, N. Pancreatic cancer heterogeneity can be explained beyond the genome. *Front. Oncol.* **2019**, *9*, 246. [[CrossRef](#)]
21. Rahemtullah, A.; Misdraji, J.; Pitman, M.B. Adenosquamous carcinoma of the pancreas: Cytologic features in 14 cases. *Cancer Cytopathol. Interdiscip. Int. J. Am. Cancer Soc.* **2003**, *99*, 372–378. [[CrossRef](#)]
22. Chan-Seng-Yue, M.; Kim, J.C.; Wilson, G.W.; Ng, K.; Figueroa, E.F.; O’Kane, G.M.; Connor, A.A.; Denroche, R.E.; Grant, R.C.; McLeod, J.; et al. Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. *Nat. Genet.* **2020**, *52*, 231–240. [[CrossRef](#)]
23. N Kalimuthu, S.; Wilson, G.W.; Grant, R.C.; Seto, M.; O’Kane, G.; Vajpeyi, R.; Notta, F.; Gallinger, S.; Chetty, R. Morphological classification of pancreatic ductal adenocarcinoma that predicts molecular subtypes and correlates with clinical outcome. *Gut* **2020**, *69*, 317–328. [[CrossRef](#)]
24. Van Timmeren, J.E.; Carvalho, S.; Leijenaar, R.T.H.; Troost, E.G.C.; van Elmpt, W.; de Ruyscher, D.; Muratet, J.-P.; Denis, F.; Schimek-Jasch, T.; Nestle, U.; et al. Challenges and caveats of a multi-center retrospective radiomics study: An example of early treatment response assessment for NSCLC patients using FDG-PET/CT radiomics. *PLoS ONE* **2019**, *14*, e0217536. [[CrossRef](#)]
25. Khalvati, F.; Zhang, Y.; Baig, S.; Lobo-Mueller, E.M.; Karanicolas, P.; Gallinger, S.; Haider, M.A. Prognostic value of CT radiomic features in resectable pancreatic ductal adenocarcinoma. *Sci. Rep.* **2019**, *9*, 5449. [[CrossRef](#)]
26. Kumar, V.; Gu, Y.; Basu, S.; Berglund, A.; Eschrich, S.A.; Schabath, M.B.; Forster, K.; Aerts, H.J.W.L.; Dekker, A.; Fenstermacher, D.; et al. Radiomics: The process and the challenges. *Magn. Reson. Imaging* **2012**, *30*, 1234–1248. [[CrossRef](#)]
27. Rashid, N.U.; Peng, X.L.; Jin, C.; Moffitt, R.A.; Volmar, K.E.; Belt, B.A.; Panni, R.Z.; Nywening, T.M.; Herrera, S.G.; Moore, K.J.; et al. Purity independent subtyping of tumors (PurIST), A clinically robust, single-sample classifier for tumor subtyping in pancreatic cancer. *Clin. Cancer Res.* **2020**, *26*, 82–92. [[CrossRef](#)]

