

TECHNISCHE UNIVERSITÄT MÜNCHEN

FAKULTÄT FÜR MEDIZIN

Klinik und Poliklinik für RadioOnkologie und Strahlentherapie
Abteilung für Experimentelle Radioonkologie und Strahlenbiologie
Klinikum rechts der Isar München

Einfluss von Bestrahlung und operativer Resektion des
Primärtumors auf extramedulläre Hämatopoese in der Leber
anhand eines orthotopen 4T1 Mammakarzinom Mausmodells.

IRIS NOEMI SCHWORM



TECHNISCHE UNIVERSITÄT MÜNCHEN

FAKULTÄT FÜR MEDIZIN

Klinik und Poliklinik für RadioOnkologie und Strahlentherapie
Abteilung für Experimentelle Radioonkologie und Strahlenbiologie
Klinikum rechts der Isar München

Einfluss von Bestrahlung und operativer Resektion des
Primärtumors auf extramedulläre Hämatopoese in der Leber
anhand eines orthotopen 4T1 Mammakarzinom Mausmodells.

IRIS NOEMI SCHWORM

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

Vorsitzender:

Prof. Dr. Ernst J. Rummeny

Prüfende der Dissertation:

1. Prof. Dr. Gabriele Multhoff

2. Prof. Dr. Stephanie E. Combs

Die Dissertation wurde am 21.11.2016 bei der Technischen Universität München eingereicht
und durch die Fakultät für Medizin am 12.07.2017 angenommen.

Inhaltsverzeichnis

Tabellenverzeichnis	iv
Abbildungsverzeichnis	vi
Abkürzungsverzeichnis	ix
1 Zielsetzung	1
2 Einführung	2
2.1 Grundlagen der Onkologie	2
2.1.1 Histologie von malignen Tumoren und Tumorzellen	2
2.1.2 Eigenschaften von Tumoren und Tumorzellen	3
2.1.3 Tumorstammzellen und Zellen des Tumor-Microenvironments	10
2.1.4 Individualisierte Tumorthherapie	11
2.2 Hämatopoese und extramedulläre Hämatopoese	12
2.3 Das Immunsystem und seine Rolle in der Onkologie	14
2.3.1 Immunsystem und Immunzellen – Funktion, Organisation und histologische Charakterisierung	14
2.3.2 Die Rolle des Immunsystems in der Onkologie	17
2.4 Das 4T1 Mammakarzinom Mausmodell	21
2.5 Grundlagen der Strahlentherapie in der Onkologie	21
2.6 Die Leber	23
2.7 Das Mammakarzinom	24
2.7.1 Epidemiologie des Mammakarzinoms	24
2.7.2 Histologie des Mammakarzinoms	24
2.7.3 Diagnostik und Therapie des Mammakarzinoms	25
2.7.4 Prognose des Mammakarzinoms	28
3 Material und Methoden	29
3.1 Material	29

3.1.1	Verbrauchsmaterialien	29
3.1.2	Geräte und Software	30
3.1.3	Chemikalien und Reagenzien, Lösungen	30
3.1.4	Antikörper.....	31
3.2	Methoden.....	32
3.2.1	Mausexperiment	32
3.2.2	Einbetten der Tumore, Lebern und Lungen	32
3.2.3	Anfertigen von histologischen Schnitten	32
3.2.4	Hämatoxylin-Eosin Färbung.....	33
3.2.5	Immunhistochemie.....	34
3.2.6	Mikroskopieren der gefärbten Schnitte	37
3.2.7	Einscannen der gefärbten Schnitte mit Aperio ScanScope	37
3.2.8	Auswerten der Scans der H.E. gefärbten Schnitte mit ImageScope	38
3.2.9	Auswerten der Daten aus ImageScope in Excel	39
3.2.10	Auswerten der Scans der immunhistochemisch gefärbten Schnitte mit Definiens.....	41
3.2.11	Auswerten der Daten aus Definiens in Excel.....	43
3.2.12	Statistische und graphische Auswertung der Daten in SPSS	43
4	Ergebnisse.....	45
4.1	Scans der H.E. gefärbten Leberschnitte	45
4.2	Scans der H.E. gefärbten Tumorschnitte.....	46
4.3	Scans der immunhistochemisch gefärbten Leberschnitte	47
4.4	Scans der immunhistochemisch gefärbten Tumorschnitte.....	50
4.5	Verteilung der Primärtumorumfängen	53
4.6	ImageScope – EMH Parameter und statistische Tests	55
4.7	Definiens – EMH Parameter und statistische Tests	60
4.8	Vergleich der Verfahren – ImageScope und Definiens	64
4.9	ImageScope – Korrelation der EMH Parameter mit dem Tumorumfang	70

4.10	Definiens – Korrelation der EMH Parameter mit dem Tumorzvolumen	76
5	Diskussion	82
5.1	Scans der H.E. gefärbten Leberschnitte	82
5.2	Scans der immunhistochemisch gefärbten Leberschnitte	82
5.3	Scans der H.E. und immunhistochemisch gefärbten Tumorschnitte	83
5.4	Vergleich des Ausmaßes an EMH in den verschiedenen Gruppen	84
5.5	Vergleich der angewandten Verfahren ImageScope und Definiens	85
5.6	Zusammenhang zwischen EMH, Therapieform und Tumorzvolumen.....	87
5.7	Ausblick.....	88
6	Zusammenfassung.....	88
	Literaturverzeichnis.....	90
	Danksagung	96
	Anhang	97
1.	Peter, L., Mateus, D., Chatelain, P., Declara, D., Schworm, N., Stangl, S., Multhoff, G., Navab, N. - Assisting the Examination of Large Histopathological Slides with Adaptive Forests - <i>Medical Image Analysis, 2017</i>	97
2.	Peter, L., Mateus, D., Chatelain, P., Schworm, N., Stangl, S., Multhoff, G., Navab, N. - Leveraging Random Forests for Interactive Exploration of Large Histological Images - <i>Proceedings of the 17th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2014</i>	97

Tabellenverzeichnis

Tabelle 1: Verbrauchsmaterialien	29
Tabelle 2: Geräte und Software	30
Tabelle 3: Chemikalien und Reagenzien.....	30
Tabelle 4: Lösungen.....	31
Tabelle 5: Primärantikörper	31
Tabelle 6: Sekundärantikörper	32
Tabelle 7: Benennung der Mäuse.....	32
Tabelle 8: Protokoll Hämatoxylin-Eosin Färbung	34
Tabelle 9: Protokoll Immunhistochemie	36
Tabelle 10: Primärantikörper mit Verdünnung und Funktion.....	37
Tabelle 11: Sekundärantikörper mit Verdünnung.....	37
Tabelle 12: Ermittelte Daten aus den H.E. gefärbten Leberscans aus ImageScope.....	40
Tabelle 13: Einteilung der Läsionen nach Größe.....	40
Tabelle 14: Berechnung der Parameter aus den H.E. gefärbten Leberscans aus ImageScope.....	40
Tabelle 15: Berechnung der Parameter aus den IHC gefärbten Leberscans aus Definiens	43
Tabelle 16: Kruskal-Wallis-Test – Tumervolumina.....	54
Tabelle 17: Mann-Whitney-U-Test – Tumervolumina	54
Tabelle 18: Kruskal-Wallis-Test – EMH, ImageScope	58
Tabelle 19: Mann-Whitney-U-Test – EMH, ImageScope.....	59
Tabelle 20: Gegenüberstellung Ellipse vs. Rechteck – ImageScope.....	59
Tabelle 21: Kruskal-Wallis-Test – EMH, Definiens.....	63
Tabelle 22: Mann-Whitney-U-Test – EMH, Definiens	64
Tabelle 23: Korrelationskoeffizienten – Läsionen pro Fläche, ImageScope.....	71
Tabelle 24: Korrelationskoeffizienten – Kleine Läsionen pro Fläche, ImageScope	72
Tabelle 25: Korrelationskoeffizienten – Mittlere Läsionen pro Fläche, ImageScope.....	73
Tabelle 26: Korrelationskoeffizienten – Große Läsionen pro Fläche, ImageScope.....	74

Tabelle 27: Korrelationskoeffizienten – Läsionsfläche pro Gesamtfläche, ImageScope.....	75
Tabelle 28: Korrelationskoeffizienten – Läsionen pro Fläche, Definiens	77
Tabelle 29: Korrelationskoeffizienten – Kleine Läsionen pro Fläche, Definiens	78
Tabelle 30: Korrelationskoeffizienten – Mittlere Läsionen pro Fläche, Definiens	79
Tabelle 31: Korrelationskoeffizienten – Große Läsionen pro Fläche, Definiens	80
Tabelle 32: Korrelationskoeffizienten – Läsionsfläche pro Gesamtfläche, Definiens	81
Tabelle 33: Vergleich von Definiens und ImageScope – Parameter aus Bland-Altman-Plot	85

Abbildungsverzeichnis

Abbildung 1: Stammbaum der Blutzellen.....	13
Abbildung 2: ImageScope: Erfassen aller Läsionen in den ausgewählten Bereichen.	38
Abbildung 3: ImageScope: Erfassen aller großen Läsionen pro Leberschnitt.	39
Abbildung 4: ImageScope, Ergebniskontrolle.	39
Abbildung 5: Definiens – Beispiel aus Snapshots, Original.	42
Abbildung 6: Definiens – Beispiel aus Snapshots, Overlay.....	42
Abbildung 7: Leberscan H.E. – Läsion.....	45
Abbildung 8: Tumorscan H.E. – Zellen.....	46
Abbildung 9: Tumorscan H.E. – Nekrose.....	46
Abbildung 10: Leberscan IHC – stark CD45 positiv.....	47
Abbildung 11: Leberscan IHC – CD45 positiv.....	47
Abbildung 12: Leberscan IHC – stark Ly6G/ Ly6C positiv.	48
Abbildung 13: Leberscan IHC – Ly6G/ Ly6C positive und negative Läsionen.....	48
Abbildung 14: Leberscan IHC – Ly6G/ Ly6C negative Läsion.....	49
Abbildung 15: Leberscan IHC – F4/80 positive Zellen.	49
Abbildung 16: Leberscan IHC – F4/80 positive Zellen in einer Läsion.....	49
Abbildung 17: Leberscan IHC – CD49b positive Zellen in einer Läsion.	50
Abbildung 18: Leberscan IHC – CD3 positive Zellen in einer Läsion.....	50
Abbildung 19: Tumorscan IHC – CD45 positive Zellen.	51
Abbildung 20: Tumorscan IHC – CD45 positive Zellen in nekrotischem Bereich.	51
Abbildung 21: Tumorscan IHC – einzelne Ly6G/ Ly6C positive Zellen.	51
Abbildung 22: Tumorscan IHC – zahlreiche Ly6G/ Ly6C positive Zellen.	52
Abbildung 23: Tumorscan IHC – F4/80 positive Zellen.....	52
Abbildung 24: Tumorscan IHC – CD49b positive Zellen.	52
Abbildung 25: Tumorscan IHC – CD3 positive Zellen.	52
Abbildung 26: Tumorscan IHC – CD3 positive Zellen.	53

Abbildung 27: Tumolvolumina.....	54
Abbildung 28: ImageScope – Läsionen pro Fläche.....	56
Abbildung 29: ImageScope – Kleine Läsionen pro Fläche.....	56
Abbildung 30: ImageScope – Mittlere Läsionen pro Fläche.....	57
Abbildung 31: ImageScope – Große Läsionen pro Fläche.....	57
Abbildung 32: ImageScope – Läsionsfläche pro Gesamtfläche.....	58
Abbildung 33: Bland-Altman-Plot: Ellipse vs. Rechteck.	59
Abbildung 34: Definiens – Läsionen pro Fläche.	61
Abbildung 35: Definiens – Kleine Läsionen pro Fläche.	61
Abbildung 36: Definiens – Mittlere Läsionen pro Fläche.	62
Abbildung 37: Definiens – Große Läsionen pro Fläche.	62
Abbildung 38: Definiens – Läsionsfläche pro Gesamtfläche.	63
Abbildung 39: ImageScope und Definiens - Läsionen pro Fläche.	65
Abbildung 40: ImageScope und Definiens – Kleine Läsionen pro Fläche.....	65
Abbildung 41: ImageScope und Definiens – Mittlere Läsionen pro Fläche.	66
Abbildung 42: ImageScope und Definiens – Große Läsionen pro Fläche.....	66
Abbildung 43: ImageScope und Definiens – Läsionsfläche pro Gesamtfläche	67
Abbildung 44: Bland-Altman-Plot Definiens vs. ImageScope – Läsionen pro Fläche.....	67
Abbildung 45: Bland-Altman-Plot Definiens vs. ImageScope – Kleine Läsionen pro Fläche.....	68
Abbildung 46: Bland-Altman-Plot Definiens vs. ImageScope – Mittlere Läsionen pro Fläche.....	68
Abbildung 47: Bland-Altman-Plot Definiens vs. ImageScope – Große Läsionen pro Fläche.....	69
Abbildung 48: Bland-Altman-Plot Definiens vs. ImageScope – Läsionsfläche pro Gesamtfläche.....	69
Abbildung 49: ImageScope – Korrelation Läsionen pro Fläche und Tumolvolumen.....	71
Abbildung 50: ImageScope – Korrelation kleine Läsionen pro Fläche und Tumolvolumen.	72
Abbildung 51: ImageScope – Korrelation mittlere Läsionen pro Fläche und Tumolvolumen.....	73
Abbildung 52: ImageScope – Korrelation große Läsionen pro Fläche und Tumolvolumen.	74
Abbildung 53: ImageScope – Korrelation Läsionsfläche pro Gesamtfläche und Tumolvolumen.....	75

Abbildung 54: Definiens – Korrelation Läsionen pro Fläche und Tumorvolumen.	77
Abbildung 55: Definiens – Korrelation kleine Läsionen pro Fläche und Tumorvolumen.....	78
Abbildung 56: Definiens – Korrelation mittlere Läsionen pro Fläche und Tumorvolumen.	79
Abbildung 57: Definiens – Korrelation große Läsionen pro Fläche und Tumorvolumen.....	80
Abbildung 58: Definiens – Korrelation Läsionsfläche pro Gesamtfläche und Tumorvolumen.	81

Abkürzungsverzeichnis

AGM	Aorta, Gonaden, Mesonephros
Ang	Angiopoetin
APC	Antigen-presenting cell
BET	Brusttherhaltende Therapie
BRCA1/2	Breast cancer 1/2
BSA	Bovines Serum Albumin
CDK	Cyclin-dependent kinase
CIS	Carcinoma in situ
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
COX	Cyclooxygenase
CSC	Cancer stem cell
DAB	3, 3' Diaminobenzidin
DC	Dendritic cell
DCIS	Ductal carcinoma in situ
dHSC	Definitive hematopoietic stem cell
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMH	Extramedulläre Hämatopoese
EMT	Epithelial–mesenchymal transition
ER	Estrogen- Rezeptor
FFPE	Formalin-fixed paraffin embedded
FGF	Fibroblast growth factor
Foxp3	Forkhead box P3
G-CSF	Granulocyte colony-stimulating factor
GM-CSF	Granulocyte-macrophage colony stimulating factor
GMP	Granulocyte/macrophage progenitor
H.E.	Hämatoxylin-Eosin
Her2	Human epidermal growth factor receptor 2
HSC	Hematopoietic stem cell
IFN	Interferon
Ig	Immunglobulin
IHC	Immunhistochemie
IL	Interleukin
KM	Knochenmark
LCIS	Lobular carcinoma in situ
LPS	Lipopolysaccharide
LT-HSC	Long-term hematopoietic stem cells
MCSF	Macrophage colony-stimulating factor
MDSC	Myeloid-derived suppressor cell
MHC	Major histocompatibility complex
MEP	Megakaryocyte/erythrocyte progenitor
MET	Mesenchymal-Epitheliale Transition
MPP	Multipotent progenitor
MRT	Magnetresonanztomographie
NK - Zelle	Natürliche Killerzelle

PAMP	Pathogen-associated molecular pattern
PBS	Phosphat buffered saline
PET	Positronen-Emissions-Tomographie
PR	Progesteron- Rezeptor
pRb	Retinoblastoma protein
RNA	Ribonucleic acid
SLNB	Sentinel-Lymph-Node-Biopsie
SMA	Smooth muscle actin
TAM	Tumor assoziierter Makrophage
TGF	Transforming growth factor
TNF	Tumor necrosis factor
Treg	Regulatorische T-Zelle
TSP	Thrombospondin
VEGF	Vascular endothelial growth factor
ZNS	Zentrales Nervensystem

1 Zielsetzung

Das Mammakarzinom ist die häufigste Krebserkrankung der Frau und nimmt deshalb einen großen Stellenwert in unserer Gesellschaft ein (Robert-Koch-Institut, 2015). Mit einer relativen 5-Jahres Überlebensrate von 88% im Jahre 2011 (Robert-Koch-Institut, 2015) ist die Prognose für Brustkrebspatientinnen vergleichsweise gut. Dennoch kommt es bei vielen Patientinnen in Laufe der Krankheit zur Ausbildung von Metastasen, die nicht kurativ behandelbar sind (Harbeck and Heywang-Köbrunner, 2011, Sauer, Mammakarzinom, 2010). Die Mechanismen dahinter sind noch immer unzureichend geklärt und Gegenstand aktueller Forschung. Dass das Immunsystem bei der Metastasierung sowie bei der Entstehung und der Progression von Tumoren eine wichtige Rolle spielt, gilt als gesichert (Grivennikov et al., 2010, Hanahan and Weinberg, 2011).

Diese Arbeit beschäftigt sich mit dem komplexen und noch immer nur teilweise entschlüsselten Zusammenspiel zwischen Immunsystem und Tumor. Ziel dieser medizinischen Doktorarbeit war die Charakterisierung von Extramedullärer Hämatopoese (EMH) in der Leber anhand eines orthotopen spontan metastasierenden 4T1 Mammakarzinom Mausmodells. EMH gilt als Quelle myeloischer Immunzellen, sogenannter „myeloid-derived suppressor cells“ (MDSCs) (Ilkovitch and Lopez, 2009), die zur Tumorprogression beitragen (Youn et al., 2008).

Zunächst wurden Tumorzellen in die Brustdrüse von 18 BALB/c Mäusen implantiert, die nach zwei Wochen in drei Gruppen von je sechs Mäusen unterteilt wurden: Die erste Gruppe erhielt keine Behandlung, in der zweiten Gruppe wurde der Primärtumor nach zwei Wochen Wachstumsphase bestrahlt und in der dritten Gruppe wurde der Primärtumor nach zwei Wochen Wachstumsphase operativ entfernt. Nach weiteren zwei Wochen wurden alle Mäuse getötet.

Sowohl von den Tumoren als auch von den Lebern, welche mir zur Verfügung gestellt wurden, habe ich histologische Schnitte erstellt. Diese habe ich mit Hämatoxylin-Eosin (H.E.) sowie immunhistochemisch mit verschiedenen Antikörpern gefärbt und anschließend eingescannt. Die Immunhistochemie (IHC) diente der sicheren Identifizierung und weiteren Klassifizierung der bereits in der Hämatoxylin-Eosin-Färbung sichtbaren EMH in der Leber. Das Ausmaß der EMH nach etablierten Therapieregimen des Mammakarzinoms wie Bestrahlung und operativer Resektion des Primärtumors mit einer Kontrolle wurde verglichen. Dazu wurde die EMH anhand der H.E. und IHC gefärbten Scans mit Aperio ImageScope und Definiens quantifiziert und mit dem Primärtumorvolumen des Tumors der jeweiligen Maus in Korrelation gesetzt. Ein weiteres Verfahren zur Quantifizierung von EMH anhand meiner H.E. gefärbter Scans wurde entwickelt (Peter et al., 2017, Peter et al., 2014).

2 Einführung

Zunächst werden Grundlagen der Onkologie erläutert. Es wird auf für diese Arbeit relevante Aspekte über die Zusammensetzung und Funktion des Blutes und der Blutbildung eingegangen. Funktion, Organisation und histologische Charakterisierung des Immunsystems und seine Rolle in der Onkologie werden ausführlich besprochen. Zudem wird Hintergrundwissen zu dem verwendeten Mausmodell vermittelt. Grundlagen der Strahlentherapie sowie Aufbau und Funktion der Leber werden skizziert. Zum Schluss werden Epidemiologie, Diagnostik, Therapie und Prognose des humanen Mammakarzinoms thematisiert.

2.1 Grundlagen der Onkologie

2.1.1 Histologie von malignen Tumoren und Tumorzellen

Ich habe in dieser Arbeit histologische Schnitte sowohl von Tumoren als auch von Lebern angefertigt und diese mittels H.E. und IHC gefärbt. Deshalb werden im Folgenden histologische Eigenschaften maligner Tumore - von Sauer (Tumorpathologie, 2010) übernommen - dargestellt:

- Gewebe, Zellen und Zellkerne sind vielgestaltig (polymorph), die Zellkerne häufig pyknotisch (verklumpt). Sie weichen vom Normalbild ab.
- Das Kern-Plasma-Verhältnis ist zugunsten des Kerns verschoben.
- Das Zellplasma ist bei der üblichen Hämatoxylin-Eosin-Färbung basophil (blau) durch einen hohen RNA-Gehalt.
- Die Kerne sind groß, verstärkt gefärbt und weisen mehrere Nukleolen (Kernkörperchen) auf.
- Mitosen (sichtbare Kernteilungen) sind häufiger als im Normalgewebe. Pathologische (abnormale) Mitosen fallen auf.
- Der Tumor respektiert die Grenze zum umgebenden gesunden Gewebe nicht. Er wächst in dieses hinein und zerstört es (infiltratives und destruktives Wachstum, „wurzelförmige“ Krebsausläufer).

Wie stark sich Tumore vom ursprünglichen Gewebe differenzieren wird im Grading wiedergegeben (G1 bis G4), wobei G1 Tumore gut differenziert und G4 Tumore entdifferenziert sind (Sauer, Tumorpathologie, 2010). Entdifferenzierte Tumore sind meist aggressiv und schlecht operabel, aber sensibel gegenüber Chemo- und Strahlentherapie, die vor allem Zellen mit einer hohen Teilungsrate schädigen (Sauer, Tumorpathologie, 2010).

2.1.2 Eigenschaften von Tumoren und Tumorzellen

Tumore und Tumorzellen unterscheiden sich nicht nur histologisch, sondern auch in ihrer Physiologie, ihrem Stoffwechsel und ihrer Interaktion mit dem umgebenden Gewebe von gesunden Zellen (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Da diese Eigenschaften und vor allem die Interaktion von Immunsystem und Tumor ein zentrales Thema dieser Arbeit sind, soll im Folgenden ein Überblick über die Tumorbilogie vermittelt werden. Hanahan und Weinberg verfassten 2000 und 2011 zwei Reviews, die die wichtigsten Forschungsergebnisse auf dem Gebiet der Onkologie der letzten Jahrzehnte zusammenfassten und die überaus komplexen Zusammenhänge der Tumorbilogie in eine verständliche und logische Ordnung brachten (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Sie beschrieben 8 „hallmarks of cancer“ als essentielle Veränderungen in der Physiologie von Tumorzellen fast aller Tumorentitäten. Die Ursache dafür, dass Tumore dennoch relativ selten sind, sehen die Autoren in den strengen und vielfältigen Schutzmechanismen der Zelle.

Die „hallmarks of cancer“ nach Hanahan and Weinberg (2011):

- „Sustaining proliferative signaling“ bedeutet die Aufrechterhaltung von Signalkaskaden, die eine fortwährende Proliferation ermöglichen
- „Evading growth suppressors“ bedeutet das Umgehen von wachstumshemmenden Signalwegen
- „Resisting cell death“ bedeutet die Resistenz gegenüber Zelltod-Mechanismen
- „Enabling replicative immortality“ bedeutet unbegrenztes Zellwachstum
- „Inducing angiogenesis“ bedeutet die Induktion von Tumorangiogenese
- „Activating invasion and metastasis“ bedeutet die Aktivierung von invasivem Tumorwachstum und Metastasierung
- „Reprogramming of energy metabolism“ bedeutet das Umprogrammieren des Energiemetabolismus von Tumorzellen
- „Evading Immune destruction“ bedeutet die Inhibition der immunologischen Abwehr gegenüber Tumorzellen

Hanahan und Weinberg beschrieben Krebs als genetische Erkrankung mit diversen Veränderungen im Genom der Zellen. Dabei nahmen sie an, dass jeder genetisch erworbene Wachstumsvorteil, den Gesetzen der Evolution folgend, weiter zur Tumorentstehung beiträgt. Sie unterschieden dominant vererbte, tumorfördernde Onkogene und rezessiv vererbte Tumorsuppressorgene, die bei der Tumorentstehung inaktiviert werden. Die beschriebenen „hallmarks of cancer“ werden durch genetische Instabilität und eine tumorfördernde Wirkung des Immunsystems ermöglicht (Hanahan and Weinberg, 2011).

2.1.2.1 Sustaining proliferative signaling

Im physiologischen Zustand ist die Proliferation von Zellen streng geregelt um die Ordnung im Zellverband sicherzustellen (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Tumorzellen stimulieren unabhängig von exogenen Faktoren ihr eigenes Wachstum (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). In diesem Zusammenhang spielen Wachstumsfaktorrezeptoren und ihre Liganden eine wichtige Rolle. In Tumorzellen kommt es zu einer Überaktivierung dieses Signalwegs durch verschiedene Mechanismen (Hanahan and Weinberg, 2011). Beispielsweise produzieren Tumorzellen die Wachstumsfaktoren, durch die sie selbst stimuliert werden (Hynes and MacDonald, 2009). Alternativ sind Wachstumsfaktorrezeptoren auf der Oberfläche der Tumorzellen überexprimiert, sodass externe Wachstumsstimuli unverhältnismäßig starke Wirkung zeigen (Hanahan and Weinberg, 2011). Ein Beispiel dafür ist der Her2/neu Rezeptor beim Mammakarzinom (Hynes and MacDonald, 2009, Lemmon and Schlessinger, 2010). Dieser Wachstumsfaktorrezeptor ist bei ca.20% der sporadischen Mammakarzinome als Folge einer Genamplifikation überexprimiert (Hynes and MacDonald, 2009) und mit einer schlechten Prognose assoziiert (Slamon et al., 1989). Die Überexpression macht man sich bei der Therapie Her2/neu positiver Mammakarzinome zunutze (siehe Kapitel 2.1.4). Neben einer Überexpression können Wachstumsfaktorrezeptoren oder intrazelluläre Signalwege, die den Rezeptoren nachgeschaltet sind, sich also „downstream“ befinden, durch Mutationen dauerhaft aktiv sein (Hynes and MacDonald, 2009, Lemmon and Schlessinger, 2010). Desweiteren stimulieren Tumorzellen andere Zellen in ihrer Umgebung, darunter Fibroblasten und inflammatorische Zellen, die wiederum Wachstumsfaktoren sezernieren (Bhowmick et al., 2004).

2.1.2.2 Evading growth suppressors

Normalerweise unterdrückt eine Reihe von intra- und extrazellulären wachstumshemmenden Faktoren die Zellproliferation und erhält die Homöostase in gesundem Gewebe (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Dabei spielen pRb und p53 als Tumorsuppressoren durch die Kontrolle des Zellzyklus eine wichtige Rolle (Sherr and McCormick, 2002). pRb wandelt inhibitorische Signale um, die von intra- und extrazellulär kommen, während p53 hauptsächlich intrazelluläre Stresssignale verarbeitet (Sherr and McCormick, 2002). Beispiele für letzteres sind Schäden am Genom und Sauerstoff-, Glukose- oder Nucleotidmangel (Hanahan and Weinberg, 2000). Eine übermäßige Zellproliferation, sei es durch Stimulation oder auch mangelnde Hemmung, kann dazu führen, dass die Zelle in einen Ruhezustand, die sogenannte Seneszenz, übergeht oder sogar apoptotisch wird (Collado and Serrano, 2010). Tumorzellen umgehen diesen Mechanismus, sei es, indem sie entweder hohe Level an proliferativen Signalen tolerieren oder indem sie stimulierende Signale unterdrücken (Collado and Serrano, 2010). Ein Beispiel dafür ist TGF- β , welches im Rahmen

der Brustentwicklung sowie auf normale epitheliale Zellen der Brustdrüse antiproliferativ wirkt (Moses and Barcellos-Hoff, 2011). Brustkrebszellen zeigen hingegen eine Resistenz gegenüber TGF- β und tolerieren hohe Konzentrationen, welche von Tumor- und Bindegewebszellen produziert werden (Moses and Barcellos-Hoff, 2011, Prud'homme, 2007). TGF- β führt im Tumor zu Modulation des Umgebungsmilieus und Immunsuppression und begünstigt Angiogenese, Tumorwachstum sowie Metastasierung (Moses and Barcellos-Hoff, 2011, Prud'homme, 2007).

2.1.2.3 Resisting cell death

Genau wie eine gesteigerte Proliferation führt auch die Vermeidung von Apoptose, also dem programmierten Zelltod, zu Tumorentstehung und Zunahme der Zellzahl (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Diese wird von Proteasen, sogenannten Caspasen, durchgeführt. Dabei gibt es zwei Hauptmechanismen, die zu ihrer Aktivierung führen (Adams, 2003, Adams and Cory, 2007): Einerseits von extrazellulär über sogenannte Todesrezeptoren wie z.B. den FAS Rezeptor (Strasser et al., 1995, Strasser et al., 2000), andererseits über intrazelluläre Stresssignale wie eine inadäquate Zytokinkonzentration oder Zellschäden (Adams, 2003). Anschließend werden die Zelle und ihre Bestandteile durch Proteolyse durch eine Reihe von weiteren Caspasen demontiert und zuletzt von umgebenden Zellen oder Makrophagen phagozytiert (Adams, 2003, Adams and Cory, 2007). In diesen komplexen Prozess sind viele pro- und antiapoptotische Faktoren involviert. Das proapoptotische Genprodukt p53 spielt dabei eine essentielle Rolle, indem es zu Zellzyklus Arrest und Reparatur führen oder auch Apoptose einleiten kann (Sherr and McCormick, 2002). Ohne ein funktionierendes p53 überleben Zellen mit Strangbrüchen, was zu Zellen mit neuen Mutationen führt (Junttila and Evan, 2009, Lowe et al., 2004). Bereits eine Metaanalyse von 1999 (Pharoah et al.) beschrieb p53 als negativen prognostischen Marker bei Patienten mit Mammakarzinom. Yadav et al. (2015) verfassten ein Review über beschriebene Biomarker, die die Prognose von Brustkrebspatienten mit triple-negativen Tumoren, also Tumoren, die weder Östrogen-Rezeptoren (ER), Progesteron- Rezeptoren (PR) noch Her2/neu-Rezeptoren exprimieren und mit einer schlechten Prognose assoziiert sind, beeinflussen können. Sie nannten unter anderem p53 als negativ prognostischen Parameter. Zugrunde liegt eine „missense“ Mutation von p53, welche zu einer stabilen Expression von p53 führt (Biganzoli et al., 2011, Chae et al., 2009). Die zweite Form des Zelltods neben der Apoptose ist die Nekrose. Nekrotische Zellen schwellen an, platzen und geben ihren Zellinhalt in die Umgebung ab (Zong and Thompson, 2006). Früher wurde Nekrose als eine unkontrollierte Form des Zelltods angesehen, die dann zustande kommt, wenn die Zelle auf massive schädliche innere und äußere Einflussfaktoren nicht mehr adäquat reagieren kann und zusammenbricht (Zong and Thompson, 2006). Mittlerweile ist man der Ansicht, dass Nekrose ein kontrollierter Prozess ist, der gezielt von der Zelle ausgelöst werden kann (Zong and Thompson,

2006). Ein essentieller Unterschied der beiden Formen von Zelltod ist, dass Nekrose im Gegensatz zur Apoptose proinflammatorisch wirkt und Reparatur- und Wachstumsmechanismen stimuliert (Zong and Thompson, 2006). Im Kontext der Tumorentstehung stimuliert Nekrose umgebende Zellen zur Proliferation und auch die durch sie angelockten Immunzellen fördern die Tumorentstehung (Grivennikov et al., 2010). Der durch sie entstandene Verlust an Tumorzellen wird so mehr als wieder wett gemacht (Hanahan and Weinberg, 2011). Im verwendeten 4T1 Mammakarzinommodell ist eine massive Nekrose in den Primärtumoren beschrieben (Tao et al., 2008). Tao et al. (2008) stellten fest, dass Nekrose in metastasierenden Tumoren ausgeprägter ist als in nicht metastasierenden Kontrolltumoren. Ebenfalls beobachteten sie eine ausgeprägte Immunreaktion in nekrotischen Tumorarealen. An dieser Stelle sollte auch die Autophagozytose genannt werden. Gerät die Zelle unter Stress, beispielsweise durch mangelnde Nährstoffversorgung oder auch Tumorthérapien wie Chemo- oder Strahlentherapie, kann die Autophagozytose, die auch in gesunden Zellen auf einem niedrigen Level stattfindet, gesteigert werden (Levine and Kroemer, 2008). Diese wirkt einerseits tumorsuppressiv, kann aber andererseits zu späten Tumorrezidiven führen, indem sich die Zelle vorübergehend in einen Ruhezustand begibt, aus dem sie, nach Therapie, wieder austreten kann (Levine, 2007, Levine and Kroemer, 2008).

2.1.2.4 *Enabling replicative immortality*

Tumorzellen unterscheiden sich von normalen Zellen durch die Fähigkeit eine unbegrenzte Zahl an Zellteilungen durchlaufen zu können (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Telomere sind repetitive DNA Sequenzen an den Enden der DNA und limitieren die Zahl der physiologisch möglichen Zellteilungen, da bei jeder Teilung ein Stück von ihnen verloren geht (Blasco, 2005). Das Enzym Telomerase, eine DNA Polymerase, ist in der Lage die Telomerenden zu verlängern und wird von den meisten Zellen, die zur unbegrenzten Zellteilung in der Lage sind, also auch von Tumorzellen, exprimiert (Blasco, 2005). Dabei ist es aber wohl so, dass Tumorzellen diese Fähigkeit erst relativ spät erlangen (Hanahan and Weinberg, 2011). Betrachtet man Frühstadien von Tumoren, gibt es viele Zellen mit verkürzten DNA Enden, die teilweise sogar schon in den Zelltod eintreten (Artandi and DePinho, 2010). So wird möglicherweise in vielen Fällen verhindert, dass Frühstadien von Tumoren weiter fortschreiten (Hanahan and Weinberg, 2011). Kommt aber beispielsweise eine Fehlfunktion von p53 hinzu, das die Integrität der DNA überwacht, kann es bei aufgebrauchten Telomeren durch DNA-Verknüpfungen zu Deletionen und Amplifikationen kommen, die zur Entstehung von Onkogenen und zum Verlust von Tumorsuppressorgenen beitragen können (Artandi and DePinho, 2010). Somit fördert ein Fehlen von Telomerase zu Beginn die Mutagenität. Sie muss aber zu einem späteren Zeitpunkt exprimiert werden, um eine unbegrenzte Teilung zu ermöglichen (Hanahan and Weinberg, 2011). Insgesamt wird die Telomerase von 85% aller Tumore exprimiert. Bei

Adenokarzinomen, also auch bei Mammakarzinomen, sogar von fast 100%, was sie als Angriffspunkt für individualisierte Tumorthherapie interessant macht (Holysz et al., 2013).

2.1.2.5 *Inducing angiogenesis*

Tumore benötigen Gefäße, um ausreichend mit Sauerstoff und Nährstoffen versorgt zu werden und um Abfallprodukte und Kohlendioxid zu entsorgen (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Physiologischerweise finden die Vaskulogenese, also die Entstehung von Gefäßen, und die Angiogenese, also das Aussprossen von Gefäßen aus bereits vorhandenen, vor allem in der Embryogenese statt (Hanahan and Weinberg, 2011). Im adulten Organismus unterliegen diese Mechanismen einer starken Regulation und sind induzierbar. So werden sie beispielsweise bei der Wundheilung und im weiblichen Zyklus angeschaltet (Hanahan and Weinberg, 2011). Im Tumor kommt es über den sogenannten „angiogenic switch“ zu einer kontinuierlichen Gefäßneubildung (Bergers and Benjamin, 2003). Dabei spielen proangiogenetische Faktoren wie z.B. VEGF-A eine entscheidende Rolle (Bergers and Benjamin, 2003). Die Menge proangiogenetischer Faktoren kann durch Onkogene wie Myc oder Ras, die in den Tumorzellen gebildet werden, aber auch durch Immunzellen, die diese Faktoren sezernieren, gesteigert werden (Hanahan and Weinberg, 2011). Auch beim Mammakarzinom spielen proangiogenetische Faktoren und Angiogenese eine wichtige Rolle. Die Dichte von Mikrogefäßen in Tumoren von Brustkrebspatienten ist mit einer negativen Prognose assoziiert (Yadav et al., 2015). Desweiteren korrelieren VEGF Konzentrationen im Serum sowie in Tumorgewebe mit negativen prognostischen Markern wie einem höheren Grading, größeren Tumoren, positivem Lymphknotenstatus, negativem Hormonrezeptorstatus sowie schlechterem Ansprechen auf Chemotherapie (Yadav et al., 2015). Auch diesen Mechanismus macht man sich bei der Therapie des Mammakarzinoms zunutze (siehe 2.1.4). Neben proangiogenetischen gibt es anti-angiogenetische Faktoren wie TSP-1, die der Tumorentstehung entgegen wirken und in Tumoren oft vermindert sind (Hanahan and Folkman, 1996). Eigenschaften von Tumorgefäßen sind ein chaotisches und abnormes, spiralenfrörmiges Wachstum, Undichtigkeit und eine hohe Rate an Proliferation und Apoptose in den Endothelzellen (Nagy et al., 2010).

2.1.2.6 *Activating invasion and metastasis*

Invasives Wachstum und Metastasierung sind weitere Schritte bei der Tumorentstehung (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Die Epithelial-Mesenchymale Transition (EMT) ist ein Mechanismus, der bei der Metastasierung von Karzinomen, also von Tumoren, die von Epithelien ausgehen, eine wichtige Rolle spielt (Hugo et al., 2007). Durch vermehrte Expression bestimmter Transkriptionsfaktoren, die auch in der Embryologie eine Rolle spielen, kommt es zur Umwandlung einer polygonal/ epithelialen Form zu einer spindelförmig/ fibroblastischen, zu einer vermehrten Expression Matrix degradierender Enzyme, zu gesteigerter Beweglichkeit und zu

vermehrter Resistenz gegenüber Apoptose (Hanahan and Weinberg, 2011). In diesem Prozess ist E-Cadherin, eines der wichtigsten interzellulären Verbindungsmoleküle, das physiologischerweise für die Integrität und Stabilität von Zellverbänden sorgt, oft herab reguliert oder inaktiviert (Berx and van Roy, 2009, van Roy and Berx, 2008). Zusätzlich sind mesenchymale Verbindungsmoleküle, die bei Zellmigrationen in der Embryogenese oder von Immunzellen eine Rolle spielen, wie beispielsweise N-Cadherin, hoch reguliert (Cavallaro and Christofori, 2004). Dabei spielt die Interaktion von Tumorzellen mit umgebenden Zellen wie mesenchymalen Stammzellen (Karnoub et al., 2007) oder Immunzellen (Grivnickov et al., 2010) eine wichtige Rolle. EMT mit den oben genannten molekularbiologischen Veränderungen, beispielsweise eine Begünstigung von Metastasierung durch vermehrte Expression von N-Cadherin, ist auch bei zahlreichen Versuchen mit metastasierenden Mammakarzinom Zelllinien in vivo und in vitro ausführlich beschrieben (Hugo et al., 2007). Eine Inaktivierung von E-Cadherin ist bei allen lobulären Mammakarzinomen zu finden und dient als histopathologisches Unterscheidungsmerkmal zu duktalem Mammakarzinomen (Böcker, 2008) (zu duktales vs. lobuläres Mammakarzinom siehe 2.7.2). Neben der EMT werden noch andere Mechanismen der Metastasierung beschrieben. Beispielsweise die sogenannte „collective invasion“, bei der Zellnester in umgebendes Gewebe eindringen. Diese Art der Metastasierung wurde bei Plattenepithelkarzinomen beobachtet, wobei diese interessanterweise selten metastatisch sind, was eher gegen die Effektivität dieser Methode spricht (Friedl and Wolf, 2010). Desweiteren wird eine amöboide Art des invasiven Wachstums beschrieben, bei der die Tumorzellen zwischen den anderen Zellen hindurch gleiten, ohne dass umliegendes Gewebe beiseite geräumt wird (Lammermann and Sixt, 2009). Das Modell der EMT versucht vor allem die Dissemination der Tumorzellen darzustellen. Wie, warum und wann sich letztlich Mikro- oder Makrometastasen bilden, ist bislang noch nicht vollständig geklärt (Hanahan and Weinberg, 2011). Am Ort des Tumorgeschehens kommt es zu Durchbrüchen durch die Basalmembran, dann zum Einbruch in Blut- und Lymphgefäße. Durch diese gelangen Tumorzellen in entfernte Organe, wo sie letztendlich wieder in das Gewebe auswandern (Talmadge and Fidler, 2010). Haben Tumorzellen Metastasen gebildet und sind ihrem ursprünglichen Milieu nicht mehr ausgesetzt, nehmen sie wieder ihre ursprüngliche Form an. Dieser Vorgang wird Mesenchymal-Epitheliale Transition (MET) genannt (Chaffer et al., 2007, Hugo et al., 2007). Tumorzellen bilden jedoch teilweise Mikrometastasen, die niemals zu makroskopischen Metastasen heranwachsen, sondern in einer Art Ruhezustand verharren (Aguirre-Ghiso, 2007). Es gibt Theorien, die diesen Zustand durch hemmende Signale des Primärtumors erklären, was die begründen würde, warum sich einige Metastasen nach der Tumorthherapie explosionsartig ausbilden (Demicheli et al., 2008). In anderen Fällen, wie beispielsweise beim Mammakarzinom oder dem Melanom, kann es noch Jahrzehnte nach erfolgreicher Therapie des Primärtumors zur Ausbildung von Makrometastasen kommen (Hanahan and Weinberg, 2011). Erklärungsmodelle dafür sind eine

ungenügende Interaktion zwischen disseminierten Tumorzellen und umgebendem Bindegewebe, eine mangelnde Fähigkeit zur Angiogenese oder eine Unterdrückung durch das Immunsystem (Aguirre-Ghiso, 2007). Es gibt außerdem Hinweise darauf, dass disseminierte Tumorzellen aufgrund eines Nährstoffmangels durch Autophagozytose in einen Ruhezustand übergehen können, aus dem sie Jahre später wieder erwachen (Kenific et al., 2010). Obwohl Metastasierung klinisch vor allem bei fortgeschrittenen Tumoren auftritt, gibt es auch Hinweise dafür, dass Tumorzellen aus Tumorfrühstadien disseminieren können (Coghlin and Murray, 2010). Das ist auch für das Mammakarzinom beschrieben (Husemann et al., 2008). Ob dadurch aber Makrometastasen entstehen, bleibt fraglich (Hanahan and Weinberg, 2011). Zusammenfassend muss man sagen, dass vor allem über die Vorgänge, die bei der Kolonisierung und Bildung von Mikro- und Makrometastasen eine Rolle spielen, noch eine Reihe offener Fragen existieren. Sicher scheint, dass dafür nicht nur Eigenschaften der Tumorzellen selbst wichtig sind, sondern die Etablierung eines günstigen Mikromilieus und ein erfolgreiches Zusammenspiel mit den ortsansässigen Bindegewebs- und Immunzellen unabdingbar sind (Hanahan and Weinberg, 2011).

2.1.2.7 Reprogramming of energy metabolism

Schon Warburg (1930) erkannte, dass Tumorzellen einige Besonderheiten in ihrem Metabolismus aufweisen, da sie trotz aeroben Bedingungen vor allem auf die normalerweise nur unter anaeroben Bedingungen ablaufende energetisch ineffizientere Glykolyse umschalten, was als aerobe Glykolyse bezeichnet wird. Um diesen Nachteil auszugleichen, importieren sie große Mengen an Glukose in das Zytosol, was sich bei der Positronen-Emissions-Tomographie (PET) zur Diagnostik von Tumoren zu Nutze gemacht wird (Hanahan and Weinberg, 2011). Ein Vorteil dieser ineffizienten Energiegewinnung ist, dass viele Metabolite, die bei der Glykolyse entstehen, in andere Stoffwechselwege eingeschleust und beispielsweise zum Aufbau von Nukleosiden und Aminosäuren genutzt werden können, was für sich teilende Zellen vorteilhaft ist (Phan et al., 2014, Vander Heiden et al., 2009). In einigen Tumoren existieren zwei Zelltypen mit unterschiedlichem Metabolismus nebeneinander. Einer produziert über Glykolyse Laktat, während der andere, besser mit Sauerstoff versorgte dieses nutzt (Kennedy and Dewhirst, 2010).

2.1.2.8 Genetic Instability

Die oben beschriebenen Veränderungen in der Physiologie der Tumorzelle basieren auf genetischen Mutationen, welche in gesundem Gewebe ein sehr seltenes Ereignis sind, in Tumorzellen aber vermehrt stattfinden (Hanahan and Weinberg, 2011). Dies kommt sowohl durch eine erhöhte Sensitivität gegenüber mutagenen Substanzen, als auch durch eine Einschränkung des DNA Überwachungs- und Reparatursystems in Tumoren zustande (Hanahan and Weinberg, 2011). Tumorzellen unterscheiden sich auch in ihrem Epigenom wie beispielsweise ihrem

Methylierungsmuster von normalen Zellen und eine Assoziation dieser Veränderungen mit genetischer Instabilität gilt als gesichert (Berdasco and Esteller, 2010).

Die Rolle des Immunsystems in der Onkologie ist für diese Arbeit besonders relevant und wird unter 2.3.2 ausführlich besprochen.

2.1.3 Tumorstammzellen und Zellen des Tumor-Microenvironments

Tumore sind sehr heterogene und komplexe Gewebe, die eine Vielzahl verschiedener Zellen enthalten (Hanahan and Weinberg, 2011). Eine wichtige Subpopulation stellen die sogenannten Tumor-initiiierenden Tumorstammzellen dar (Lobo et al., 2007). Sie haben ein bestimmtes Muster an Oberflächenmarkern und können sowohl Stammzellen als auch herkömmliche Tumorzellen hervorbringen (Dalerba et al., 2007). Ihre Entstehung ist nicht eindeutig geklärt bzw. unterscheidet sich möglicherweise je nach Tumorentität. Eine Theorie ist, dass Gewebestammzellen bei der Tumorentstehung zu Tumorstammzellen werden, die dann ähnliche Eigenschaften haben wie die Ursprungszellen: Sie haben die Fähigkeit zur Selbsterneuerung sowie zur Bildung weiter differenzierter Nachkommen. Es ist aber auch möglich, dass sie aus Vorstufen von Tumorzellen hervorgehen, die dann Stammzelleigenschaften erwerben (Hanahan and Weinberg, 2011). Es wird diskutiert, dass im Rahmen der EMT Tumorstammzellen zu normalen Tumorzellen werden können und umgekehrt (Gupta et al., 2009). Man vermutet, dass Zellen mit Stammzellcharakter einer Chemo- oder Strahlentherapie schwerer zugänglich sind und bei späten Tumorrezidiven eine Rolle spielen (Housman et al., 2014, Singh and Settleman, 2010). Auch Endothelzellen, Perizyten und Fibroblasten tragen zur zellulären Diversität in Tumoren bei und spielen in der Tumorbio­logie eine wichtige Rolle (Hanahan and Weinberg, 2011). Endothelzellen in Tumoren unterscheiden sich durch ihre Transkriptionsprofile und Signalwege von denen in gesundem Gewebe (Nagy et al., 2010, Ruoslahti, 2002) und tragen durch den „angiogenic switch“, der es ihnen ermöglicht, neue Gefäße zu bilden, zur Tumorangiogenese bei (Hanahan and Weinberg, 2011). Perizyten umgeben Endothelzellen von Arteriolen, Venolen und Kapillaren fingerförmig und sorgen für die Integrität und Stabilität der Gefäße (Hanahan and Weinberg, 2011). In Tumoren beeinflussen Perizyten Endothelzellen und deren Fähigkeit zu Angiogenese (Bergers and Song, 2005). Man vermutet sogar, dass sie eine Rolle bei Metastasierung spielen können (Raza et al., 2010). In Tumoren sind neben den Fibroblasten, die auch in normalem Gewebe zu finden sind, Myofibroblasten, die α -smooth muscle actin (SMA) exprimieren und in gesundem Gewebe selten sind, zu nennen. Myofibroblasten und rekrutierte Fibroblasten fördern Angiogenese, Proliferation, Invasion und Metastasierung (Bhowmick et al., 2004).

2.1.4 Individualisierte Tumorthherapie

Aus den oben erläuterten Erkenntnissen entstand die Hoffnung, Tumorforschung und -therapie individueller und präziser gestalten zu können (Hanahan and Weinberg, 2011). Ziel der sogenannten „targeted therapy“ ist Tumorzellen durch eine gezielte Störung der oben beschriebenen essentiellen Mechanismen anzugreifen und dabei gesunde Zellen möglichst weitgehend zu schonen (Hanahan and Weinberg, 2011). Beispiele dafür sind VEGF Inhibitoren, die einer Vaskularisierung entgegen wirken, Inhibitoren der aeroben Glykolyse, EGFR Inhibitoren, die Wachstumsfaktoren blockieren, antiinflammatorische Medikamente usw. (Hanahan and Weinberg, 2011).

Mammakarzinome beispielsweise, die den Wachstumsfaktorrezeptor Her2/neu überexprimieren, werden mit dem monoklonalen Antikörper Trastuzumab, einem EGFR Inhibitor, therapiert (Leitlinienprogramm Onkologie, 2012). Dieser blockiert die extrazelluläre Domäne der Her2-Rezeptors und somit seine Aktivierung. Desweiteren verhindert er dessen Dimerisierung, führt über Aktivierung einer zellvermittelten Zytotoxizität zur Tumorzelllyse und zur Internalisierung des Rezeptors in die Zelle (Boekhout et al., 2011). Desweiteren ist Bevacicumab, ein VEGF Inhibitor, als „targeted therapy“ von der European Medicines Agency (2016) bei metastasiertem Mammakarzinom in Kombination mit Chemotherapie zugelassen. Die Wirksamkeit von Bevacicumab ist jedoch umstritten. Kumler et al. (2014) schrieben in ihrem Review, dass es bei teilweise verlängertem progressionsfreien Überleben kein Hinweis auf ein verlängertes Gesamtüberleben gab. Auch die amerikanische US-Arzneibehörde zog die Zulassung für Bevacicumab bei Mammakarzinom aufgrund fehlender Verlängerung des Gesamtüberlebens sowie beträchtlicher Nebenwirkungen zurück (Food and drug association, 2011).

Leider stellte sich der Effekt der „targeted therapy“ in der Mehrzahl der Fälle als vorübergehend heraus und dem anfänglich guten Ansprechen folgt fast immer ein Rezidiv (Hanahan and Weinberg, 2011). Hanahan and Weinberg (2011) erklärten diesen Effekt durch redundante Signalwege. Inhibiert man einen davon erfolgreich, sind andere noch aktiv und durch den entstehenden Selektionsdruck können durch neue Mutationen, epigenetische Modifikationen oder Veränderungen im tumorassoziierten Bindegewebe weitere Signalwege eröffnet werden. Eine andere Ursache für Rezidive könnte sein, dass Tumore, wenn eine „hallmark capability“ durch individualisierte Tumorthherapie ausgeschaltet wird, vermehrt auf andere zurückgreifen und somit ihr Überleben sichern (Hanahan and Weinberg, 2011). Daraus folgt, dass eine erfolgreiche, dauerhafte Tumorthherapie an vielen Regulationsmechanismen gleichzeitig angreifen müsste (Hanahan and Weinberg, 2011).

2.2 Hämatopoese und extramedulläre Hämatopoese

In dieser Arbeit wird EMH in der Leber anhand eines orthotopen spontan metastasierenden 4T1 Mammakarzinom Mausmodells quantifiziert. Deshalb soll im Folgenden zunächst ein Überblick über die physiologische Hämatopoese im Knochenmark gegeben werden. Anschließend wird EMH definiert und ihr Vorkommen erläutert.

Lüllmann-Rauch (Blut und Blutbildung, 2006) schrieb über die Hämatopoese, also die Blutbildung, Folgendes: Die Hämatopoese findet ausschließlich im Knochenmark statt. Eine Ausnahme bildet die Vermehrung der Lymphozyten, die vor allem in den lymphatischen Organen lokalisiert ist. Man unterscheidet rotes von gelbem Knochenmark, wobei die Hämatopoese nur im roten Knochenmark stattfindet. Beim Kind füllt das rote Knochenmark fast alle Knochen aus, beim Erwachsenen findet man es nur in wenigen Knochen wie u.a. Sternum, Wirbelkörper und Beckenkamm. Die Hämatopoese wird durch Zytokine, also „lösliche Proteine, die der Kommunikation zwischen den Abwehrzellen dienen“ (Lüllmann-Rauch, Blut und Blutbildung, 2006), reguliert.

Die Hämatopoese wird meist als hierarchisches System beschrieben, in dem alle Blutzellen von einer Gruppe multipotenter hämatopoetischer Stammzellen (HSCs) im Knochenmark abstammen, die einerseits die Fähigkeit zur lebenslangen Selbsterneuerung und andererseits zur Differenzierung zu Progenitorzellen der verschiedenen Zelllinien besitzen (Orkin, 2000).

Shizuru et al. (2005) unterschieden „long-term hematopoietic stem cells“ (LT-HSC), die sich lebenslang selbst erneuern können und „short-term“ (ST)-HSCs, die diese Eigenschaft nur für 6-8 Wochen tragen. Dabei gehen aus den HSCs myeloische (CMPs, „common myeloid progenitors“) und lymphatische (CLPs, „common lymphoid progenitors“) Progenitorzellen hervor, welche sich dann zu Vorläufern der jeweiligen Zellgruppen entwickeln (siehe Abbildung 1). Ob eine HSC den Weg der Selbsterneuerung oder der Differenzierung wählt, wird von einer Reihe von Zytokinen und Transkriptionsfaktoren gesteuert (Shizuru et al., 2005). Shizuru et al. (2005) nannten den Wnt-Signalweg als zentrales Element für die Fähigkeit zur Selbsterneuerung von HSCs. In Anwesenheit von Zytokinen wie beispielsweise IL-11, IL-6, „granulocyte colony-stimulating factor“ (G-CSF), Thrombopoietin, „macrophage colony-stimulating factor“ (M-CSF) oder „granulocyte-macrophage colony-stimulating factor“ (GM-CSF) kommt es zu einer Differenzierung der HSCs zu Progenitorzellen (Shizuru et al., 2005).

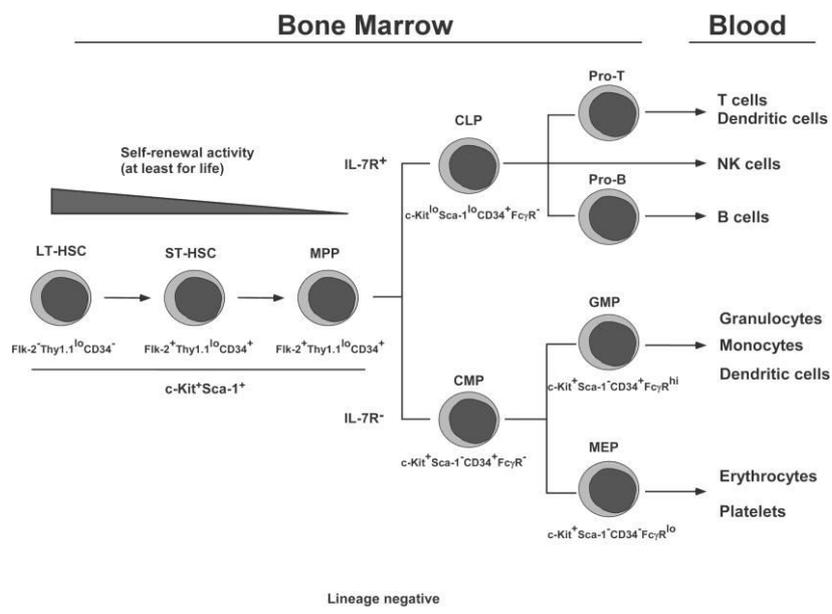


Abbildung 1: Stammbaum der Blutzellen. LT-HSC, long-term hematopoietic stem cells; short-term ST-HSC, short-term HSC; MPP, multipotent progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte/macrophage progenitor; MEP, megakaryocyte/erythrocyte progenitor. Abbildung und Bezeichnungen aus Shizuru et al. (2005); angelehnt an (Lüllmann-Rauch, Blut und Blutbildung, 2006);

Medvinsky et al. (2011) schrieben in ihrem Review Folgendes: Grundsätzlich wird zwischen embryonaler und definitiver Hämatopoese unterschieden. Aus verschiedenen Gründen ist es schwierig, die Entwicklung der Hämatopoese in der Embryonalzeit zu untersuchen, weshalb sich die meisten Forschungsergebnisse aus Mausversuchen ergaben. In der embryonalen Hämatopoese werden vorübergehend primitive Erythrozyten, also kernhaltige, große Erythrozyten, und myeloische Zellen gebildet. Definitive Hämatopoese definieren die Autoren als das Vorhandensein von definitiven hämatopoetischen Stammzellen (dHSCs), die alle Blutzellreihen bilden können und die Fähigkeit zur Selbsterneuerung haben. Als erste Quelle hämatopoetischer Zellen gelten die Blutinseln im Dottersack. Dort entstehen zunächst primitive Blutzellen, die aber wohl nicht aus dHSCs hervorgehen. Es wird diskutiert, dass dHSCs zunächst in der sogenannten AGM-Region (Aorta, Gonaden, Mesonephros) entstehen.

Cumano and Godin (2007) stellten die Rolle der Leber in der fetalen Hämatopoese folgendermaßen dar: Um ca. den 10. Tag post conceptionem siedeln sich hämatopoetische Zellen in der Leber fetaler Mäuse an. Man geht davon aus, dass es sich dabei zum großen Teil um Progenitorzellen aus dem Dottersack handelt, die sich in der Leber unter Erythropoietin Exposition zu kernlosen Erythrozyten differenzieren. Außerdem findet in der Leber die Proliferation und endgültige Differenzierung von HSCs mit all ihren Oberflächenmarkern statt. In der Leber herrscht ein Milieu, das sich deutlich von dem im Knochenmark unterscheidet: Es begünstigt eine starke Proliferation der Stammzellen. Zusätzlich herrscht eine weniger strenge hierarchische Organisation in ihrem

Differenzierungspotential. Dabei ist es aufgrund fehlender bekannter Oberflächenmarker noch nicht gelungen, das Zusammenspiel zwischen Leberzellen und hämatopoetischen Zellen darzustellen. Aus der Leber wandern HSCs und Progenitorzellen wohl kurz vor der Geburt in das Knochenmark, welches schließlich den einzigen verbleibenden Ort der Hämatopoese darstellt.

Dort gibt es dreimal so viele Zellen der Granulopoese wie der Erythropoese, obwohl auf 1000 Erythrozyten im peripheren Blut nur ein Leukozyt kommt, was an der sehr kurzen Lebensdauer der Granulozyten von maximal 3 Tagen in der Peripherie liegt (Lüllmann-Rauch, Das Immunsystem, 2006). Erythrozyten haben hingegen eine Verweildauer von 120 Tagen (Lüllmann-Rauch, Blut und Blutbildung, 2006).

Extramedulläre Hämatopoese (EMH) ist definiert als abnormale Blutbildung außerhalb der Grenzen des Knochenmarks und kann bei einer Reihe von Krankheiten beobachtet werden (Shawker et al., 1987). Krams et al. (2013) schrieben, dass eine Verdrängung der normalen Blutbildung zur Reaktivierung der extramedullären Blutbildung führen kann. Als Ursachen nannten sie chronisch myeloproliferative Erkrankungen, die akute myeloische Leukämie sowie ausgedehnte Knochenmarkmetastasen und maligne Lymphome. Bei der EMH siedeln sich differenzierte Vorläuferzellen in der Milz und in den Sinusoiden der Leber an. Anders als im Knochenmark gibt es dort keinen Schutz gegen das Ausschwemmen unreifer Blutzellen, weshalb sich im peripheren Blut granulopoetische und erythropoetische Zellen aller Reifungsstufen finden (pathologische Linksverschiebung) (Krams et al., 2013). Neben Leber und Milz können auch Lymphknoten betroffen sein oder sich paraspinale EMH Herde ausbilden, die sich dann sogar auf den Spinalkanal ausweiten und Rückenschmerzen und eine Querschnittssymptomatik auslösen können (Hashmi et al., 2014).

2.3 Das Immunsystem und seine Rolle in der Onkologie

Die Rolle des Immunsystems in der Onkologie ist für diese Arbeit besonders relevant und wird deshalb ausführlich besprochen. Zum besseren Verständnis soll davor ein Überblick über Funktion, Organisation und Histologie des Immunsystems und der Immunzellen gegeben werden.

2.3.1 Immunsystem und Immunzellen – Funktion, Organisation und histologische Charakterisierung

Man unterscheidet das unspezifische und das spezifische Immunsystem, wobei diese funktionell eng miteinander verbunden sind (Delves and Roitt, 2000, Parkin and Cohen, 2001). Die angeborene Immunität zeichnet sich durch ihre Schnelligkeit aus, führt aber manchmal durch ihre unspezifische Reaktion zu Schäden am Gewebe. Die erworbene Immunität ist im Gegensatz dazu erst nach einigen Tagen bis Wochen voll einsatzbereit (Parkin and Cohen, 2001). Zur unspezifischen Abwehr zählen phagozytierende Zellen wie Granulozyten, Makrophagen und Monozyten sowie Zellen, die

Mediatoren sezernieren (Mastzellen, Basophile, Eosinophile), natürliche Killerzellen (NK-Zellen) sowie Zytokine, Komplement und Akute Phase Proteine (Delves and Roitt, 2000, Parkin and Cohen, 2001). Ein zentraler Punkt der unspezifischen Abwehr ist die Rekrutierung von neutrophilen Granulozyten aus dem Knochenmark und der Peripherie, beispielsweise durch ortsansässige Makrophagen, die Zytokine wie G-CSF und GM-CSF sezernieren (Parkin and Cohen, 2001). Erkennen von pathogenen Mikroorganismen anhand von Rezeptoren (z.B. Toll-like-Rezeptoren), Verstärkung der Entzündungsreaktion durch Sekretion von Zytokinen und Phagozytose, also die Aufnahme von Partikeln oder kleinen Zellen in die phagozytierende Zelle, sind Abwehrmechanismen von Granulozyten und Makrophagen (Lüllmann-Rauch, Das Immunsystem, 2006). Makrophagen können in die Subtypen M1 und M2 unterteilt werden (Sica et al., 2008): Der M1 Typ sezerniert vor allem pro-inflammatorische Zytokine, exprimiert „major histocompatibility complex“ (MHC) Moleküle und ist in der Bekämpfung von Pathogenen und Tumorzellen beteiligt. Der M2 Typ wirkt eher anti-inflammatorisch und reguliert MHC Moleküle herunter. Neutrophile Granulozyten spielen in der Immunabwehr von Bakterien eine besonders wichtige Rolle (Lüllmann-Rauch, Blut und Blutbildung, 2006).

Colotta et al. (1992) schrieben, dass neutrophile Granulozyten, einmal in einen Entzündungsherd migriert, von dort nicht mehr ins Blut zurückkehren, die kürzeste Halbwertszeit aller Leukozyten haben und in-vitro schnell apoptotisch werden. Die Arbeitsgruppe fand heraus, dass die Überlebenszeit der neutrophilen Granulozyten von bestimmten Zytokinen und mikrobiellen Substanzen, die im Entzündungsherd vermehrt vorhanden sind, verlängert werden kann. Sie nannten unter anderem GC-SF, GM-CSF, TNF, IL-1 β , IL-6, IFN- γ und LPS.

Die Effektivität der Neutrophilen kann um ein Vielfaches gesteigert werden, wenn die Partikel durch Faktoren des Komplementsystems oder Antikörper opsoniert sind (Parkin and Cohen, 2001). Komplementfaktoren sind Glykoproteine, die in einem Kaskadensystem organisiert sind, das, einmal ausgelöst, zu einer massiven Verstärkung der eigenen Wirkung führt (Parkin and Cohen, 2001). Eosinophile dienen der Abwehr von Parasiten (Parkin and Cohen, 2001). Mastzellen und basophile Granulozyten tragen IgE spezifische Rezeptoren, deren Quervernetzung nach Bindung von IgE zur Degranulation und Freisetzung zahlreicher Mediatoren wie Histamin und Serotonin führt (Parkin and Cohen, 2001). Außerdem spielen sie eine zentrale Rolle bei allergischen Erkrankungen (Delves and Roitt, 2000). Die Interaktion von Mastzellen und basophilen Granulozyten mit dem erworbenen Immunsystem ist Gegenstand aktueller Forschungen. Mastzellen können die Lebensdauer und Proliferation von B-Zellen beeinflussen sowie die Differenzierung zu IgA produzierenden Plasmazellen stimulieren (Merluzzi et al., 2010) und deren IgE Synthese induzieren (Pawankar et al., 1997). Basophile Granulozyten können B-Zell stimulierende T-Helfer Zellen induzieren (Denzel et al., 2008)

sowie das Überleben von B-Zellen verlängern (Rodriguez Gomez et al., 2010). Lüllmann-Rauch (Das Immunsystem, 2006) schrieb, dass NK- Zellen Virus-infizierte oder entartete körpereigene Zellen erkennen und sie abtöten können. Dabei spielen wie bei den später beschriebenen zytotoxischen T-Zellen Granzyme, die zur Gruppe der Serin Proteasen gehören, und Perforine, also „pore-forming proteins“ eine wichtige Rolle (Voskoboinik et al., 2015).

Voskoboinik et al. (2015) schrieben dazu Folgendes: Granula der zytotoxischen Lymphozyten, zu denen NK-Zellen sowie zytotoxische T-Zellen gezählt werden, enthalten Perforin und Granzym. Perforine werden in die Membran der Zielzellen eingebaut, sodass Granzyme eingeschleust werden können. Letztere induzieren dann Apoptose in den Zielzellen.

Die Zellen des angeborenen Immunsystems können fremde Zellen durch sogenannte „pathogen-associated molecular patterns“ (PAMPs) wie beispielsweise Lipopolysaccharide (LPS) erkennen, die an bestimmte Rezeptoren an ihrer Oberfläche binden (Parkin and Cohen, 2001). Zytokine, darunter vor allem Chemokine, und Adhäsionsmoleküle spielen in der Koordination und Rekrutierung aller Immunzellen eine essentielle Rolle (Parkin and Cohen, 2001).

Die Hauptvertreter der spezifischen Immunabwehr sind B-Lymphozyten und T-Lymphozyten (Delves and Roitt, 2000, Parkin and Cohen, 2001). Beide entstehen aus Progenitorzellen im Knochenmark, wobei nur B-Zellen während ihrer ganzen Entwicklung dort verbleiben, T-Zellen jedoch zum Teil im Thymus reifen (Delves and Roitt, 2000, Parkin and Cohen, 2001). Beide besitzen Antigen-spezifische Rezeptoren auf ihrer Zelloberfläche, die durch Rekombination von DNA Abschnitten, die für die Antigenbindestelle kodieren, entstehen (Delves and Roitt, 2000, Parkin and Cohen, 2001). Dieser Vorgang wird als VDJ Rekombination bezeichnet und findet auch bei der Bildung der Antikörper von Plasmazellen, also reifen B-Lymphozyten, statt. Sie ermöglicht die extrem große Variabilität unterschiedlicher Antikörper bzw. B- und T-Zell-Rezeptoren (Delves and Roitt, 2000). Nach ihrer Reifung verlassen naive B- und T-Zellen Knochenmark bzw. Thymus und siedeln sich in den lymphatischen Organen wie Lymphknoten oder Milz an (Parkin and Cohen, 2001). T-Zellen werden in CD8+ zytotoxische T-Zellen und CD4+ T-Helfer-Zellen unterteilt (Grivennikov et al., 2010). Letztere beinhalten Th1, Th2, Th17, regulatorische T-Zellen (Tregs) und NK-Zellen (Grivennikov et al., 2010). Tregs sind definiert als CD25+CD4+ Zellen, die den Transkriptionsfaktor „Forkhead box P3 transcription factor“ (Foxp3) exprimieren und eine wichtige Rolle bei der Selbsttoleranz spielen (Nishikawa and Sakaguchi, 2010). So sind sie beispielsweise bei der Entwicklung von Autoimmunerkrankungen beteiligt (Nishikawa and Sakaguchi, 2010).

Parkin and Cohen (2001) schrieben desweiteren: Aktivierte B-Zellen, sogenannte Plasmazellen, produzieren Antikörper, die Toxine neutralisieren, Komplement aktivieren, Bakterien opsonieren und entartete und Virus infizierte Zellen markieren können. Somit verstärken sie die Wirkung des

angeborenen Immunsystems. Im Gegensatz zu T-Zellen verbleiben sie meist in lymphatischen Organen. T-Zellen können zwischen lymphatischen Organen migrieren. Durch ein spezielles Milieu aus Antigen-präsentierenden Zellen (APCs), wie beispielsweise Dendriten (DCs), und Zytokinen wird das Zusammentreffen von Antigen und passendem Lymphozyten ermöglicht. Antigene werden über MHC Moleküle präsentiert, wobei CD8+ T-Zellen präsentierte Moleküle an MHC I Rezeptoren und CD4+ T-Zellen Moleküle auf MHC II Rezeptoren erkennen. CD4+ T-Helfer-Zellen haben eine koordinierende Funktion und spielen auch bei der B-Zell-Aktivierung eine Rolle.

CD8+ zytotoxische Zellen wirken ähnlich wie NK-Zellen und können entartete oder virusbefallene Zellen durch Perforine und Granzyme erkennen und abtöten (Voskoboinik et al., 2015).

Lüllmann-Rauch (Blut und Blutbildung, 2006) beschrieb die Histologie der Immunzellen: Ausgereifte Granulozyten sind 10-12 µm groß und enthalten zahlreiche Granula in ihrem Zytoplasma. Je nachdem, wie sich das Zytoplasma in der Pappenheimfärbung anfärbt, unterscheidet man neutrophile (fliederfarbendes Zytoplasma), basophile und eosinophile Granulozyten. Der Zellkern besteht aus mehreren Segmenten (deshalb auch segmentkerniger Granulozyt). Vorstufen werden wegen ihrem C-förmigen und noch nicht segmentierten Kern stabkernige Granulozyten genannt. Die unterschiedlichen Lymphozyten lassen sich im Blutaussstrich nicht sicher unterscheiden. Sie sind zwischen 7,5 und 15 µm groß, haben einen runden, chromatindichten Kern und einen schmalen hellblauen Zytoplasmasaum. Monozyten sind die größten Leukozyten (bis 20 µm), haben einen nierenförmigen Zellkern und blaugraues Zytoplasma.

2.3.2 Die Rolle des Immunsystems in der Onkologie

„Evading immune destruction“ ist ein weiteres „hallmark of cancer“, das Hanahan and Weinberg (2011) ausführlich beschrieben. Es gibt Hinweise darauf, dass Immunzellen sowohl einen fördernden als auch einen hemmenden Effekt auf die Tumorentstehung und -progression haben (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011). Nahezu jeder Tumor wird von Zellen des angeborenen und des erworbenen Immunsystems infiltriert (Pages et al., 2010), wobei vor allem das angeborene Immunsystem tumorfördernd wirkt. Die Frage, ob und inwiefern das Immunsystem daran beteiligt ist, neu entstehende Neoplasien, fortgeschrittene Tumore oder Metastasen zu entdecken und zu eliminieren, ist nicht endgültig geklärt. Eine weit verbreitete Theorie ist, dass nur wenige Tumorzellen dem Überwachungssystem des Immunsystems entgehen (Hanahan and Weinberg, 2011). Dafür spricht, dass immunsupprimierte Patienten vermehrt Tumore entwickeln (Vajdic and van Leeuwen, 2009). Da es sich bei diesen jedoch meist um Virus-induzierte Tumore handelt, lässt vermuten, dass es bei diesem Mechanismus eher darum geht, dass immunsupprimierte Patienten von Viren befallene Zellen schlechter abwehren können (Hanahan and Weinberg, 2011). Zudem konnten Tumore aus immunsupprimierten Mäusen schlecht in immunkompetente

übertragen werden, was umgekehrt gut möglich war (Kim et al., 2007). Kim et al. (2007) folgerten aus diesen Beobachtungen, dass vor allem stark immunogene Tumore von einem intakten Immunsystem entdeckt und eliminiert werden, wohingegen es schwach immunogene Tumore schaffen, ihm zu entgehen. Dennoch erfasst diese Theorie nicht die Komplexität der Zusammenhänge, da es beispielsweise auch hoch immunogene Tumore gibt, die es durch verschiedene Mechanismen schaffen, das Immunsystem auszuschalten (Hanahan and Weinberg, 2011).

Auch Grivennikov et al. (2010) veröffentlichen ein Review, welches die Erkenntnisse über das Zusammenspiel zwischen Immunsystem, Entzündung und Krebs zusammenfasste. Einige essentielle Fakten daraus sollen im Folgenden wiedergegeben werden. Grivennikov et al. (2010) führten an, dass das Immunsystem bei ausnahmslos allen Schritten der Tumorentstehung, aber auch bei der Entdeckung und Elimination von Tumorzellen sowie dem Ansprechen auf Tumortherapie beteiligt ist. Es gibt sowohl chronische Entzündungen wie z.B. die chronische Entzündung in der Lunge, die durch das Rauchen entsteht, die zur Tumorentstehung führen können (Punturieri et al., 2009), als auch chronische Entzündungen wie beispielsweise die Psoriasis, die eher mit einem verringerten Krebsrisiko vergesellschaftet sind (Nickoloff et al., 2005). Die Ursachen für dieses Phänomen sind nicht bekannt. Im Tumor selbst verursachen Entzündungszellen sozusagen immer ein tumorförderndes Milieu (Mantovani et al., 2008). Die Genprodukte von Onkogenen wie Ras und Myc führen beispielsweise zu einer Veränderung des Tumormikromilieus, indem sie durch tumorfördernde Chemokine und Zytokine beispielsweise Leukozyten rekrutieren und einen „angiogenic switch“ induzieren (Soucek et al., 2007, Sparmann and Bar-Sagi, 2004). Leukozyten wiederum produzieren Zytokine, Chemokine, Wachstumsfaktoren, Prostaglandine sowie reaktive Stickstoff- und Sauerstoffspezies (Grivennikov et al., 2010). Tumorzellen und Immunzellen beeinflussen und stimulieren sich also wechselseitig. Tumortherapien wie beispielsweise Chemotherapie oder Radiotherapie führen zu einer massiven Aktivierung des Immunsystems, beispielsweise als Folge von Tumornekrose (Zong and Thompson, 2006).

Grivennikov et al. (2010) schrieben über die Immunzellen im Tumor Folgendes: Im Tumor finden sich alle Arten an Immunzellen, jedoch vor allem Tumor assoziierte Makrophagen (TAM) und T-Zellen. TAMs werden meist dem M2 Typ zugeordnet (Biswas et al., 2006), wobei tumorfördernde Zytokine eher dem Zytokinmuster des M1 Typs entsprechen. Der Übergang zwischen den beiden Typen ist jedoch fließend und nicht genetisch determiniert (Grivennikov et al., 2010). Ein hoher Anteil an TAMs ist bei vielen Tumoren, inklusive dem Mammakarzinom, mit einer schlechten Prognose (Bingle et al., 2002) sowie mit Metastasierung und vermehrter Tumorangiogenese assoziiert (Condeelis and Pollard, 2006). Fast alle T-Zellen können je nach Tumorentität tumorfördernd oder -hemmend wirken, was möglicherweise mehr vom Zytokinmuster, das im Tumormilieu herrscht, abhängt als von

der Zusammensetzung und Anzahl der Immunzellen (Grivennikov et al., 2010). Die einzigen T-Zellen, die, soweit man weiß, nie tumorfördernd wirken, sind NK-Zellen (Grivennikov et al., 2010). DeNardo et al. (2009) zeigten in einem Mammakarzinommausmodell, dass CD4+Th2 Zellen TAMs stimulieren, die wiederum Metastasierung und Angiogenese fördern.

Desweiteren sollten Tregs erwähnt werden, über die Nishikawa and Sakaguchi (2010) Folgendes schrieben: Tregs finden sich vermehrt in Tumoren und spielen eine Rolle bei der Suppression der anti-Tumor Antwort des Immunsystems über noch nicht genau verstandene Mechanismen. Sowohl Zell-Zell-Interaktionen mit APCs wie DCs und Effektor T-Zellen als auch die Produktion immunsuppressiver Zytokine scheinen dabei eine Rolle zu spielen.

Neutrophile Granulozyten können je nach Differenzierung und der Anwesenheit von TGF- β einen tumorfördernden oder –hemmenden Effekt haben (Fridlender et al., 2009). Mastzellen scheinen bei einigen Tumorentitäten Angiogenese und Metastasierung zu fördern (de Visser et al., 2006). B-Zellen tragen durch Aktivierung der angeborenen Abwehr, beispielsweise durch Sekretion von Zytokinen, zur Aufrechterhaltung einer chronischen Entzündung im Tumor bei (de Visser et al., 2006).

Desweiteren sollten die sogenannten „myeloid-derived suppressor cells“ (MDSCs) erwähnt werden, die für diese Arbeit eine besonders wichtige Rolle spielen. Sie wurden bei fast allen Tumor Mausmodellen in unterschiedlichem Ausmaß und bei vielen menschlichen Tumorerkrankungen beschrieben (Youn et al., 2008). Diaz-Montero et al. (2009) wiesen bei einer Reihe von Tumorentitäten, auch beim Mammakarzinom, im Vergleich mit gesunden Probanden erhöhte MDSCs im peripheren Blut nach. Auch bei entzündlichen Darmerkrankungen, Verbrennungen, Infektionen, Transplantationen und posttraumatisch ist das Vorkommen von MDSCs beschrieben (Ilkovitch and Lopez, 2009, Youn et al., 2008). MDSCs sind eine Gruppe myeloischer Zellen, die aus Vorläufern von Makrophagen, Granulozyten, dendritischen Zellen und myeloischen Zellen früherer Reifungsstufen besteht (Youn et al., 2008) und in Mäusen als Gr-1+CD11b+ definiert wird (Gabrilovich et al., 2007). MDSCs können nochmals in Zellen mit einem granulozytären Phänotyp, die Ly6G Marker exprimieren und einen monozytären Phänotyp, die Ly6C Marker exprimieren, unterteilt werden (Movahedi et al., 2008, Youn et al., 2008). Movahedi et al. (2008) unterschieden polymorphnukleäre MDSCs (PMN-MDSCs), die vorwiegend Ly6G exprimieren und monozytäre MDSCs (MO-MDSCs), die F4/80, CD115 aber auch den Monozytenmarker Ly6C exprimieren. Youn et al. (2008) beschrieben, dass die MDSCs für die unterdrückte und fehlgeleitete Immunreaktion bei Tumorpatienten eine wichtige Rolle spielen. MDSCs haben eine immunsuppressive Wirkung und hemmen sowohl die Aktivierung von T-Zellen als auch die Antigen-spezifische CD4+ und CD8+ T-Zell Antwort (Talmadge, 2007). Die Mechanismen dahinter sind noch nicht abschließend geklärt, beinhalten aber beispielsweise den L-Arginin Metabolismus und die Freisetzung reaktiver Sauerstoffspezies, Stickstoffmonoxids sowie immunsuppressiver Zytokine (Talmadge, 2007).

Movahedi et al. (2008) schrieben Folgendes: Sowohl der granulozytäre als auch der monozytäre Typ hemmen die Antigen spezifische T-Zellantwort (s.o), erreichen dies aber durch unterschiedliche Mechanismen. Bei beiden Unterformen spielt die Hemmung der T-Zellaktivität durch Interferon- γ (IFN- γ) eine wichtige Rolle. IFN- γ , ein von T-Zellen und NK-Zellen produziertes Zytokin mit zahlreichen Wirkmechanismen im Immunsystem (Boehm et al., 1997), wirkt dabei über STAT1 abhängige oder unabhängige Signalwege. STAT1 ist ein Transkriptionsfaktor, dessen Bildung durch Stimulation durch Interferone gesteigert wird. Die immunsuppressive Wirkung von MO-MDSCs ist im Gegensatz zu der von PMN-MDSCs zumindest zum Teil von STAT1 abhängig. MO-MDSCs exprimieren das STAT1 abhängige Gen iNOS, dessen Effektormolekül NO wiederum eine wichtige immunsuppressive Rolle innehat. Der Mechanismus bei PMN-MDSCs ist ein anderer. Zwar spielt auch hier IFN- γ eine wichtige Rolle, aber über STAT1 unabhängige Signalwege.

Liu et al. (2007) untersuchten die Unterdrückung der Toxizität von NK-Zellen durch MDSCs: Diese scheint durch Zell-Zell Interaktionen vermittelt zu werden. Sie fanden Hinweise darauf, dass MDSCs über die Hemmung der JAK-3 vermittelten Aktivierung von STAT5 die IL-2 vermittelte Zytotoxizität von NK-Zellen supprimieren. Daraus scheint eine verminderte Expression von Perforinen, aber nicht von Granzym B zu folgen. Desweiteren zeigten Li et al. (2009) in einem orthotopen Leberzellkarzinom Mausmodell, dass MDSCs über zellgebundenes TGF- β zu NK-Zell Anergie führen.

MDSCs induzieren bei der Differenzierung der Makrophagen über u.a. IL-10 den immunsuppressiven M2 Phänotyp (Parker et al., 2015, Sinha et al., 2007) und über IL-10 und TGF- β regulatorische T-Zellen (Tregs) (Huang et al., 2006).

Es wird diskutiert, dass MDSCs sich weiter zu Tumor assoziierten Makrophagen (Kusmartsev and Gabrilovich, 2005) oder Endothelzellen differenzieren können und so die Angiogenese im Primärtumor fördern (Yang et al., 2004). Die Bildung von MDSCs wird wohl durch stimulierende Faktoren des Primärtumors wie GM-CSF, VEGF, IL-1 β und Prostaglandine hervorgerufen (Movahedi et al., 2008). Bei nicht tumortragenden Mäusen sind MDSCs nur in geringem Maße im Knochenmark und in der Milz zu finden und haben dort keine immunsuppressive Wirkung (Youn et al., 2008). In verschiedenen Tumormodellen wurden MDSCs vor allem in Milz und Leber, aber auch in den Tumoren selbst, im Blut und im Knochenmark beschrieben (Ilkovitch and Lopez, 2009).

Ilkovitch and Lopez (2009) untersuchten die Rolle der Leber als Ort der MDSC Akkumulation und der daraus resultierenden Immunsuppression bei tumortragenden Mäusen. Sie fanden EMH vor allem perivaskulär aber auch im gesamten Leberparenchym und zeigten, dass es sowohl in metastasierenden wie auch in nicht metastasierenden Tumormodellen zu EMH in der Leber kam und schlossen daraus, dass sie eher eine Folge der Stimulation durch den Primärtumor als Teil der Metastasierung ist. Desweiteren zeigten sie, dass Kupffer'sche Sternzellen durch vermehrte

Expression des T-Zell supprimierenden Oberflächenmoleküls PD-L1 die T-Zell Antwort unterdrücken und dass die hepatischen MDSCs diesen Mechanismus unterstützen. Möglicherweise trifft dies auch für Hepatozyten und sinusoidale Endothelzellen zu. Ob die MDSCs in der Leber gebildet werden oder sich dort nur akkumulieren ist noch nicht abschließend geklärt.

2.4 Das 4T1 Mammakarzinom Mausmodell

Die 4T1 Zelllinie ist eine der wenigen Mammakarzinom Zelllinien im Mausmodell, die in dieselben Organe metastasiert, wie auch humane Mammakarzinome, also in Lunge, Leber, Knochen, Milz und später auch gelegentlich in Gehirn, Herz und Gastrointestinaltrakt und eignet sich deshalb gut als Modell für das fortgeschrittene Mammakarzinom (Tao et al., 2008). Tao et al. (2008) implantierten 4T1 Mammakarzinomzellen in die Brustdrüse weiblicher BALB/c Mäuse. Ca. 3-4 Wochen später kam es zu einem Rückgang des Tumorwachstums, der mit massiver Nekrose und Leukozyteninfiltration im Tumor einherging. In der 2. Woche nach Implantation kam es zur Ausbildung von EMH in der Leber und der Milz der Mäuse und erhöhten Leukozytenzahlen sowie dem Auftreten unreifer myeloischer Zellen im peripheren Blut. Es wird angenommen, dass es sich dabei um MDSCs handelt (Movahedi et al., 2008). Metastasen wurden zuerst in der Lunge (nach ca. 3 Wochen), dann auch in Leber, Niere und Knochen gefunden (nach ca. 3-6 Wochen) (Tao et al., 2008). Es wird spekuliert, dass EMH, Entzündung und Metastasierung eng miteinander verbunden sind und sich gegenseitig verstärken. Beispielsweise war die Nekrose in metastasierenden Tumoren ausgeprägter als in nicht metastasierenden Kontrolltumoren und nach Gabe von entzündungshemmenden Medikamenten wie COX-2 Inhibitoren kam es seltener zu Metastasierung (Tao et al., 2008). Zusätzlich sezernierten die Tumorzellen selbst eine Reihe von Faktoren, die die Bildung von Entzündungszellen stimulierten, diese in den Tumor lockten und sie aktivierten. Dazu gehörten u.a. VEGF-C und Ang2, die Angiogenese, Lymphangiogenese, Tumolvaskularisation und Metastasierung begünstigen und chemotaktisch Makrophagen und Monozyten rekrutieren sowie G-CSF und GM-CSF, die die Hämatopoese stimulieren (Tao et al., 2008). Durch EMH in der Leber entstandene MDSCs können, wie unter 2.3.2 beschrieben das Mikromilieu im Tumor beeinflussen und Metastasierung begünstigen.

2.5 Grundlagen der Strahlentherapie in der Onkologie

In dieser Arbeit wurde EMH in der Leber nach operativer Resektion bzw. Bestrahlung des Primärtumors charakterisiert. Die Strahlentherapie ist essentieller Bestandteil der Therapie des Mammakarzinoms (siehe 2.7.3). Deshalb soll im Folgenden skizziert werden, wie und warum ionisierende Strahlung zum Zelltod und damit zur Abnahme des Tumorumfanges führt. Man unterscheidet direkte Strahlenschäden, z.B. direkte Schädigung des Gewebes und der DNA, von

indirekten Strahlenschäden durch reaktive Sauerstoffspezies als Folge der Interaktion ionisierender Strahlung mit Wassermolekülen (Barcellos-Hoff et al., 2005).

Eriksson and Stigbrand (2010) erläuterten die Mechanismen, durch die Bestrahlung zum Zelltod führt: Bestrahlung führt zur „mitotic catastrophe“, also zum Zelltod während oder als Folge einer fehlerhaften Mitose. Sie tritt oft erst Tage nach Behandlungsbeginn ein. Durch fehlerhafte Zellzykluskontrollpunkte oder Amplifikation von Zentrosomen können Zellen mit anomalen oder mehrfach vorhandenen Zellkernen oder Nukleoli entstehen. Diese Zellen werden mit zeitlicher Verzögerung apoptotisch oder nekrotisch oder treten in die Seneszenz ein. Die eingeschränkte Fähigkeit zur Apoptose, beispielsweise durch Mutationen von p53, gehört wie oben beschrieben zu den „hallmarks of cancer“ und deshalb spielt sie bei soliden Tumoren nur eine untergeordnete Rolle.

Sauer (Biologische Grundlagen der Strahlentherapie von Tumoren, 2010) erläuterte die Grundlagen der Strahlentherapie so: Ziel der Strahlentherapie ist die vollständige Vernichtung des Tumors, was heißt, dass alle Tumorzellen inklusive der Tumorstammzellen abgetötet werden müssten. Dabei entsteht meist der Konflikt, dass die Strahlendosis, die dazu nötig wäre, höher ist als die Strahlendosis, die dem umliegenden gesunden Gewebe zugemutet werden kann. Das Bestrahlungsschema richtet sich also nach der Strahlentoleranz des benachbarten Gewebes. Das Mammakarzinom hat eine mittlere Strahlenempfindlichkeit und benötigt nach Sauer (Mammakarzinom, 2010) zur kurativen Bestrahlung eine Dosis von 60 – 70 Gy. Die Geschwindigkeit, mit der sich Tumore nach Bestrahlung zurückbildeten und das histologische Bild des Tumors im Allgemeinen lassen nur eingeschränkte Aussagen über deren Strahlensensibilität zu. So können auch Tumore, die sich nach Bestrahlung nur langsam zurückbilden oder sogar erst noch wachsen, strahlensensibel sein. Genauso können beispielsweise bestimmte Lymphome, die sonst sehr strahlensensibel sind, strahlenresistent sein oder bestimmte Sarkome, die normalerweise eher strahlenresistent sind, sensibel auf Bestrahlung reagieren. Desweiteren können auch innerhalb einer Tumorentität verschiedene Fälle eine unterschiedliche Strahlensensibilität aufweisen. Sauer (Biologische Grundlagen der Strahlentherapie von Tumoren, 2010) nannte folgende Faktoren, die eine Strahlenresistenz bewirken:

- Großes Tumolvolumen: Im besten Falle werden bei einer Fraktion nur 50 % der klonogenen Tumorzellen inaktiviert
- Hypoxie: Fehlender Sauerstoffverstärkungsfaktor
- Repopulierung: Gesteigerte Proliferation von Tumorzellen nach Tumorteilentfernung und Bestrahlung
- Reparatur: Hohes Reparaturvermögen von Tumoren

- Geringe primäre Tumorphilierung: Nur wenige Zellen befinden sich in einer strahlensensiblen Phase des Zellzyklus
- Intrinsische Strahlenresistenz: Genetisch bedingte Strahlenresistenz

Es gibt eine Reihe von Ansätzen, die darauf abzielen die Strahlensensibilität zu erhöhen und Resistenzbildung zu minimieren. Einer davon ist die fraktionierte Bestrahlung, also die tägliche Applikation geringer Dosen über einen längeren Zeitraum (Bernier et al., 2004).

Bernier et al. (2004) stellten die vier R's der Strahlenbiologie, die man sich bei der fraktionierten Bestrahlung zunutze macht – „repair, redistribution, repopulation, reoxygenation“ –so dar: Das gesunde Gewebe proliferiert langsam und subletale DNA Schäden können repariert werden, bevor die Zellen in den Zellzyklus eintreten. Anders verhalten sich Tumorzellen. Diese teilen sich schnell und durch Bestrahlung entstandene nicht reparierte DNA Schäden können bei der Zellteilung zum Tod der Zelle führen („repair“). Des Weiteren sind Zellen, die sich in der späten S-Phase des Zellzyklus befinden, besonders resistent gegenüber ionisierenden Strahlen. Eine wiederholte Bestrahlung erhöht die Wahrscheinlichkeit, Zellen außerhalb der S-Phase zu treffen („redistribution“). Zusätzlich kommt es nach Radiotherapie zu einer vermehrten Repopulierung („repopulation“) der Tumorzellen. Deshalb sollte zwischen zwei Bestrahlungen nicht zu viel Zeit vergehen. Nach Bestrahlung kommt es außerdem zu einer Reoxygenierung („reoxygenation“) des Gewebes, die wiederum zur besseren Strahlenempfindlichkeit des Tumors beiträgt. Hypoxie führt zu Strahlenresistenz.

Individuelle Einflussgrößen wie z.B. Patientenalter, Allgemeinzustand oder Zusatzerkrankungen haben Einfluss auf das Therapieschema und die maximal mögliche Bestrahlungsdosis (Sauer, Biologische Grundlagen der Strahlentherapie von Tumoren, 2010).

2.6 Die Leber

Lüllmann-Rauch (Leber, 2006) erläutert Aufbau, Funktion und Histologie der Leber wie folgt: Die Leber ist die größte Drüse des Körpers und wiegt um die 1500 g. 80 % des Gewichtes machen die Leberepithelzellen (Hepatozyten) aus. Sie ist ein zentrales Stoffwechselorgan und für die Gallenproduktion und die Produktion zahlreicher anderer Stoffe zuständig. Das Kapillarbett der Leber ist zwischen das Kapillarbett des Magen-Darm-Trakts und die systemische Zirkulation geschaltet, was bedeutet, das im Darm resorbierte Stoffe direkt den Leberzellen zugeführt werden. Außerdem ist sie für die Entgiftung vieler körpereigener und körperfremder Stoffe zuständig. Neben den Hepatozyten sind noch leberspezifische Makrophagen, die sogenannten Kupffer'schen Sternzellen, die Fremdpartikel, Bakterien sowie alte Blutzellen phagozytieren und Ito-Zellen zu nennen. Letztere weisen große Lipidtropfen auf, die Vitamin A enthalten, und werden diskutiert das intralobuläre Bindegewebe zu produzieren. „Das Zentralvenen-Läppchen gilt traditionell als die histologische

Baueinheit der Leber“ (Lüllmann-Rauch, Leber, 2006). Dieses kann man sich als Sechseck vorstellen. An 3 Ecken dieses Sechsecks befinden sich die sogenannten Glisson-Trias: Äste von Pfortader, Arterie und Gallengang. Im Zentrum des Zentralvenenläppchens verläuft die Zentralvene, die Blut aus der Zentralarterie sowie der Pfortader erhält.

2.7 Das Mammakarzinom

Ich habe anhand des unter 2.4 beschriebenen 4T1 Mammakarzinommodells unter anderem den Einfluss von Operation und Bestrahlung auf EMH in der Leber untersucht. Diese sind, wie im Folgenden erläutert, essentielle Bestandteile der Therapie des humanen Mammakarzinoms. Neben weiteren Aspekten der Therapie sollen Epidemiologie, Diagnostik sowie Prognose des Mammakarzinoms erläutert werden.

2.7.1 Epidemiologie des Mammakarzinoms

Da nur einer von 100 an Mammakarzinom Erkrankten männlich ist (Robert-Koch-Institut & Statistisches Bundesamt, 2005), beziehen sich die folgenden Daten ausschließlich auf weibliche Tumorpatientinnen. Das Robert-Koch-Institut (2015) berichtete, dass Brustkrebs mit 30,8% die häufigste Krebsneuerkrankung und mit 17,5% die häufigste krebsbedingte Todesursache bei Frauen in Deutschland ist. Aktuelle Zahlen gehen davon aus, dass jede 8.Frau in ihrem Leben an einem Mammakarzinom erkrankt (Robert-Koch-Institut, 2015). Brustkrebspatientinnen sterben meist nicht aufgrund der lokalen Tumorerkrankung, sondern an systemischen Folgen wie z.B. einer Metastasierung. Da Metastasen noch nicht kurativ behandelt werden können (Sauer, Mammakarzinom, 2010), sind die Vorgänge, die hinter der Entwicklung der Erkrankung stehen, von großem wissenschaftlichen Interesse.

2.7.2 Histologie des Mammakarzinoms

Böcker (2008) schrieb über die Histologie des Mammakarzinoms Folgendes: Beim Mammakarzinom handelt es sich fast ausschließlich um Adenokarzinome, also Tumore, die von Drüsengewebe ausgehen. Histologisch werden je nach Ursprungsgewebe duktale und lobuläre Neoplasien unterschieden. Duktale Neoplasien entstehen aus dem Epithel der Drüsengänge, lobuläre aus dem Epithel der Drüsenläppchen. Invasive Karzinome entstehen aus Vorläuferläsionen, sogenannten Carcinoma in situ (CIS), die die Basalmembran nicht durchbrechen. Das duktale Carcinoma in situ (DCIS) macht 95% aller CIS aus, das lobuläre Carcinoma in situ (LCIS) 5%. Auch Invasiv-duktales Karzinome sind mit 80-90% deutlich häufiger als invasiv-lobuläre Karzinome.

Das DCIS und das LCIS unterscheiden sich in ihrem Ausbreitungsmuster. Das LCIS neigt zu multizentrischem, das DCIS zu multifokalem Wachstum (Leitlinienprogramm Onkologie, 2012).

Das Leitlinienprogramm Onkologie (2012) empfiehlt folgende Zuordnung von Multifokalität und Multizentrität nach Faverly et al. (1994), obwohl keine international einheitliche Definition vorliegt:

- Multifokalität: Auftreten von getrennten Karzinomherden in einem Quadranten bzw. nach Faverly (Faverly, DR et al. 1994) bei einem Abstand zwischen den Herden von weniger als 4 cm.
- Multizentrität: Auftreten von getrennten Karzinomherden in mehr als einem Quadranten bei einem Abstand von mindestens 4 cm zwischen den Herden.

2.7.3 Diagnostik und Therapie des Mammakarzinoms

2.7.3.1 Komplementäre Diagnostik

Die Diagnostik setzt sich aus klinischer Untersuchung, apparativer Diagnostik, operativer Abklärung und pathomorphologischer Beurteilung zusammen (Leitlinienprogramm Onkologie, 2012). Die Verdachtsdiagnose Mammakarzinom kann durch einfache Maßnahmen wie Inspektion, Palpation und Prüfung der Mamillensekretion gestellt werden, wobei die Tastuntersuchung als alleinige Früherkennungsmaßnahme unzureichend ist (Leitlinienprogramm Onkologie, 2012, Sauer, Mammakarzinom, 2010). Bei der Früherkennung des Mammakarzinoms ist als bedeutendste Methode die Mammographie zu nennen. Zu dieser werden seit 2005 alle Frauen zwischen dem 50. Und 69. Lebensjahr alle 24 Monate eingeladen (Robert-Koch-Institut, 2010). Die Mammographie ist die „einzige für die Erkennung von Brustkrebsvorstufen oder frühen Tumorstadien allgemein als wirksam anerkannte Methode“ (Leitlinienprogramm Onkologie, 2012). Die Mammographie kann durch die Sonographie und bei Rezidivverdacht durch die Kontrastmittel-MRT ergänzt werden (Sauer, Mammakarzinom, 2010). Die Kontrastmittel-MRT sollte auch bei erhöhtem familiärem Risiko bei jungen Frauen mit dichtem Brustdrüsenkörper eingesetzt werden (Leitlinienprogramm Onkologie, 2012). Um die Verdachtsdiagnose eines Mammakarzinoms zu bestätigen, ist eine histologische Sicherung nötig. Sie wie folgt durchgeführt (Harbeck and Heywang-Köbrunner, 2011): „Bei tastbaren oder sonographisch darstellbaren Läsionen wird eine Stanzbiopsie unter sonographischer Kontrolle durchgeführt. Nicht palpable Läsionen (z.B. Mikrokalk) sollen unter mammographischer Kontrolle („stereotaktisch“) perkutan biopsiert werden.“ Anhand der Biopsie kann die Bestimmung des histologischen Typs und das Grading, also die „Einstufung des Malignitätsgrades aufgrund histologischer und zytologischer Kriterien“ (Böcker, 2008), erfolgen (Harbeck and Heywang-Köbrunner, 2011). Mittels immunhistochemischer Färbungen werden der Hormonrezeptor- (Progesteron, Östrogen) und der Her2/neu-Status bestimmt (Leitlinienprogramm Onkologie, 2012).

Vor der Therapie ist je nach Größe des Tumors eine Staging Untersuchung mittels Röntgen-Thorax, Skelett Szintigraphie und Leberonographie zum Ausschluss einer Fernmetastasierung sinnvoll (Harbeck and Heywang-Köbrunner, 2011, Leitlinienprogramm Onkologie, 2012). Frauen mit einem erhöhten familiären Risiko, an Brustkrebs zu erkranken, beispielsweise Trägerinnen einer BRCA-1 oder BRCA-2 Mutation, bedürfen einer besonderen Betreuung. Für sie besteht die Möglichkeit einer intensivierten Früherkennung oder einer prophylaktischen bilateralen Mastektomie. Zusätzlich wird ihnen eine prophylaktische beidseitige Salpingo-Oophorektomie um das 40. Lebensjahr als primäre Prävention eines Ovarialkarzinoms empfohlen (Leitlinienprogramm Onkologie, 2012).

2.7.3.2 Primäre Behandlung des Mammakarzinoms

Die Therapie präinvasiver Läsionen ist laut Richtlinien wie folgt empfohlen (Leitlinienprogramm Onkologie, 2012): Das DCIS sollte mit einem Resektionsrand von 2 mm operativ entfernt werden um die Rate an Lokalrezidiven zu senken. Anschließend wird die gesamte betroffene Brust bestrahlt. Eine Axilladisektion muss nicht durchgeführt werden. Der relativ große Sicherheitsabstand beim DCIS ist eine Folge davon, dass es häufig ein multifokales, diskontinuierliches Wachstumsmuster zeigt (siehe 2.7.2). Die Empfehlungen, ob beim LCIS eine Exzision erfolgen sollte, sind uneinheitlich und abhängig von pathologischen Kriterien.

Auch bei der Behandlung des invasiven Mammakarzinoms steht die vollständige Entfernung des Tumors, d.h. eine R0 Resektion mit einem Sicherheitsabstand von einem Millimeter, an erster Stelle (Leitlinienprogramm Onkologie, 2012). Davon ausgenommen sind das inflammatorische Mammakarzinom, lokal weit fortgeschrittene exulzierende Tumore oder Tumore bei Patientinnen, die aufgrund ihres Alters und/oder ihrer Komorbiditäten als inoperabel gelten (Sauer, Mammakarzinom, 2010). Heutzutage machen brusterhaltende Operationen (BET) mit anschließender Bestrahlung den Großteil der Eingriffe aus. Diese sind der modifiziert radikalen Mastektomie, also dem Entfernen der gesamten betroffenen Brust, bezüglich des Überlebens gleichwertig (Leitlinienprogramm Onkologie, 2012). Eine Bestimmung des histologischen Nodalstatus ist unumgänglich und wird mittels „sentinel lymph node“ Biopsie (SLNB), also einer Biopsie des Wächterlymphknotens, durchgeführt (Leitlinienprogramm Onkologie, 2012).

Böcker (2008) schrieb dazu Folgendes: Hinter der Untersuchung des Wächterlymphknotens steht die Annahme, dass Tumorzellen bei der lymphogenen Metastasierung zunächst in einem oder selten auch mehreren der nächsten Lymphknoten-Stationen Absiedlungen bilden, bevor sie weitere Lymphknoten befallen. Zur Identifikation des Wächterlymphknotens wird blauer Farbstoff und/oder ein radioaktiver Tracer in die Tumorumgebung injiziert. Diese verbreiten sich über die Lymphabflusswege und reichern sich im SLN an. Intraoperativ kann er mittels einer Sonde geortet werden, wobei die blaue Farbmarkierung die Auffindung erleichtert.

Ist der Wächterlymphknoten nach histopathologischer Untersuchung tumorfrei, kann auf eine Entnahme weiterer axillärer Lymphknoten verzichtet werden (Leitlinienprogramm Onkologie, 2012, Sauer, Mammakarzinom, 2010). Bei Befall erfolgt eine Axilladissektion, also eine Entnahme und histopathologische Untersuchung von mindestens 10 axillären Lymphknoten, welche mit einer deutlich erhöhten Morbidität wie z.B. einem Lymphödem des Arms einhergeht (Leitlinienprogramm Onkologie, 2012).

Postoperativ kann der Tumor anhand der TNM Klassifikation klassifiziert werden. Diese beinhaltet Tumorgröße (T), Lymphknotenstatus (N) und das Vorliegen von Fernmetastasen (M) (Sauer, Mammakarzinom, 2010). Anschließend an eine BET ist immer eine Bestrahlung der gesamten betroffenen Brust und der angrenzenden Thoraxwand indiziert, um das Lokalrezidivrisiko zu senken (Leitlinienprogramm Onkologie, 2012). Bei Mastektomien ist eine individuelle Entscheidung über eine Bestrahlung möglich (Sauer, Mammakarzinom, 2010), wobei diese auch hier das Lokalrezidivrisiko senkt (Leitlinienprogramm Onkologie, 2012). Eine Bestrahlung des parasternalen und supraklavikulären Lymphabflussgebietes bzw. der axillären Lymphknoten kann je nach Tumorlokalisation und Befall der Lymphknoten nötig sein (Sauer, Mammakarzinom, 2010). Die Dosis in der erhaltenen Brust beträgt ca. 50 Gy bei 5 wöchentlichen Einzeldosen von je 1,8-2 Gy täglich (Leitlinienprogramm Onkologie, 2012).

In der Interdisziplinären S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms (Leitlinienprogramm Onkologie, 2012) werden die Möglichkeiten systemischer Therapien beschrieben: Als systemische Therapien stehen die endokrine Therapie, die Anti-Her2-Antikörpertherapie und die Chemotherapie zur Verfügung. Diese können alleine, in Kombination oder sequentiell angewendet werden. Die endokrine Therapie, also die sogenannte „Antihormontherapie“, wirkt nur bei ER und/oder PR-positiven Tumoren (etwa 70-80% der Tumore) und wird meist in Anschluss an eine Operation für die Dauer von 5 Jahren eingesetzt. Prämenopausale Patientinnen erhalten Tamoxifen, einen selektiven Estrogen Rezeptor Modulator (SERM), postmenopausale Patientinnen Aromatasehemmer. Bei Her2/neu-positivem Mammakarzinom, also bei Karzinomen, die den Wachstumsfaktorrezeptor Her2 überexprimieren, wird eine einjährige adjuvante Therapie mit dem Antikörper Trastuzumab, der sich gegen diesen Rezeptor richtet, empfohlen. Eine postoperative Chemotherapie sollte bei erhöhtem Rezidivrisiko, beispielsweise bei Lymphknotenbefall, aber auch bei hormonrezeptornegativen Tumoren, für die Dauer von 16-24 Wochen durchgeführt werden. Desweiteren kann eine neoadjuvante Chemotherapie, also eine Chemotherapie vor der Operation, eingesetzt werden, wenn dadurch möglicherweise eine brusterhaltende Operation zu erreichen ist. Taxane und Anthrazykline sind die häufigsten Chemotherapeutika, die in der Therapie des Mammakarzinoms Verwendung finden.

2.7.3.3 Rezidivtherapie des Mammakarzinoms

Das Risiko für lokoregionäre Rezidive ist in den ersten beiden Jahren nach Primärtherapie am höchsten (Harbeck and Heywang-Köbrunner, 2011). Im Gegensatz zum metastasierten Mammakarzinom handelt es sich hierbei um eine potentiell kurativ behandelbare Krankheit, wobei die Patientinnen ein deutlich erhöhtes Risiko für eine Fernmetastasierung haben (Harbeck and Heywang-Köbrunner, 2011). Die Therapie der Wahl ist hier nach BET eine sekundäre Mastektomie, nach primärer Mastektomie eine möglichst vollständige Entfernung des lokalen Rezidivs und gegebenenfalls eine zusätzliche Bestrahlung (Leitlinienprogramm Onkologie, 2012).

2.7.3.4 Metastasiertes Mammakarzinom

Bei ca. 5% aller Patientinnen lassen sich bereits bei Erstdiagnose eines Mammakarzinoms Fernmetastasen feststellen, bei weiteren 30% im Verlauf, meist in den ersten 5 Jahren nach Erstdiagnose (Harbeck and Heywang-Köbrunner, 2011). Metastasen treten vor allem in Knochen, Leber, Lunge und ZNS auf (Harbeck and Heywang-Köbrunner, 2011). Das Mammakarzinom kann auch noch nach Jahrzehnten zu Metastasen führen (Hanahan and Weinberg, 2011). Das metastasierte Mammakarzinom ist nicht heilbar und die Therapie orientiert sich an palliativen Gesichtspunkten (Harbeck and Heywang-Köbrunner, 2011).

2.7.4 Prognose des Mammakarzinoms

Sauer (Mammakarzinom, 2010) nannte als prognostische Faktoren Tumorgroße, Lymphknotenstatus, Grading, Hormonrezeptorstatus, Überexpression des membranständigen Rezeptors Her2/neu und Proliferationsaktivität des Karzinoms. Sowohl die 10-Jahres-Rezidivfreiheit als auch die 10-Jahres-Gesamtüberlebensrate belaufen sich auf 20-90% und variieren somit sehr, was daher kommt, dass die Prognose stark von den oben genannten Faktoren abhängt (Sauer, Mammakarzinom, 2010).

3 Material und Methoden

3.1 Material

3.1.1 Verbrauchsmaterialien

Tabelle 1: Verbrauchsmaterialien

Verbrauchsmaterialien	Hersteller
Paraffin	Carl Roth, Karlsruhe, Deutschland
Einbettförmchen	Dia Tec, Hallstadt, Deutschland
Objektträgerkästen	Carl Roth, Karlsruhe, Deutschland
Klingen	Leica Biosystems, Wetzlar, Deutschland
Pinsel	Faber-Castell, Stein, Deutschland
Objektträger: Super Frost Plus	Menzel GmbH & Co. KG,
Menzel-Gläser	Braunschweig, Deutschland
Färbekästen nach Schiefferdecker	Carl Roth, Karlsruhe, Deutschland
Färbekammer Stain Tray	Carl Roth, Karlsruhe, Deutschland
Transferpipetten 3,5 ml	Sarstedt AG & Co., Nümbrecht, Deutschland
Einschlussmittel für die mikroskopische Technik: Roti-Histokitt	Carl Roth, Karlsruhe, Deutschland
Deckgläschen für gefärbte Schnitte 24 mal 40 mm	Menzel GmbH & Co. KG, Braunschweig, Deutschland
Cellstar Tubes 50 ml	Greiner Bio-One, Kremsmünster, Österreich
Cellstar Tubes 14 ml	Greiner Bio-One, Kremsmünster, Österreich
Fettstift: Dako Cytomation Pen	Dako Cytomation, Glostrup, Dänemark
Reagiergefäß 1,5 ml	Sarstedt AG & Co., Nümbrecht, Deutschland
serologische Pipette 5 ml	Greiner Bio-One, Kremsmünster, Österreich
serologische Pipette 10 ml	Greiner Bio-One, Kremsmünster, Österreich
serologische Pipette 25 ml	Greiner Bio-One, Kremsmünster, Österreich
Pipettenspitzen: ep T.I.P.S. Standard 20-300 µl	Eppendorf AG, Hamburg, Deutschland
Pipettenspitzen: ep T.I.P.S. Standard 0,5-20 µl	Eppendorf AG, Hamburg, Deutschland
Pipettenspitzen 1000 µl	Sarstedt AG & Co., Nümbrecht, Deutschland
Pipettenspitzen 200 µl	Sarstedt AG & Co., Nümbrecht, Deutschland

3.1.2 Geräte und Software

Tabelle 2: Geräte und Software

Geräte	Hersteller
Einbettstation Histocentre 2 mit Kühlplatte	Shandon Southern Products Ltd., Astmoor, England
Mikrotom	Microm HM 400, Walldorf, Deutschland
Wasserbad	
Heizschrank Heraeus T6060	Heraeus Instruments, Hanau, Deutschland
Vortex-Schüttler: Vortexgenie 2	Scientific Industries Inc., NY, USA
Mikrowelle	Sharp Electronics, Hamburg, Germany
Waage Ca Kern EW	Kern & Sohn GmbH, Balingen, Deutschland
Pipettierhilfe: Pipetboy 2	Integra Biosciences AG, Fernwald, Deutschland
Pipette: 10-100 µl Pipette	Gilson, Middleton, WI, USA
Pipette: 20-200 µl Pipette	Gilson, Middleton, WI, USA
Pipette: 100-1000 µl Pipette	Gilson, Middleton, WI, USA
Kühlschrank Profi-line	Liebherr, Biberach, Deutschland
Mikroskop Axio Imager.M2	Zeiss, Jena, Deutschland
Software Aperio ImageScope Version 11.2.0.780	Informer Technologies, Inc.
Software Definiens	Definiens AG, München, Deutschland
Aperio ScanScope CS2	Leica Biosystems Nussloch GmbH, Nussloch, Deutschland

3.1.3 Chemikalien und Reagenzien, Lösungen

Tabelle 3: Chemikalien und Reagenzien

Chemikalien und Reagenzien	Hersteller	Lagerung
Ethanol 99,8 %	Merck, Darmstadt, Deutschland	RT
saures Hämalaun nach Meyer	Merck, Darmstadt, Deutschland	RT
Eosin	Merck, Darmstadt, Deutschland	RT
Xylol	Merck, Darmstadt, Deutschland	RT
destilliertes H ₂ O	Milli-Q Reinstwasseranlage, Millipore Corp., Billerica, MA., USA	RT
Phosphat Buffered Saline	Gibco Invitrogen GmbH, Karlsruhe, Deutschland	RT

Tabelle 4: Lösungen

Lösungen	Hersteller	Zusammensetzung	Lagerung
Natriumazid-Lösung	Carl Roth, Karlsruhe, Deutschland (Natriumazid); Sigma-Aldrich, St. Louis, MO, USA (BRI J 30 %; Wasserstoffperoxid); Merck, Darmstadt, Deutschland (Tris (hydromethyl)-aminomethan; HCl 37 %)	0,2 g Natriumazid Lösung auf 200 ml Autopuffer 1x, 4-6 ml Wasserstoffperoxid 30% dazugeben, vor jedem Gebrauch 200 µl auf 200 ml Wasserstoffperoxid zugeben (endvolumenabhängig)	RT
Autopuffer 10 x	Merck, Darmstadt, Deutschland (Tris (hydromethyl)-aminomethan; HCl 37%); Sigma-Aldrich, St. Louis, MO, USA (BRI J 30 %)	30,285g Tris (hydromethyl)-aminomethan in 450 ml H ₂ O millipore lösen mit reiner HCl37% (ca. 17ml) auf pH7,5 einstellen 12,5ml BRIJ 35 (30%) zugeben auf 500ml auffüllen	2-8°
Tris-Puffer 10 x	Merck, Darmstadt, Deutschland	30,285g Tris (hydromethyl)-aminomethan in 450 ml H ₂ O millipore lösen mit reiner HCl37% (ca. 17ml) auf pH7,5 einstellen auf 500ml auffüllen	2-8°
BSA 10 %	Carl Roth GmbH + Co., Karlsruhe, Deutschland	5g BSA(Fraktion V, nicht Alc.-präzipitiert) in 50ml PBS lösen, ein wenig Natriumazid zugeben	2-8°
Dako Target Retrieval Solution 10 * concentrate	DAKO A/S, Kopenhagen, Dänemark		2-8°
DAB Gebrauchslösung	DAKO A/S, Kopenhagen, Dänemark		2-8°

3.1.4 Antikörper

Tabelle 5: Primärantikörper

Primär Antikörper	Hersteller	Klon	Ursprung
Anti-CD 3	Zymed	SP7	rabbit
Anti-F480	Acris	BM8	rat
Anti-CD 49b	Bio Legend	DX-5	rat
Anti-CD 45	Bio Legend	30-F11	rat
Anti-Ly6G/ Ly6C (Gr-1)	Bio Legend	RB6-8C5	rat

Tabelle 6: Sekundärantikörper

Sekundär Antikörper	Firma	Klon	Ursprung
Anti-rat	Dako	P0450	rabbit
Anti-rabbit	Dako	P0217	swine

3.2 Methoden

3.2.1 Mausexperiment

Je 250.000 4T1 Mammakarzinomzellen wurden am Tag 0 in die Brustdrüse von 18 weiblichen BALB/c Mäusen implantiert. Nach 2 Wochen Tumorwachstum wurden die 18 Mäuse in drei Gruppen mit je 6 Mäusen unterteilt: Die erste Gruppe diente als Kontrollgruppe. In der zweiten Gruppe wurde der Tumor nach 2 Wochen Wachstumsphase einmalig mit einer Dosis von 30 Gy bestrahlt. In der dritten Gruppe wurde der Tumor bereits nach 2 Wochen Wachstumsphase operativ entfernt und fixiert. 4 Wochen nach der Implantation der Tumorzellen wurden die Mäuse getötet. Tumore, rechter Leberlappen und Lunge wurden entnommen und fixiert. Das Volumen der Tumore wurde bestimmt. Einer unbehandelten Maus ohne Tumor wurde ebenfalls der rechte Leberlappen entnommen. In Tabelle 7 wird die Benennung der Mäuse erläutert. Das Formalin-fixed paraffin embedded (FFPE) Material von Tumor und Leber wurde mir zur Verfügung gestellt.

Tabelle 7: Benennung der Mäuse

Kontrollmäuse	K1 - K6
Bestrahlte Mäuse	30 Gy 1- 30 Gy 4
Operierte Mäuse	Op 1 - Op 6
Unbehandelte Maus	M1

3.2.2 Einbetten der Tumore, Lebern und Lungen

In einer Einbettstation wurden die fixierten Gewebeproben in Einbettförmchen mit 70° C heißem flüssigem Paraffin übergossen und auf einer Kühlplatte bei -10° C ausgehärtet.

3.2.3 Anfertigen von histologischen Schnitten

Sowohl von den Leberlappen als auch von den Primärtumoren wurden histologische Schnitte angefertigt. Vor dem Schneiden mit dem Rotationsmikrotom wurden die Paraffinblöcke mit den Gewebeproben für einige Minuten auf einer Kühlplatte bei -10° C gekühlt. Das dient dazu das

Paraffin zu erhärten und das Schneiden zu erleichtern. Anschließend wurde der entsprechende Paraffinblock in das Mikrotom eingespannt. Es wurden Ultradünnschnitte von 3 µm Dicke angefertigt. Diese wurden mit einem Pinsel zum Strecken in ein Wasserbad mit destilliertem Wasser von 40° C überführt und anschließend auf Objektträger aufgezogen. Die Objektträger mit den Schnitten wurden dann über Nacht bei 65° C in einem Heizschrank getrocknet. Von den 16 Lebern wurden im Abstand von 50 µm je 10 Folgeschnitte angefertigt. Die Lebern der Mäuse K1, K2 und Op 1 wurden vollständig untersucht. Von den restlichen Lebern wurden je zwischen 70 und 140 Schnitte angefertigt. Von M1 wurden ca. 20 Schnitte erstellt. Von den 16 Primärtumoren wurden zunächst ca. 12 Folgeschnitte angefertigt. Dann wurden ca. 300 µm verworfen und erneut 2 Folgeschnitte erstellt. Dieser Vorgang wurde noch einmal wiederholt. Bei den Tumoren der operierten Mäuse wurde aufgrund ihres geringen Volumens weniger verworfen.

3.2.4 Hämatoxylin-Eosin Färbung

Sowohl von den Leberlappen als auch von den Primärtumoren wurden histologische Schnitte mit einer H.E.-Färbung analog dem in Tabelle 8 dargestellten Protokoll gefärbt. Zunächst wurden die Schnitte in Xylol deparaffiniert. Anschließend wurden die Gewebe über eine absteigende Ethanolreihe in Wasser überführt. Dann wurden die Schnitte mit Hämatoxylin gefärbt. Als nächstes wurden die Gewebe mit warmem laufendem Leitungswasser im Waschbecken gebläut. Der Kalk im Leitungswasser führt dazu, dass die Färbung fixiert wird. Danach wurden die Schnitte in Eosin gefärbt und wieder in destilliertes Wasser gegeben. Anschließend wurden die Objektträger in eine aufsteigende Ethanolreihe überführt und am Ende in Xylol eingelegt. Zuletzt wurden die Schnitte mit Eukitt eingedeckelt.

Tabelle 8: Protokoll Hämatoxylin-Eosin Färbung

Schritt	Lösungsmittel	Zeit	Bemerkungen
1	Xylol I	10'	
2	Xylol II	10'	
3	EtOH abs. I	5'	
4	EtOH abs. II	5'	
5	EtOH 96%	5'	
6	EtOH 70%	5'	
7	EtOH 50%	5'	
8	H ₂ O _{dest}	5'	
9	Hämatoxylin	1'	saures Hämatoxylin nach Ehrlich oder Hämalaun nach Mayer
10	warmes Leitungs-H ₂ O	10'	
11	Eosin	3'	plus 2 Tropfen Eisessig pro 200 ml Ansatz
12	H ₂ O _{dest}	5'	
13	EtOH 50%	5'	
14	EtOH 70%	5'	
15	EtOH 96%	5'	
16	EtOH abs III	5'	
17	Xylol III	5'	
18			mit geeignetem Eindeckmedium (Eukitt) eindecken

3.2.5 Immunhistochemie

Sowohl von den Leberlappen als auch von den Primärtumoren wurden histologische Schnitte analog dem in Tabelle 9 dargestellten Protokoll immunhistochemisch mit den in Tabelle 10 genannten Antikörpern gefärbt. Zunächst wurden die Schnitte in Xylol deparaffiniert. Anschließend wurden die Gewebe über eine absteigende Ethanolreihe in Wasser überführt. Dann wurden die Objektträger zur Antigenrenaturierung in Target Retrieval Solution (pH 6) in einer Mikrowelle gekocht. Um die Lösung von den Schnitten zu waschen, wurden die Objektträger in destilliertes Wasser gegeben. Dann wurden sie in eine Färbekammer gelegt und die Gewebestücke mit einem Fettstift umrandet. Das dient dazu, dass die folgenden Lösungen und Antikörper auf dem Gewebe verbleiben, es feucht halten und nicht von dem Objektträger herunterfließen. Als nächstes wurde mit einer Einmalpipette Natriumazid Lösung auf die Schnitte gegeben. Durch das Natriumazid wird die intrazelluläre Peroxidase inaktiviert, was falsch-positive Farbreaktionen durch 3, 3'-Diaminobenzidin (DAB) verhindert. Dann wurden die Objektträger zunächst in Tris-Puffer und anschließend in Detergenzhaltigen Auto-Puffer überführt. Ersterer dient dazu, die Natriumazid Lösung von dem Gewebe zu

waschen, letzterer verbessert die Aufnahmefähigkeit des Gewebes für die nachfolgenden Antikörper. Dann wurde mit einer Einmalpipette 10 %ige Bovine Serum Albumin-Lösung (BSA-Lösung) in Phosphat Buffered Saline (PBS) auf das Gewebe gegeben, um unspezifische Bindungen zu vermeiden. Nach einer halben Stunde ließ ich diese durch Schräghalten der Schnitte abtropfen und gab mit einer Pipette den entsprechend verdünnten Primärantikörper auf die Schnitte. Anschließend wurde die Färbekammer entweder über Nacht kühl gestellt oder für zwei Stunden bei Raumtemperatur belassen. Dann wurden die Schnitte zum Abwaschen der Primärantikörper in Autopuffer überführt, bevor der entsprechend verdünnte Sekundärantikörper auf die Schnitte gegeben wurde (siehe Tabelle 11). War der Ursprung des Primärantikörpers beispielsweise die Spezies „Ratte“, so wurde als Sekundärantikörper ein anti-Ratten-Immunglobulin verwendet. Nachdem die Schnitte wiederum mit Autopuffer gereinigt worden waren, wurde mit einer Pipette in der Feuchtekammer DAB-Gebrauchslösung auf die Gewebestücke gegeben. Diese dient der Färbung der Sekundärantikörper. Danach wurden die Objektträger erst in Leitungswasser, dann in destilliertes Wasser gegeben. Nun wurden die Schnitte mit Hämatoxylin gefärbt und die Objektträger anschließend mit warmem laufendem Leitungswasser im Waschbecken überspült. Danach wurden die Objektträger in eine aufsteigende Ethanolreihe überführt und anschließend in Xylol eingelegt. Zuletzt wurden die Schnitte mit Eukitt eingedeckelt.

Tabelle 9: Protokoll Immunhistochemie

Schritt	Lösungsmittel	Zeit	Bemerkungen
1	Xylol A I	30'	abh. vom Gewebe
2	Xylol A II	30'	abh. vom Gewebe
3	EtOH abs. I	10'	
4	EtOH abs.II	10'	
5	EtOH 96%I	5'	
6	EtOH 96% II	5'	
7	EtOH 70%	5'	
8	EtOH 50%	5'	
9	H ₂ O _{dest}	2'	
10	Target Retrieval Solution (1:10)	20'	In Mikrowelle bei 630 Watt kochen
11	Target Retrieval Solution	15'	abkühlen lassen
12	H ₂ O _{dest} (od. PBS)	5'	
13	Fettstift-Markierungen		
14	Natriumazid Lösung	30'	
15	Tris-Puffer	10'	
16	Auto-Puffer	5'	
17	BSA 10% in PBS	30'	in Färbekammer
18	Primär-Antikörper in Feuchtekammer	2h bei RT oder über Nacht bei 4°	in Färbekammer
19	Auto-Puffer	20'	
20	sekundär-Antikörper	2h in RT	in Färbekammer
21	Auto-Puffer	20'	
22	DAB-Gebrauchslösung	7'	in Färbekammer
23	Leitungs-H ₂ O	3'	
24	H ₂ O _{dest}	3'	
25	Hämatoxylin	30''	
26	warmes Leitungs-H ₂ O	10'	
27	EtOH 30%	5'	
28	EtOH 50%	5'	
29	EtOH 70%	5'	
30	EtOH 96% II	5'	
31	EtOH 96% I	5'	
32	EtOH abs.II	5'	
33	EtOH abs.I	5'	
34	Xylol B I	10'	
35	Xylol B II	10'	
36			mit geeignetem Eindeckmedium (Eukitt) eindeckeln

Tabelle 10: Primärantikörper mit Verdünnung und Funktion

Primär Antikörper	Ursprung	Verdünnung 1% BSA in PBS	Verwendung
Anti-CD 3	rabbit	1 : 100	T-Zell-Marker
Anti-F480	rat	1 : 100	Makrophagen-Marker
Anti-CD 49b	rat	1 : 400	NK-Zell-Marker
Anti-CD 45	rat	1 : 1000	Pan-Leukozytenmarker
Anti-Ly6C/G	rat	1 : 1000	Neutrophilen-Marker

Tabelle 11: Sekundärantikörper mit Verdünnung

Sekundär Antikörper	Ursprung	Verdünnung 1%BSA in PBS
anti rat	rabbit	1 : 100
anti rabbit	swine	1 : 100

3.2.6 Mikroskopieren der gefärbten Schnitte

Unter einem Lichtmikroskop wurden die H.E. und immunhistochemisch gefärbten Schnitte der Lebern und Tumore betrachtet.

3.2.7 Einscannen der gefärbten Schnitte mit Aperio ScanScope

Zur Erleichterung der histologischen Begutachtung sowie zur digitalen Dokumentation wurden die histologisch präparierten Gewebe eingescannt. Es entstanden pro Leber einmal 6 eingescannte Folgeschnitte (ein H.E. Schnitt und 5 IHC Schnitte, die mit den in Tabelle 10 genannten Antikörpern gefärbt waren) und einmal 2 eingescannte Folgeschnitte (ein H.E. Schnitt und ein immunhistochemisch CD45 spezifisch gefärbter Schnitt). Das gilt auch für die Leber der unbehandelten Maus. Von den Primärtumoren wurden H.E. und IHC gefärbte Schnitte eingescannt. Der Aperio Scan Scope Scanner dient dazu histologische Schnitte digital speichern, betrachten und bearbeiten zu können. Zunächst wurde am Rechner das Programm Aperio ScanScope geöffnet. Dann wurden je 5 Objektträger in die Objektträgerhaltevorrückung eingeordnet. Erst wurden Übersichtsaufnahmen der gesamten Objektträger erstellt, anhand derer die relevanten Bereiche für hochauflösende Aufnahmen definiert wurden. Nach der Benennung der Datei sowie der Durchführung eines Weißabgleichs bei allen 5 Schnitten wurde der Scanvorgang gestartet.

3.2.8 Auswerten der Scans der H.E. gefärbten Schnitte mit ImageScope

Je Leber wurden zwei Scans von H.E. gefärbten Schnitten mit Aperio ImageScope bearbeitet. Da eine manuelle Auswertung des gesamten Scans einen zu hohen zeitlichen Aufwand bedeutet hätte, wurden, um einen objektiven Überblick zu erlangen, zwischen 4 und 5 Rechtecke mit einer mittleren Fläche von ca. 3 mm² gleichmäßig über den Scan verteilt betrachtet. In diesen Rechtecken wurden alle Bereiche, die als relevant erachtet und als Läsionen bezeichnet wurden, manuell umrandet. Läsionen, die über die Grenze des Rechtecks hinausgingen, wurden vollständig umrandet. Abbildung 2 zeigt, wie vorgegangen wurde. Nun wurden die Flächen der markierten Rechtecke und der Läsionen über eine Textdatei in eine Excel Datei exportiert. Die sehr großen Läsionen (>10.00 µm²) wurden im gesamten Leberscan erfasst (siehe Abbildung 3). Ergebnis dieser Analyse waren nun einerseits die Flächen der einzelnen Rechtecke und die Flächen der entsprechenden umrandeten Läsionen in den Rechtecken und andererseits die Fläche des gesamten Leberscans und die Flächen der großen Läsionen im gesamten Leberscan. Um eine möglichst große Objektivität zu erlangen und sicherzustellen, dass durch die gewählten Ausschnitt kein Bias entstand, wurde die Zählung an je einem H.E. Scan jeder Gruppe mit Ellipsen statt Rechtecken wiederholt (siehe Abbildung 4). Auch diese Daten wurden in eine Exceldatei exportiert und ausgewertet.

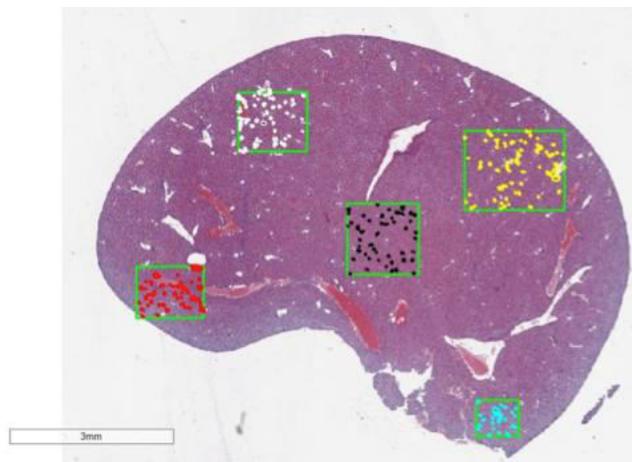


Abbildung 2: ImageScope: Erfassen aller Läsionen in den ausgewählten Bereichen. Über den H.E. gefärbten Scan eines Leberschnitts verteilte Rechtecke, in denen alle Läsionen markiert wurden.

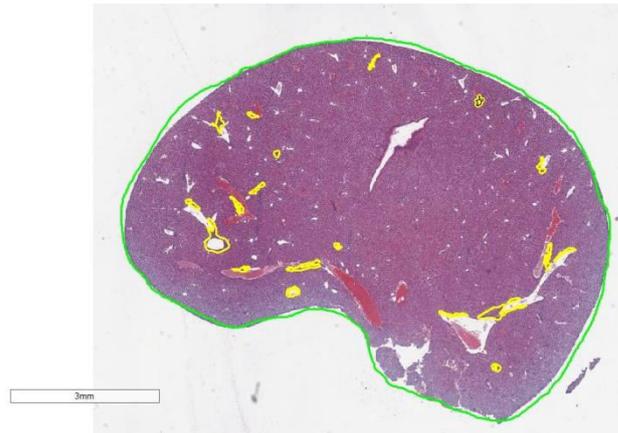


Abbildung 3: ImageScope: Erfassen aller großen Läsionen pro Leberschnitt. Umrandeter H.E. gefärbter Scan eines Leberschnitts, in dem alle großen Läsionen markiert wurden.

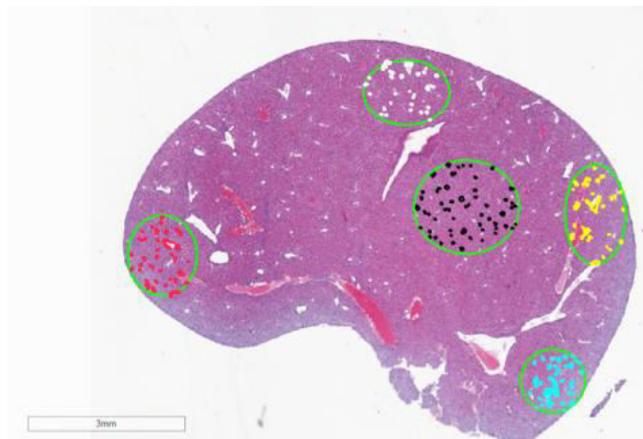


Abbildung 4: ImageScope, Ergebniskontrolle. Über den H.E. gefärbten Scan eines Leberschnitts verteilte Ellipsen, in denen alle Läsionen markiert wurden.

3.2.9 Auswerten der Daten aus ImageScope in Excel

In Excel wurden die vorangegangenen Quantifizierungen auf Fläche normiert. Zunächst entstand je eine Excel Tabelle pro in ImageScope bearbeiteten Scan mit den in Tabelle 12 dargestellten Parametern. Es wurde eine Unterteilung in kleine, mittlere und große Läsionen getroffen (siehe Tabelle 13). Die Aufteilung in Rechtecke wurde nun nicht weiter verwendet, sie diente vor allem der Beurteilung der Vergleichbarkeit der verschiedenen Bereiche. Nun wurden für jeden Scan die in Tabelle 14 dargestellten Parameter berechnet und dann Mittelwert und Standardabweichung aus beiden Scans der einzelnen Lebern der Versuchstiere ermittelt.

Tabelle 12: Ermittelte Daten aus den H.E. gefärbten Leberscans aus ImageScope

Parameter	Bemerkung
Gesamtfläche in mm ²	Summe aller Rechteckflächen
Läsionsfläche in mm ²	Summe aller Läsionsflächen
Schnittfläche in mm ²	Fläche des gesamten eingescannten Schnittes
Anzahl der Läsionen	Summe aller gezählten Läsionen
Anzahl der kleinen Läsionen	Summe aller gezählten kleinen Läsionen
Anzahl der mittleren Läsionen	Summe aller gezählten mittleren Läsionen
Anzahl der großen Läsionen	Summe aller gezählten großen Läsionen

Tabelle 13: Einteilung der Läsionen nach Größe

Bezeichnung der Läsionen	Fläche in µm²
klein	500 – 1.999
mittel	2.000-10.000
groß	>10.000

Tabelle 14: Berechnung der Parameter aus den H.E. gefärbten Leberscans aus ImageScope

Parameter	Bemerkung
Läsionen/mm ²	Summe aller Läsionen aus den Rechtecken geteilt durch die Summe aller Rechteckflächen
kleine Läsionen/mm ²	Summe aller kleinen Läsionen aus den Rechtecken geteilt durch die Summe aller Rechteckflächen
mittlere Läsionen/mm ²	Summe aller mittleren Läsionen aus den Rechtecken durch die Summe aller Rechteckflächen
große Läsionen/mm ²	Summe aller großen Läsionen aus dem Leberschnitt geteilt durch die Schnittfläche
Anteil der Läsionsfläche an der Summe der Rechteckflächen in %	Prozentualer Anteil der Läsionsfläche eines Schnittes an der Gesamtfläche des Leberschnittes

3.2.10 Auswerten der Scans der immunhistochemisch gefärbten Schnitte mit Definiens

Pro Leber wurden zwei IHC Scans mit Definiens Tissue Studio bearbeitet. Ich verwendete die CD45 spezifisch gefärbten Scans, da fast alle Läsionen CD45 positiv sind. Diese waren Nachbarschnitte der H.E. Scans. Definiens Tissue Studio ist eine digitale Bildanalysesoftware, mit der IHC Scans analysiert werden können. Das Programm macht es möglich, bestimmte Biomarker und/oder besondere histologische Merkmale in einem vordefinierten Bereich der Scans zu entdecken und zu quantifizieren. Dabei kann man bestimmte Zielgrößen festlegen. Dazu zählen beispielsweise die Intensität der IHC-Färbung und die Größe der Bereiche. Ziel war es, die Zahl der Läsionen pro Leberfläche sowie die Größe der Läsionen und den Anteil der Läsionsfläche an der Gesamtleberfläche herauszufinden. Zunächst wurden einzelne Dateien oder ganze Ordner importiert und die „Workspace“ benannt. Dann wurden „Regions of interest“ (ROIs) definiert (ROI 1: „Leberfläche; ROI 2: „No“). Nun wurde die Leberfläche des geladenen Bildes umfahren, was nun unsere „Region of interest“, also der Bereich, der analysiert werden sollte, war. Größere Artefakte konnten auf die gleiche Weise umrandet, als ROI 2 bezeichnet und somit ausgeschlossen werden. Anschließend wurden die Grundeinstellungen zur Analyse des Scans, die „Marker Detection Settings“, festgelegt. Als „Threshold Hematoxylin“ wurde 0.08, als „Threshold Marker“ 0.14, als „Minimum Area in μm^2 “ 100 gewählt. „Threshold Hematoxylin“ legt den Grenzwert des Blauwertes fest und bestimmt somit ab welchem Blauwert eine Zelle als solche erkannt wird. So werden beispielsweise weiße Bereiche erkannt und nicht fälschlicherweise dem Gewebe zugeordnet. Wir wählten diesen relativ niedrig um das Gewebe vollständig zu erfassen. „Threshold Marker“ hingegen definiert den Grenzwert des Braunwertes, ab dem ein Bereich als immunhistochemisch positiv gewertet wird. Die „Minimum Area in μm^2 “ legt die minimale Größe fest, ab der ein Bereich als positiv gewertet wird. Bei diesem Wert hat man allerdings nur eine relativ geringe Spielbreite. Anschließend wurden anhand eines kleinen Ausschnittes des Scans die bisherigen Einstellungen validiert. Über „Marker Area Classification“ wurde eine Reihe weiterer Parameter definiert um unspezifisch positive Bereiche noch einmal heraus zu filtern. Bei „Classification“ wurde als „Intensity Measure“ → „Optical Density“ gewählt. Unter „Thresh Low/Medium“ und „Thresh Medium/High“ wurde jeweils der größte Wert verwendet. Unter „Exclusion“ wurden Eigenschaften definiert, die aus der Analyse ausgeschlossen werden sollten. Es wurden alle positiven Bereiche, die kleiner oder gleich $500 \mu\text{m}^2$ sind, ausgeschlossen. Desweiteren wurde eine „Brown Chromogen Intensity“ zwischen <0.5 und <0.8 bei den verschiedenen Scans angegeben, je nachdem, wie viele Artefakte es gab und wie intensiv die Färbung war. Dadurch wurden Bereiche ausgeschlossen, die nicht stark genug braun gefärbt waren. Nun konnte eine erneute Validierung stattfinden und bestimmte Bereiche, die kritisch erschienen, wurden erneut genauer betrachtet. Anschließend wurde festgelegt, welche Parameter und Resultate exportiert werden sollten. Zuletzt wurde die „Solution“ gespeichert. So konnte man sie, wollte man weitere

Scans analysieren, wieder verwenden. Nun wurde die Analyse des Scans gestartet. Als Ergebnis erhielt man unter anderem Originalausschnitte mit dem zugehörigen „Overlay“, also mit den Bereichen, die Definiens als positiv gewertet hatte (siehe Abbildung 5 und Abbildung 6). Desweiteren erhielt man Exceltabellen, die den Anteil der Läsionsfläche an der Gesamtfläche enthielten sowie Exceltabellen, in denen alle Läsionen mit ihrer Größe aufgelistet waren. Auch die Größe der Leberfläche wurde angegeben.

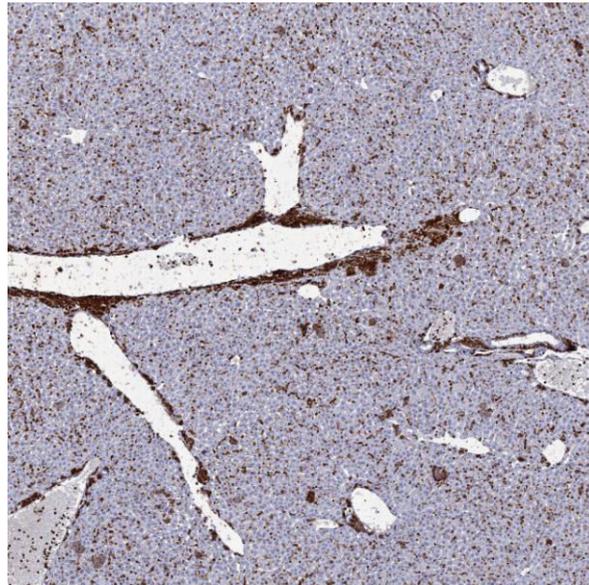


Abbildung 5: Definiens – Beispiel aus Snapshots, Original. Ausschnitt aus einem Scan eines CD45 gefärbten Leberschnitts – erstellt durch Definiens. Zu sehen sind zahlreiche CD45 positive Bereiche, vor allem perivaskulär.

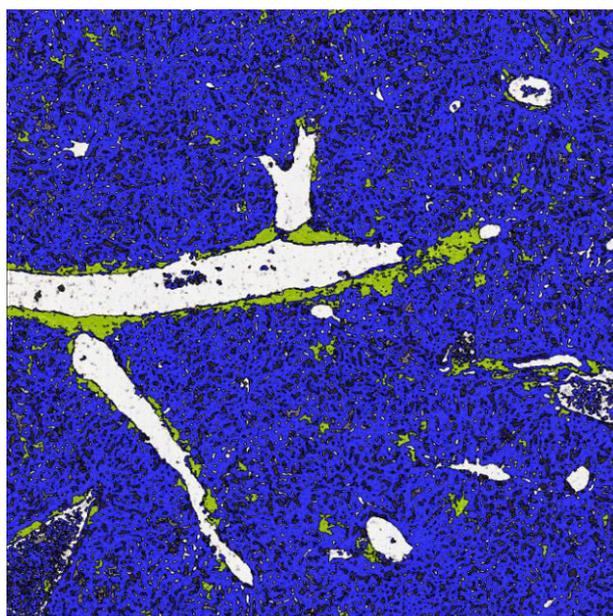


Abbildung 6: Definiens – Beispiel aus Snapshots, Overlay. Zu Abbildung 5 korrespondierender Overlay. Gelb markiert sind alle Bereiche, die von Definiens als positiv erkannt wurden.

3.2.11 Auswerten der Daten aus Definiens in Excel

Für jeden in Definiens bearbeiteten Scan wurde eine Exceltabelle angelegt. Dort wurden folgende Werte eingetragen:

- Leberfläche in mm² (entspricht der ROI „Leberfläche“)
- „Marker Area“ in %
- Alle Läsionen der ROIs

Die Läsionen wurden ihrer Größe nach sortiert und gezählt. Genau wie bei der Analyse der Daten aus ImageScope wurden die Läsionen je nach Größe in 3 Gruppen aufgeteilt (siehe Tabelle 13). Jede Excel Tabelle enthielt die in Tabelle 15 dargestellten Parameter, von denen Mittelwert und Standardabweichung der beiden Scans einer Leber ermittelt wurden.

Tabelle 15: Berechnung der Parameter aus den IHC gefärbten Leberscans aus Definiens

Parameter	Definition
Läsionen/mm ²	Summe aller Läsionen aus der ROI Leberfläche geteilt durch die Leberfläche
kleine Läsionen/mm ²	Summe aller kleinen Läsionen aus der ROI Leberfläche geteilt durch die Leberfläche
mittlere Läsionen/mm ²	Summe aller mittleren Läsionen aus der ROI Leberfläche geteilt durch die Leberfläche
große Läsionen/mm ²	Summe aller großen Läsionen aus der ROI Leberfläche geteilt durch die Leberfläche
% „Marker Area“	Prozentualer Anteil der Läsionsfläche an der Gesamtfläche des Leberschnittes

3.2.12 Statistische und graphische Auswertung der Daten in SPSS

Die statistischen Auswertungen wurden nach Beratung durch das Institut für medizinische Statistik und Epidemiologie, Klinikum Rechts der Isar der TU München, durchgeführt. In SPSS wurden die Daten aus Excel graphisch dargestellt und statistisch ausgewertet. Zunächst wurden sowohl die Primärtumorvolumina als auch die Ergebnisse aus ImageScope und Definiens in SPSS eingelesen. Dabei wurde der Mittelwert aus den je zwei pro Leber und Verfahren analysierten Scans als beste Schätzung für den wahren Wert verwendet. Von den Primärtumorvolumina jeder der drei Gruppen

(Kontrolle, Bestrahlung, Operation) wurde eine Deskriptive Statistik angefertigt. Anschließend wurden die Primärtumorvolumina der verschiedenen Gruppen in einem Boxplot Diagramm gegenüber gestellt. Aufgrund der kleinen Fallzahl wurde keine Normalverteilung angenommen. Mittels Kruskal-Wallis-Test für unabhängige Stichproben wurde untersucht, ob es statistisch signifikante Unterschiede zwischen den Medianen der Primärtumorvolumina der verschiedenen Gruppen gibt. War dies der Fall, wurde der Mann-Whitney-U-Test für unabhängige Stichproben angewendet, um herauszufinden, zwischen welchen Gruppen statistisch signifikante Unterschiede vorlagen. Genau so wurde bei den Ergebnissen aus ImageScope und Definiens vorgegangen. Dabei wurden zunächst jeweils die Mittelwerte der EMH Parameter aus zwei H.E. bzw. CD45 spezifisch gefärbten IHC Scans derselben Leber verwendet. Aus den Mittelwerten der EMH Parameter aller Mäuse einer Gruppe wurden Boxplot Diagramme erstellt und anschließend untersucht ob statistisch signifikante Unterschiede zwischen den Gruppen bestehen. Um die Daten aus Ellipse und Rechteck sowie aus ImageScope und Definiens vergleichen zu können, wurden Bland-Altman-Plots erstellt. Der Bland-Altman-Plot ist ein sinnvolles graphisches Verfahren zur Darstellung der Übereinstimmung unterschiedlicher Messmethoden, also des „agreements“ (Grouven et al., 2007). Dabei wurden die Mittelwerte aus beiden Messmethoden gegen ihre Differenz aufgetragen. Das Diagramm enthält zusätzlich die Verzerrung, also den Mittelwert über alle Differenzen, sowie die Übereinstimmungsgrenzen $d \pm 2 \times s$ wobei d den Mittelwert der Differenzen und s die Standardabweichung der Differenzen bezeichnet (obere und untere Linie). Das Diagramm veranschaulicht die Streuung der Differenzen der einzelnen Messwertpaare. Der Mittelwert eines Messwertpaares wird dabei als der bestmögliche Schätzer für den wahren Wert angesehen. 95% der Punkte liegen zwischen den Übereinstimmungsgrenzen.

Zusätzlich wurden die Ergebnisse aus ImageScope und Definiens in Boxplot Diagrammen gegenüber gestellt. Desweiteren wurden die Ergebnisse aus ImageScope und Definiens (siehe Tabelle 14 und Tabelle 15) in Streudiagrammen gegen das Tumolvolumen aufgetragen. Zusätzlich wurden die Korrelationskoeffizienten nach Spearman bestimmt und auf statistische Signifikanz getestet.

4 Ergebnisse

4.1 Scans der H.E. gefärbten Leberschnitte

In den H.E. Scans der Lebern sind über den ganzen Schnitt verteilte, sehr unterschiedlich große Ansammlungen von Zellen zu sehen (vgl. Abbildung 7). Die Zellen dieser Bereiche unterscheiden sich in vielerlei Hinsicht von den umgebenden Hepatozyten. Sie sind kleiner und haben einen deutlich schmaleren Zytoplasmasaum. Auffallend ist vor allem ihr Zellkern. Er nimmt fast das gesamte Zytoplasma ein, ist teils gelappt, teils rundlich und deutlich kompakter, also basophiler, als der Zellkern der Hepatozyten (vgl. Abbildung 7). Da im Versuch ein metastasierendes Tumormausmodell mit Metastasen in Lunge, Leber und Gehirn verwendet wurde, erwartete man in den Lebern Metastasen als Absiedlungen des Primärtumors. Die Zellen hingegen, die in den Lebern der tumortragenden Mäuse zu sehen sind, sind Immun- und keine Tumorzellen.

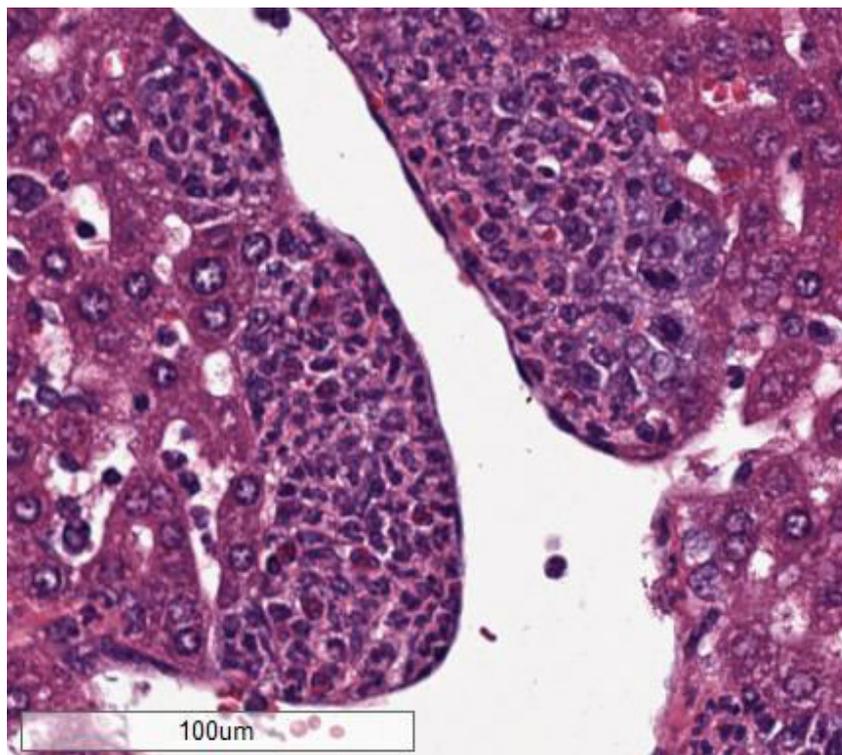


Abbildung 7: Leberscan H.E. – Läsion. Läsionen von Zellen mit schmalen Zytoplasmasaum, einem das gesamte Zytoplasma einnehmenden, teils gelappten, teils rundlichen, deutlich kompakten und basophilen Zellkern, umgeben von Hepatozyten (K5)

4.2 Scans der H.E. gefärbten Tumorschnitte

Wie bereits unter 2.1.1 beschrieben, bieten die Tumorzellen und Zellkerne in den Scans der Tumorschnitte ein polymorphes Bild. Das Kern-Plasma Verhältnis ist zugunsten eines großen Zellkerns verschoben und das Zellplasma basophil (siehe Abbildung 8). Desweiteren sind in den Scans der H.E. gefärbten Tumorschnitte ausgeprägte Nekrosen zu sehen (siehe Abbildung 9).

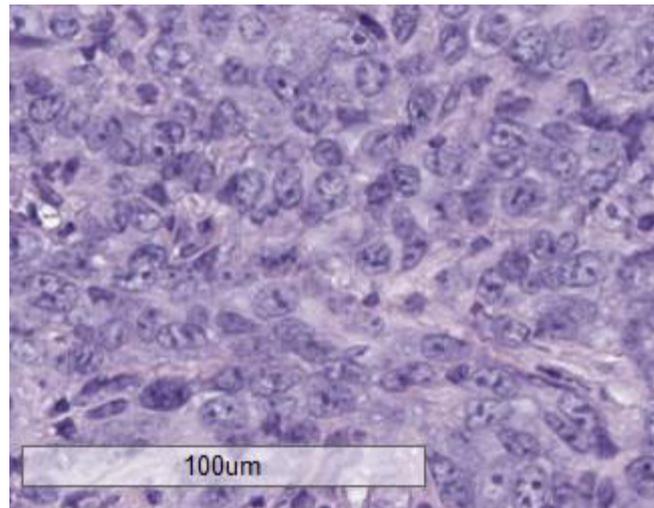


Abbildung 8: Tumorscan H.E. – Zellen. Polymorphe 4T1 Tumorzellen und Zellkerne. Verschobenes Kern-Plasma Verhältnis zugunsten eines großen Zellkerns und basophiles Zytoplasma (K1)

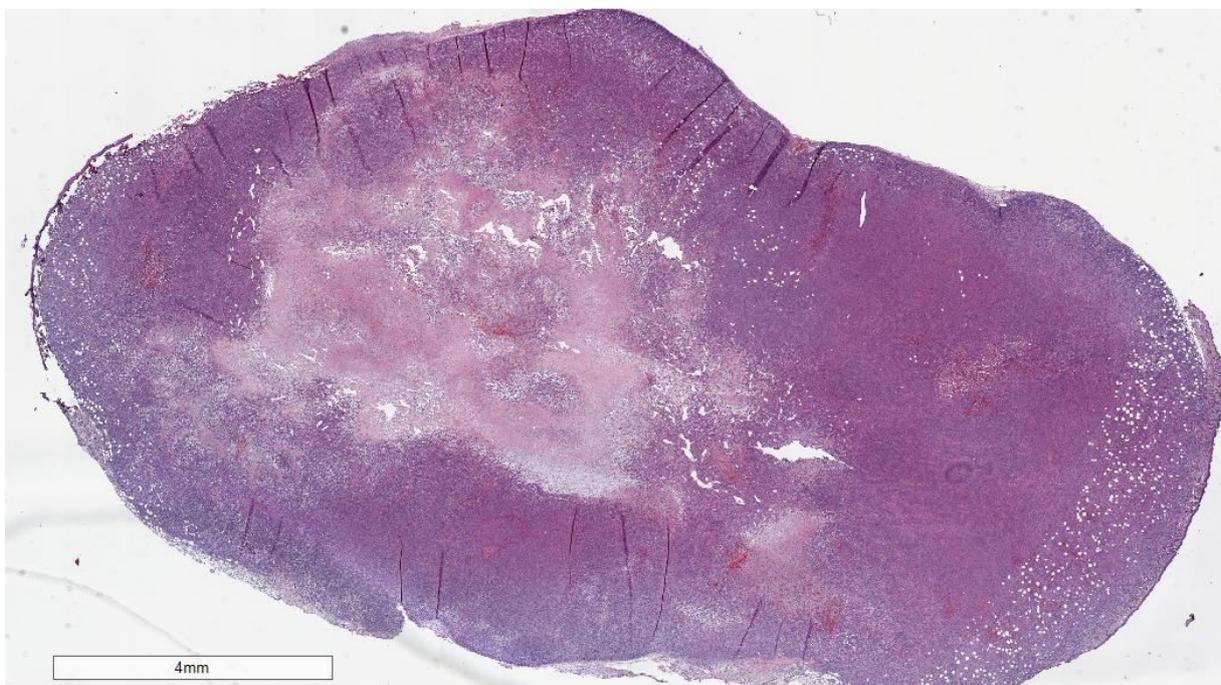


Abbildung 9: Tumorscan H.E. – Nekrose. Zentrale Nekrose (heller Bereich) in einem Primärtumor. Auch peripher einzelne nekrotische Bereiche (K6)

4.3 Scans der immunhistochemisch gefärbten Leberschnitte

Ziel der immunhistochemischen Färbung war es, die Zellen in den Zellansammlungen der Leber zu identifizieren und sie anhand ihrer Oberflächenmarker weiter zu differenzieren.

CD45 ist als Panleukozytenmarker bekannt und wird von Leukozyten aller Reifungsstufen exprimiert (Lai et al., 1998). Man sieht, dass fast alle Zellansammlungen in den Lebern stark CD45 positiv sind (vergleiche Abbildung 10 und Abbildung 11).

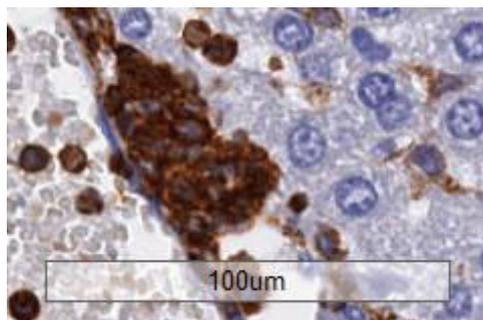


Abbildung 10: Leberscan IHC – stark CD45 positiv. Stark CD45 gefärbte Läsion in der Leber perivaskulär (K1)

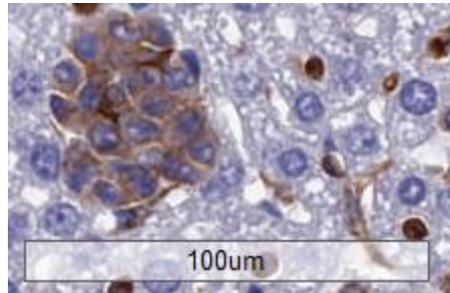


Abbildung 11: Leberscan IHC – CD45 positiv. CD45 positive Zellansammlung umgeben von Leberparenchym (K1)

Des Weiteren wollte ich den Verdacht bestätigen, dass es sich bei den EMH Herden in der Leber um MDSCs handelt. Tatsächlich sind fast alle Zellen in den Läsionen Ly6G/Ly6C positiv (siehe Abbildung 12 bis Abbildung 14). Desweiteren sind Ly6G/Ly6C negative Zellen zu sehen (siehe Abbildung 13 und Abbildung 14).

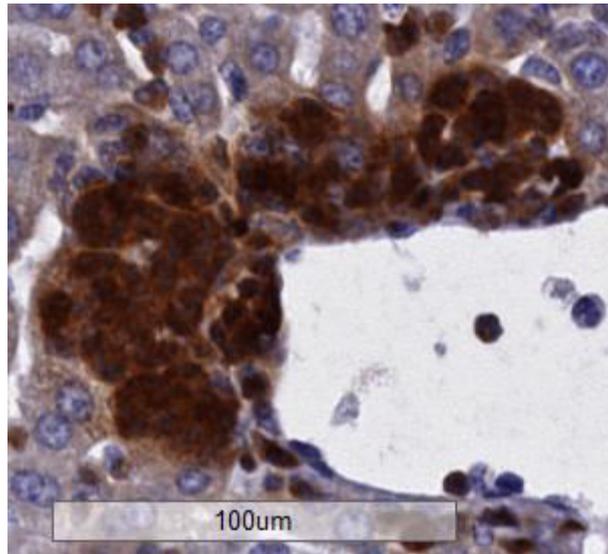


Abbildung 12: Leberscan IHC – stark Ly6G/ Ly6C positiv. Stark Ly6G/ Ly6C positive Läsion in der Leber perivaskulär (K5)

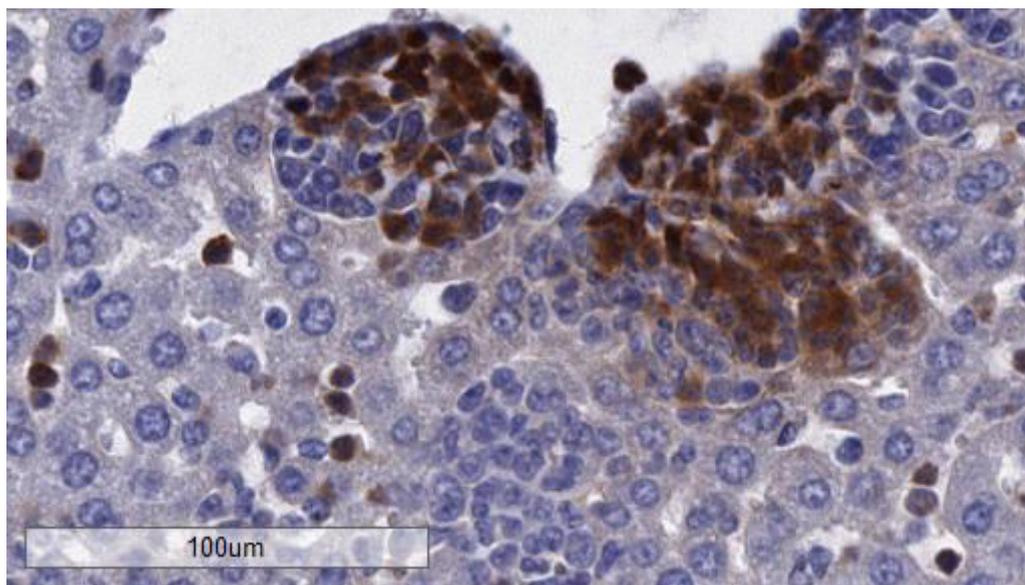


Abbildung 13: Leberscan IHC – Ly6G/ Ly6C positive und negative Läsionen. Stark Ly6G/ Ly6C positive und Ly6G/ Ly6C negative Zellansammlungen in der Leber perivaskulär (K5)

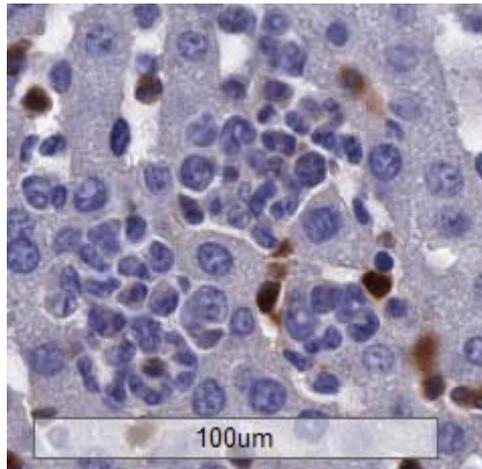


Abbildung 14: Leberscan IHC – Ly6G/ Ly6C negative Läsion. Ly6G/ Ly6C negative Zellansammlung umgeben von Leberparenchym (K5)

Es fällt auf, dass F4/80 positive Zellen die gesamte Leber durchziehen (siehe Abbildung 15). Auch in einigen Läsionen sind F4/80 positive Zellen zu finden (siehe Abbildung 16). F4/80 wird im Mausversuch als Marker für Makrophagen verwendet (Lai et al., 1998).

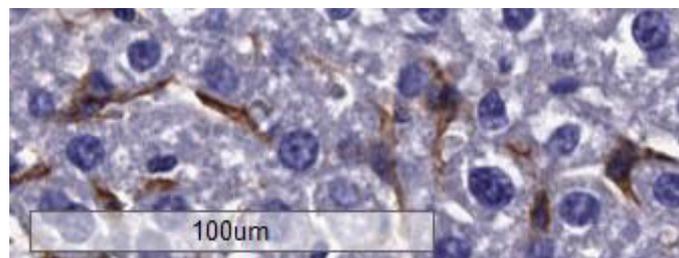


Abbildung 15: Leberscan IHC – F4/80 positive Zellen. F4/80 positive Zellen, die die Leber durchziehen (30 Gy 3)

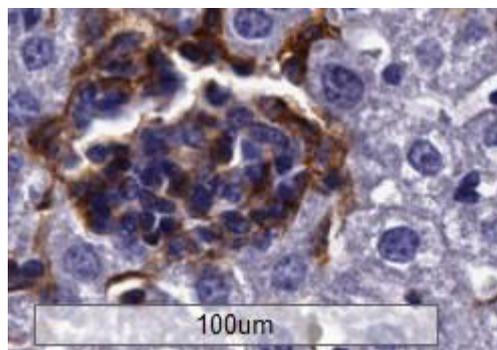


Abbildung 16: Leberscan IHC – F4/80 positive Zellen in einer Läsion. F4/80 positive Zellen in einer Läsion umgeben von Leberparenchym (30 Gy 3)

In den Läsionen und im Lebergewebe sind selten CD49b und CD3 positive Zellen zu sehen (siehe Abbildung 17 und Abbildung 18). CD49b dient der Identifikation von NK-Zellen und CD3 ist ein T-Zell-Marker (Stangl et al., 2011).

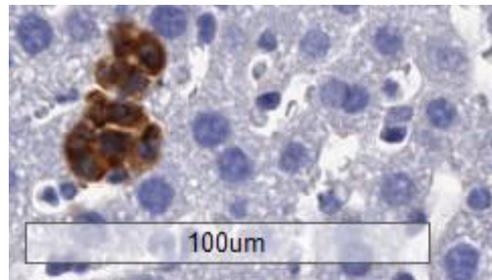


Abbildung 17: Leberscan IHC – CD49b positive Zellen in einer Läsion. CD49b positive Zellen in einer Läsion umgeben von Hepatozyten (K2)

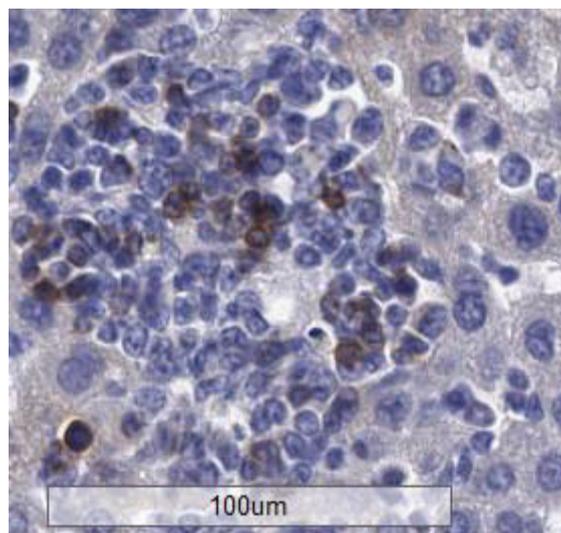


Abbildung 18: Leberscan IHC – CD3 positive Zellen in einer Läsion. CD3 positive Zellen in einer Läsion umgeben von Hepatozyten (K2)

4.4 Scans der immunohistochemisch gefärbten Tumorschnitte

Schnitte der Primärtumore wurden ebenfalls immunohistochemisch gefärbt. Einerseits um sicherzustellen, dass die Tumorzellen selbst keine der verwendeten Oberflächenmarker tragen, sich also durch die IHC von den Zellen in den Läsionen der Leber unterscheiden, und andererseits, um sich ein Bild von den tumorinfiltrierenden Immunzellen im verwendeten Modell zu machen. Die Tumorzellen lassen sich mit keinem der verwendeten Marker anfärben (siehe Abbildung 19 bis Abbildung 26). Im Tumor sind viele CD45 positive Zellen zu sehen. Diese infiltrieren sowohl den

Tumor selbst als auch das umgebende Bindegewebe sowie die nekrotischen Bereiche (siehe Abbildung 19 und Abbildung 20).

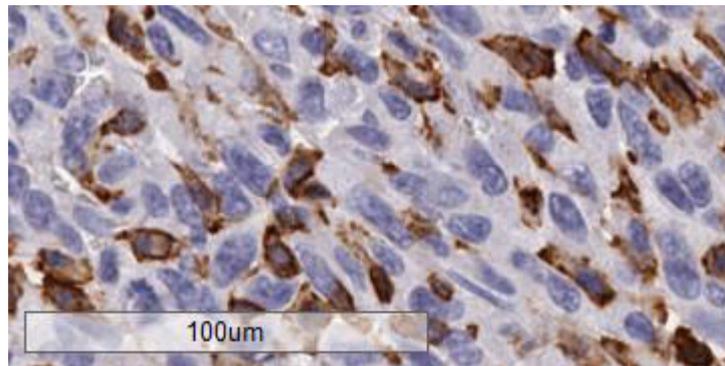


Abbildung 19: Tumorscan IHC – CD45 positive Zellen. CD45 positive Zellen umgeben von Tumorzellen (Op 4)

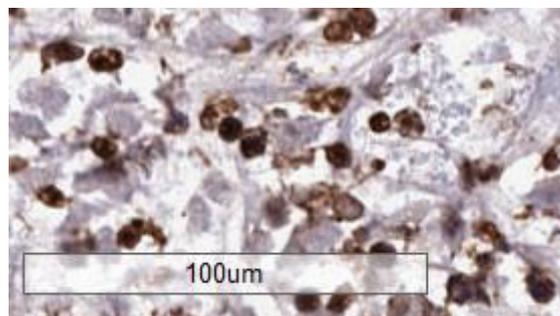


Abbildung 20: Tumorscan IHC – CD45 positive Zellen in nekrotischem Bereich. CD45 positive Zellen umgeben von Tumorzellen (K6)

Desweiteren sind zahlreiche Ly6G/Ly6C positive Zellen zu sehen (siehe Abbildung 21 und Abbildung 22). F4/80 positive Zellen durchziehen das Gewebe (siehe Abbildung 23). Desweiteren findet man einige CD49b und CD3 positive Zellen (siehe Abbildung 24 bis Abbildung 26).

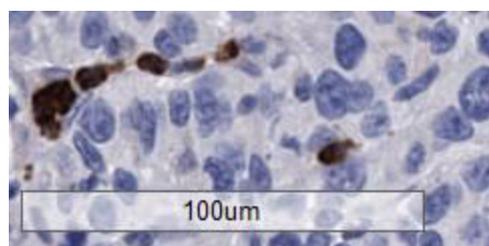


Abbildung 21: Tumorscan IHC – einzelne Ly6G/ Ly6C positive Zellen. Ly6G/ Ly6C positive Zellen umgeben von Tumorzellen (Op 4)

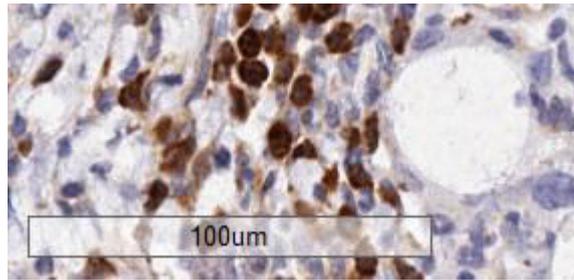


Abbildung 22: Tumorscan IHC – zahlreiche Ly6G/ Ly6C positive Zellen. Zahlreiche Ly6G/ Ly6C positive Zellen im an den Tumor angrenzenden Bindegewebe (Op 6)

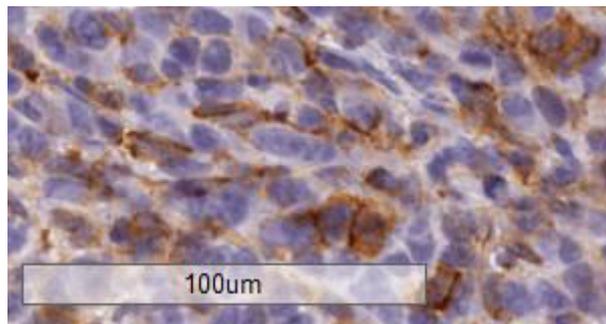


Abbildung 23: Tumorscan IHC – F4/80 positive Zellen. F4/80 positive Zellen, die das Gewebe durchziehen, umgeben von Tumorzellen (Op 6)

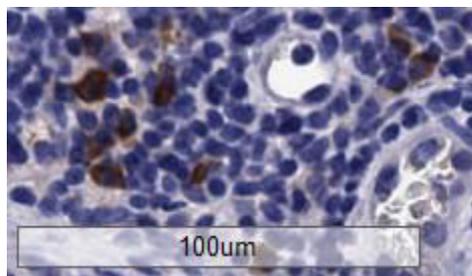


Abbildung 24: Tumorscan IHC – CD49b positive Zellen. CD49b positive Zellen im den Tumor umgebenden Bindegewebe (Op2)

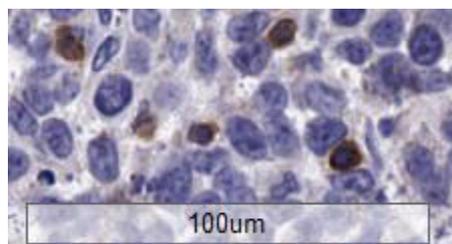


Abbildung 25: Tumorscan IHC – CD3 positive Zellen. CD3 positive Zellen umgeben von Tumorzellen (K5)

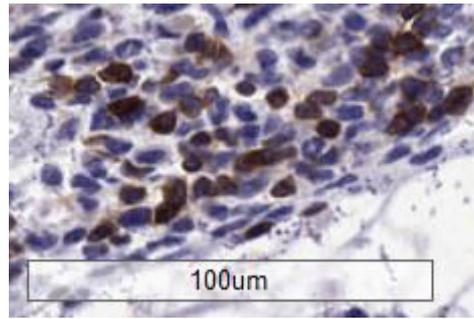


Abbildung 26: Tumorscan IHC – CD3 positive Zellen. CD3 positive Zellen im den Tumor umgebenden Bindegewebe (K5)

4.5 Verteilung der Primärtumorvolumina

Anhand der Abbildung 27 lässt sich das Volumen der Primärtumorvolumina in den verschiedenen Gruppen vergleichen. Die Tumore der Kontrollmäuse, die keinerlei Behandlung erhalten haben, haben erwartungsgemäß das größte Volumen. Die Tumorvolumina der operierten Mäuse sind durchschnittlich am kleinsten, da die Tumore in dieser Gruppe nur eine Wachstumsphase von zwei statt vier Wochen wie in den beiden anderen Gruppen hatten. In einem Kruskal-Wallis-Test für unabhängige Stichproben wurde statistisch nachgewiesen, dass die Tumorvolumina nicht in allen Gruppen gleich sind ($p=0.005$) (siehe Tabelle 16). Aus dem anschließenden Mann-Whitney-U-Test ergab sich ein statistisch signifikanter Unterschied zwischen den Tumorvolumina der Kontrollgruppe und der bestrahlten Gruppe ($p=0.038$) sowie zwischen der Kontrollgruppe und der operierten Gruppe ($p=0.002$), nicht jedoch zwischen der bestrahlten und der operierten Gruppe ($p=0.114$) (siehe Tabelle 17).

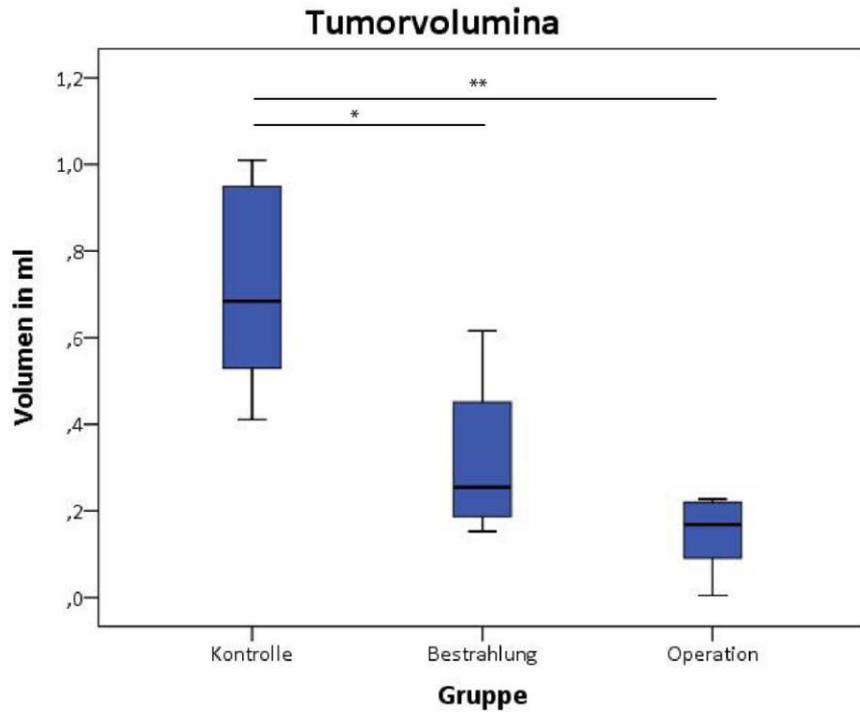


Abbildung 27: Tumorvolumina. Verglichen werden die Primärtumorvolumina der Gruppen Kontrolle, Bestrahlung und Operation. Signifikante Unterschiede zwischen den Primärtumorvolumina Kontrolle/ Bestrahlung ($p=0.038$) und Kontrolle/ Operation ($p=0.002$) (siehe Tabelle 17).

Tabelle 16: Kruskal-Wallis-Test – Tumorvolumina

Kruskal-Wallis-Test bei unabhängigen Stichproben	p-Wert
Kontrolle, Bestrahlung, Operation	0.005**

* Signifikanzniveau bei 0.05.
 ** Signifikanzniveau bei 0.01.

Tabelle 17: Mann-Whitney-U-Test – Tumorvolumina

Mann-Whitney-U-Test bei unabhängigen Stichproben	p-Wert
Kontrolle vs. Bestrahlung	0.038*
Kontrolle vs. Operation	0.002**
Bestrahlung vs. Operation	0.114

* Signifikanzniveau bei 0.05.
 ** Signifikanzniveau bei 0.01.

4.6 ImageScope – EMH Parameter und statistische Tests

Das Ausmaß der EMH in den verschiedenen Behandlungsgruppen sollte graphisch veranschaulicht und auf statistisch signifikante Unterschiede untersucht werden.

Anhand Abbildung 28 bis Abbildung 32 lässt sich das Ausmaß der EMH in der Leber in den drei Behandlungsgruppen, wie es mittels ImageScope ermittelt wurde, vergleichen. Abbildung 28 zeigt die Zahl der Läsionen pro Fläche. Anschließend folgt in Abbildung 29 bis Abbildung 31 die Darstellung der Läsionen pro Fläche in den drei Untergruppen - kleine, mittlere und große Läsionen. Abbildung 32 zeigt den prozentualen Anteil der Läsionsfläche an der Gesamtfläche. Die meisten Läsionen aus jeder der drei Gruppen liegen im Bereich der kleinen Läsionen (vergleiche Abbildung 29). In der Kontrollgruppe gibt es durchschnittlich ca. 15 Läsionen pro mm^2 , in den anderen Gruppen unter 5 pro mm^2 . Die Gruppe der großen Läsionen ist die zahlenmäßig kleinste Gruppe, es gab in jeder Gruppe weniger als eine Läsion pro mm^2 (siehe Abbildung 31). Große Läsionen sind also eher selten. Betrachtet man Abbildung 28 bis Abbildung 31, so sieht man, dass es in allen Untergruppen in der Kontrollgruppe die meisten Läsionen gibt. Tendenziell hat die bestrahlte Gruppe die zweitgrößte Zahl an Läsionen und die operierte Gruppe die geringste Zahl an Läsionen. In einem Kruskal-Wallis-Test für unabhängige Stichproben wurde statistisch nachgewiesen, dass die Parameter in den verschiedenen Gruppen nicht gleich sind (siehe Tabelle 18). Aus dem anschließenden Mann-Whitney-U-Test ergaben sich statistisch signifikante Unterschiede zwischen Kontrollgruppe und bestrahlter Gruppe und Kontrollgruppe und operierter Gruppe bei allen ermittelten Parametern (vergleiche Tabelle 19). Zwischen bestrahlter Gruppe und operierter Gruppe gibt es bei keinem der ermittelten Parameter einen signifikanten Unterschied (vergleiche Tabelle 19). Auch in der unbehandelten Maus lässt sich EMH finden (siehe Abbildung 28 bis Abbildung 32). Hier wurden die Mittelwerte von zwei Scans nur einer Maus aufgetragen. Deshalb konnten keine statistischen Tests durchgeführt werden.

Man sieht, wenn man Tabelle 20 betrachtet, dass die Werte zwischen Ellipse und Rechteck vergleichbar sind. In einem Bland-Altman-Plot wurden die Mittelwerte der Messwerte Ellipse und Rechteck gegen die Differenz aus Ellipse und Rechteck aufgetragen (siehe Abbildung 33). Die Werte aus der Methode Ellipse ergeben in 95% der Fälle einen Wert, der um 3.6 größer bzw. um 4.3 kleiner ist als der Wert aus der Methode Rechteck.

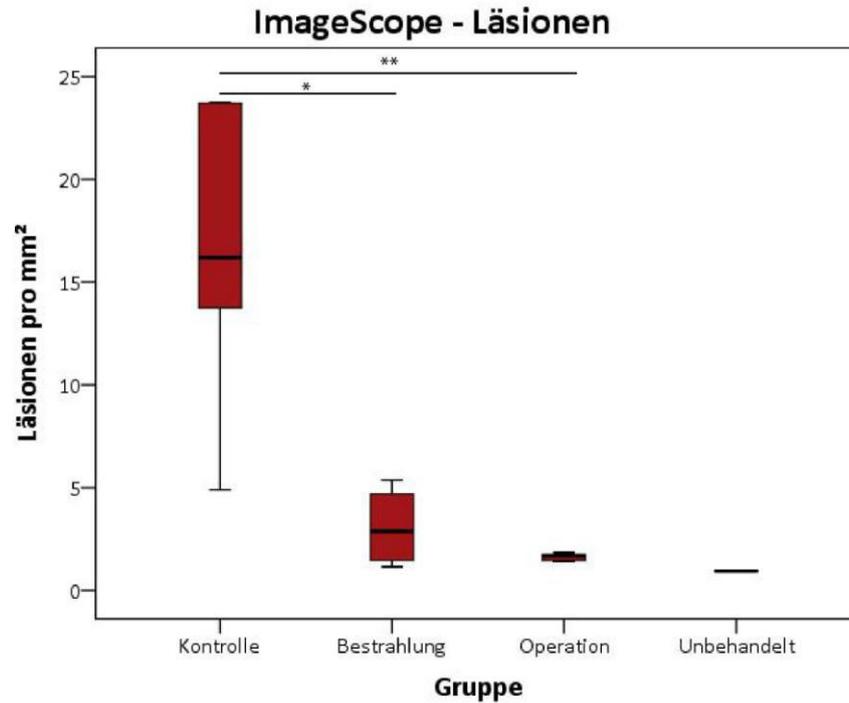


Abbildung 28: ImageScope – Läsionen pro Fläche. Verglichen wird die Anzahl der Läsionen pro mm² der verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation ($p=0.002$) und Kontrolle/ Bestrahlung ($p=0.019$) (siehe Tabelle 19).

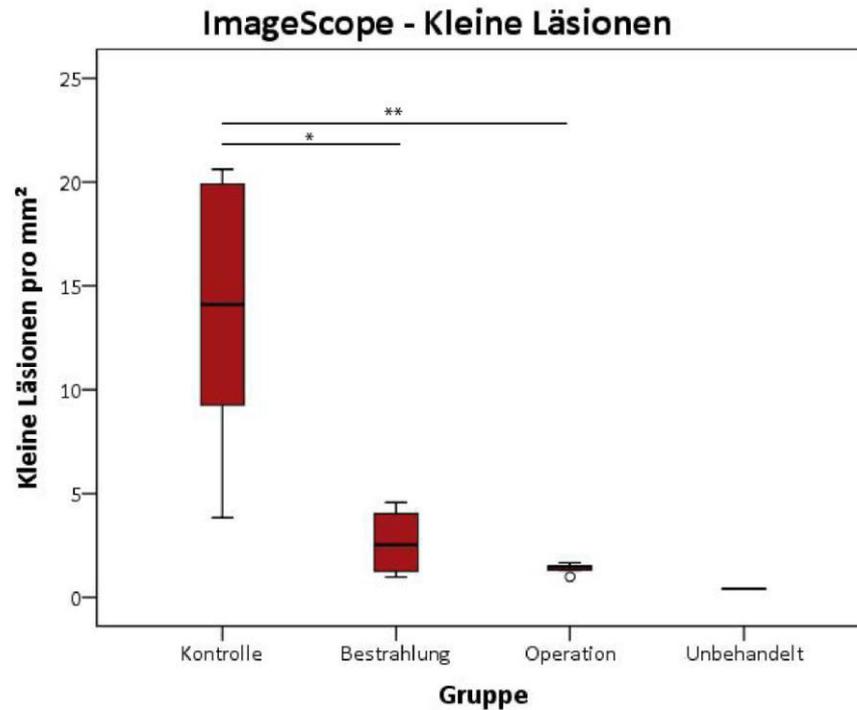


Abbildung 29: ImageScope – Kleine Läsionen pro Fläche. Verglichen wird die Anzahl der kleinen Läsionen pro mm² in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation ($p=0.002$) und Kontrolle/ Bestrahlung ($p=0.019$) (siehe Tabelle 19).

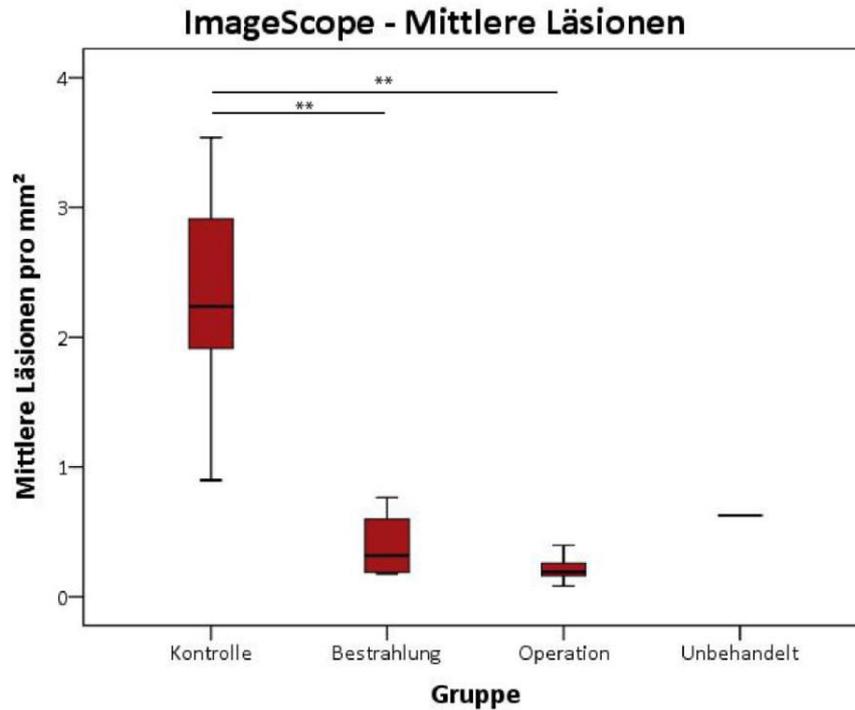


Abbildung 30: ImageScope – Mittlere Läsionen pro Fläche. Verglichen wird die Anzahl der mittleren Läsionen pro mm² in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation ($p=0.002$) und Kontrolle/ Bestrahlung ($p=0.01$) (siehe Tabelle 19).

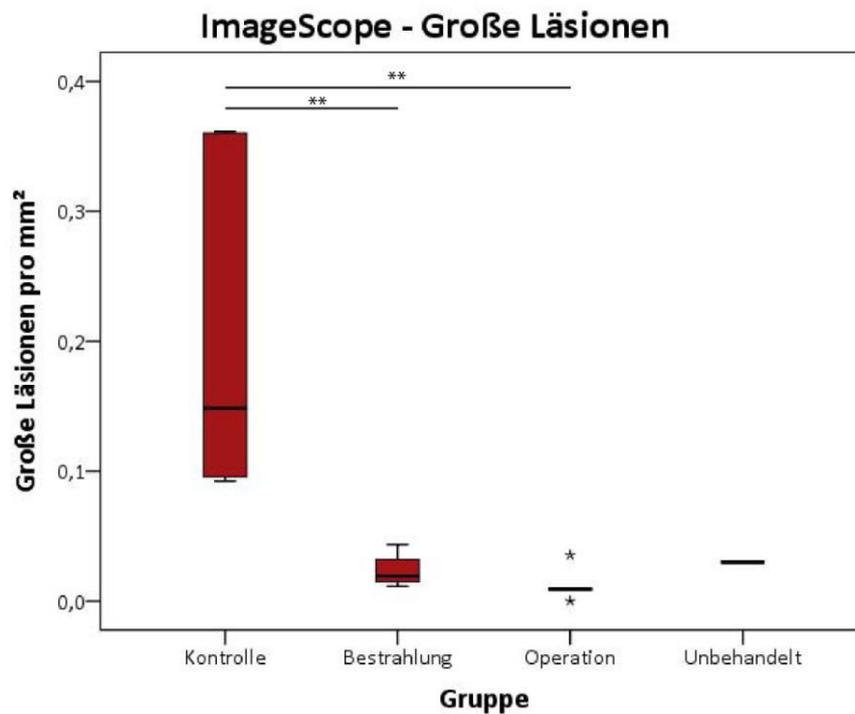


Abbildung 31: ImageScope – Große Läsionen pro Fläche. Verglichen wird die Anzahl der großen Läsionen pro mm² in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation ($p=0.002$) und Kontrolle/ Bestrahlung ($p=0.01$) (siehe Tabelle 19).

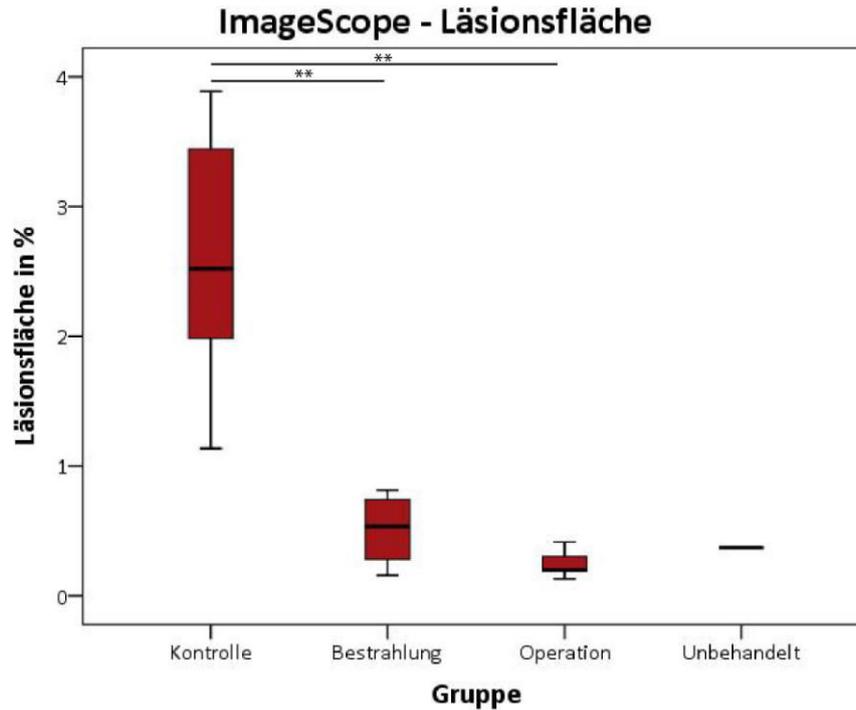


Abbildung 32: ImageScope – Läsionsfläche pro Gesamtfläche. Verglichen wird der Anteil der Läsionsfläche an der Gesamtfläche in Prozent in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation ($p=0.002$) und Kontrolle/ Bestrahlung ($p=0.01$) (siehe Tabelle 19).

Tabelle 18: Kruskal-Wallis-Test – EMH, ImageScope

ImageScope	p- Werte				
	Läsionen pro mm ²	Kleine Läsionen pro mm ²	Mittlere Läsionen pro mm ²	Große Läsionen pro mm ²	Läsionsfläche pro Gesamtfläche in %
Kruskal-Wallis-Test bei unabhängigen Stichproben					
Kontrolle, Bestrahlung, Operation	0.006**	0.005**	0.004**	0.002**	0.004**

* Signifikanzniveau bei 0.05.

** Signifikanzniveau bei 0.01.

Tabelle 19: Mann-Whitney-U-Test – EMH, ImageScope

ImageScope	p- Werte				
Mann-Whitney-U-Test bei unabhängigen Stichproben	Läsionen pro mm ²	Kleine Läsionen pro mm ²	Mittlere Läsionen pro mm ²	Große Läsionen pro mm ²	Läsionsfläche pro Gesamtfläche in %
Kontrolle vs. Bestrahlung	0.019*	0.019*	0.01**	0.01**	0.01**
Kontrolle vs. Operation	0.002**	0.002**	0.002**	0.002**	0.002**
Bestrahlung vs. Operation	0.476	0.352	0.257	0.67	0.257

* Signifikanzniveau bei 0.05.

** Signifikanzniveau bei 0.01.

Tabelle 20: Gegenüberstellung Ellipse vs. Rechteck – ImageScope

Parameter	Schnitte					
	K4 75 R	K4 75 O	30 Gy 2 65 R	30 Gy 2 65 O	Op 5 135 R	Op 5 135 O
Herdfläche/Gesamtfläche in %	3.78	3.71	0.57	0.36	0.18	0.42
Herde/mm ²	26.79	23.52	2.14	2.22	1.39	2.36
kleine Herde/mm ²	23.65	18.39	1.97	1.86	1.12	1.77
mittlere Herde/mm ²	2.98	5.13	0.16	0.32	0.27	0.48

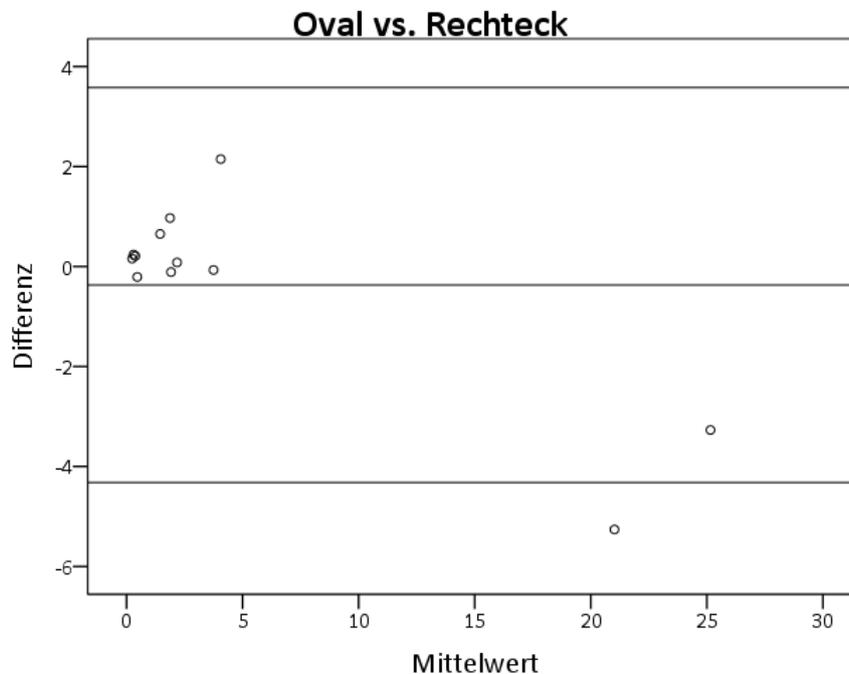


Abbildung 33: Bland-Altman-Plot: Ellipse vs. Rechteck. Dargestellt sind die Mittelwerte der Parameter aus Ellipse und Rechteck aufgetragen gegen die Differenzen aus Ellipse und Rechteck (vergleiche Tabelle 20). Außerdem dargestellt die Verzerrung als Mittelwert aller Differenzen (mittlere Linie) sowie die Übereinstimmungsgrenzen $d \pm 2 \times s$, wobei d den Mittelwert der Differenzen und s die Standardabweichung der Differenzen bezeichnet (obere und untere Linie).

4.7 Definiens – EMH Parameter und statistische Tests

Auch die Ergebnisse aus Definiens wurden graphisch dargestellt und die EMH Parameter auf statistisch signifikante Unterschiede zwischen den Gruppen untersucht.

Abbildung 34 bis Abbildung 38 zeigen das Ausmaß der EMH in der Leber in den drei Behandlungsgruppen. Abbildung 34 zeigt die Zahl der Läsionen pro Fläche. Anschließend folgt in Abbildung 35 bis Abbildung 37 die Darstellung der Läsionen pro Fläche in den drei Untergruppen - kleine, mittlere und große Läsionen. Abbildung 38 zeigt den prozentualen Anteil der Läsionsfläche an der Gesamtfläche. Wieder sind die kleinen Läsionen in jeder der drei Gruppen die meisten (vergleiche Abbildung 35) und die großen Läsionen die wenigsten (siehe Abbildung 37). In der Kontrollgruppe gibt es durchschnittlich ca. 8 Läsionen pro mm^2 , in den anderen Gruppen unter 5 pro mm^2 . Genau wie bei den Ergebnissen aus ImageScope gibt es in allen Untergruppen in der Kontrollgruppe die meisten Läsionen (Abbildung 34 bis Abbildung 37). Tendenziell hat die bestrahlte Gruppe die zweitgrößte Zahl und die operierte Gruppe die geringste Zahl an Läsionen. In einem Kruskal-Wallis-Test für unabhängige Stichproben wurde statistisch nachgewiesen, dass die Parameter in den verschiedenen Gruppen nicht gleich waren (siehe Tabelle 21). Aus dem anschließenden Mann-Whitney-U-Test ergaben sich statistisch signifikante Unterschiede zwischen Kontrollgruppe und bestrahlter Gruppe und Kontrollgruppe und operierter Gruppe bei fast allen ermittelten Parametern (vergleiche Tabelle 22). Nur bei den kleinen Läsionen pro Fläche gibt es keinen statistisch signifikanten Unterschied zwischen Kontrollgruppe und bestrahlter Gruppe (vergleiche Tabelle 22). Zwischen bestrahlter und operierter Gruppe gibt es bei den mittleren Läsionen pro Fläche einen signifikanten Unterschied (vergleiche Tabelle 22). Auch in der unbehandelten Maus lässt sich EMH finden (siehe Abbildung 34 bis Abbildung 38). Da wiederum nur die Leber einer unbehandelten Maus untersucht wurde, konnten die Ergebnisse nicht in die Signifikanztests einbezogen werden.

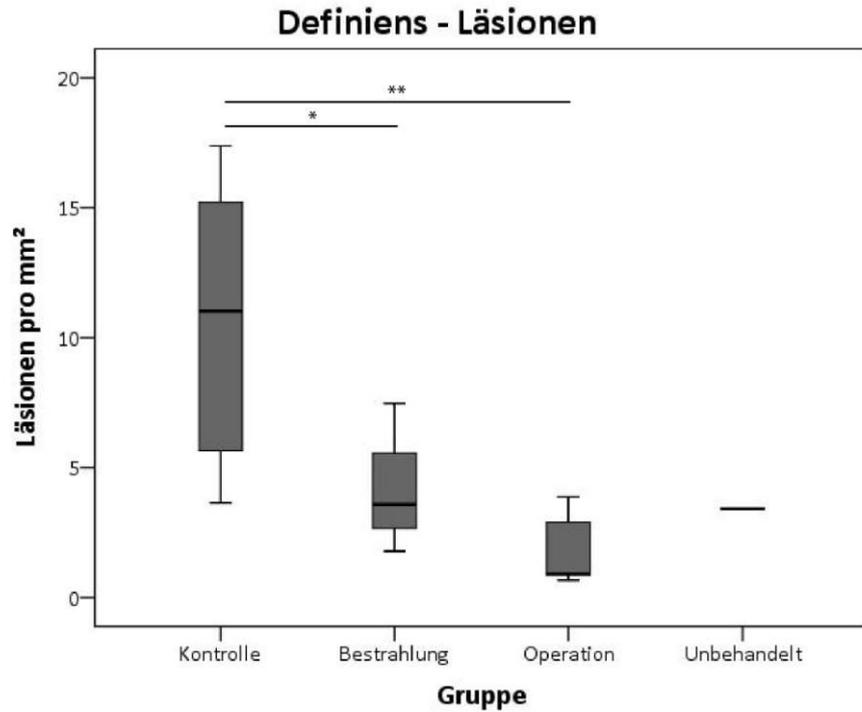


Abbildung 34: Definiens – Läsionen pro Fläche. Verglichen wird die Anzahl der Läsionen pro mm² in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation ($p=0.004$) und Kontrolle/ Bestrahlung ($p=0.038$) (siehe Tabelle 22).

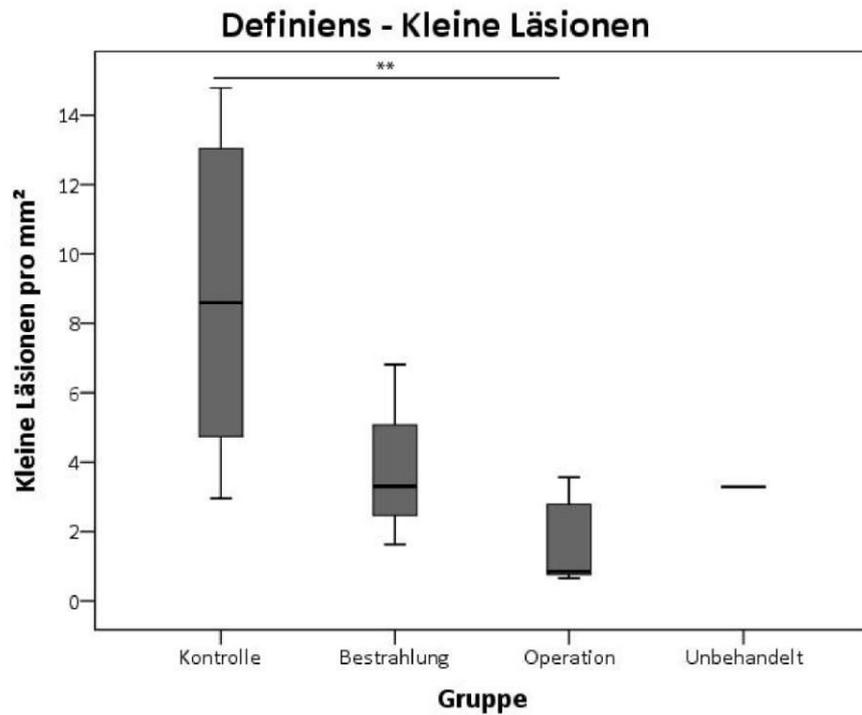


Abbildung 35: Definiens – Kleine Läsionen pro Fläche. Verglichen wird die Anzahl der kleinen Läsionen pro mm² in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation ($p=0.004$) (siehe Tabelle 22).

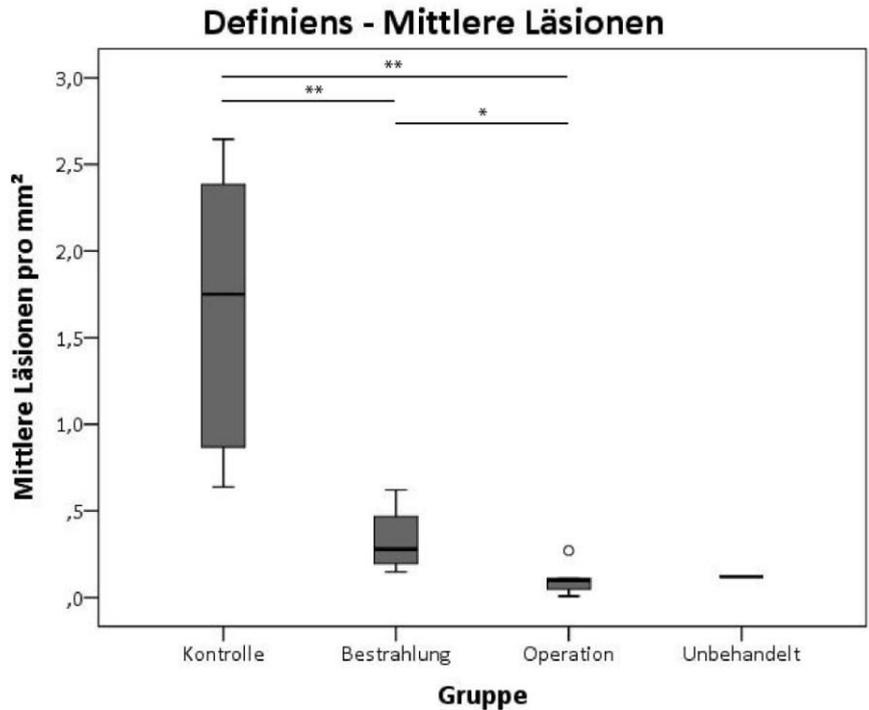


Abbildung 36: Definiens – Mittlere Läsionen pro Fläche. Verglichen wird die Anzahl der mittleren Läsionen pro mm² in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation (p=0.002), Kontrolle/ Bestrahlung (p=0.01) und Bestrahlung/ Operation (p=0.038) (siehe Tabelle 22).

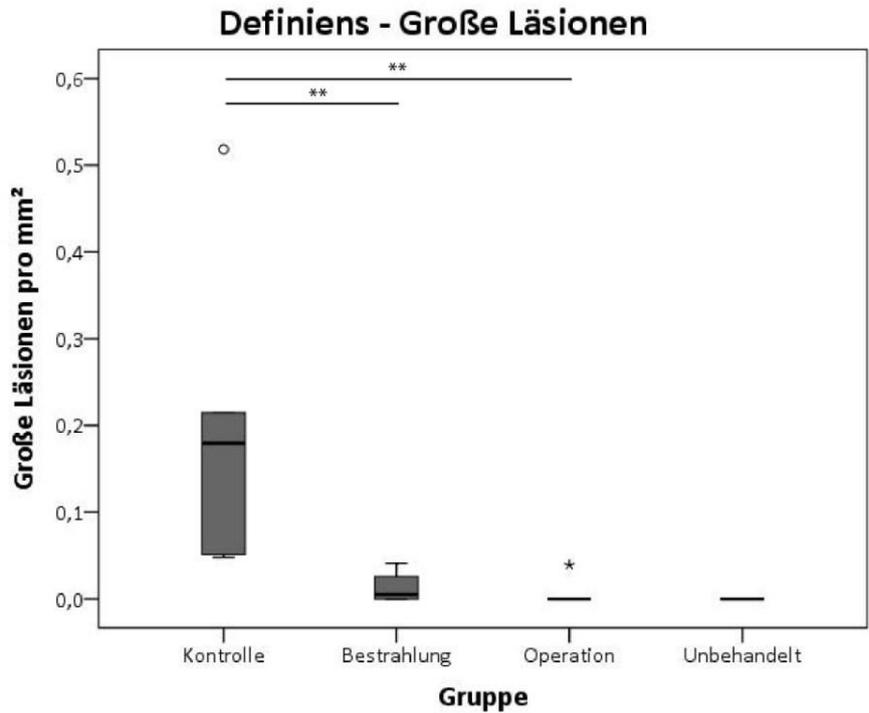


Abbildung 37: Definiens – Große Läsionen pro Fläche. Verglichen wird die Anzahl der großen Läsionen pro mm² in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation (p=0.002) und Kontrolle/ Bestrahlung (p=0.01) (siehe Tabelle 22).

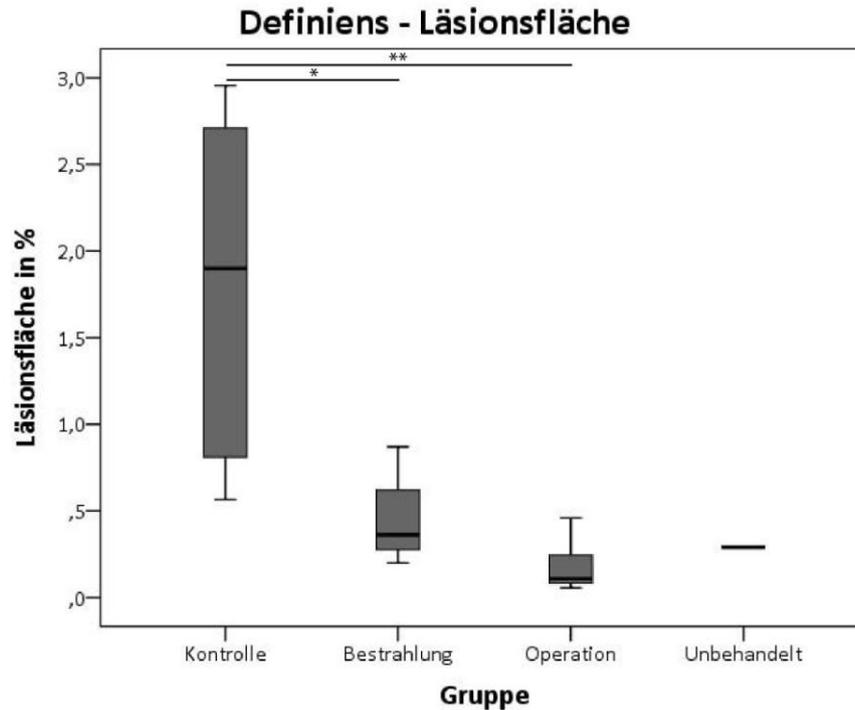


Abbildung 38: Definiens – Läsionsfläche pro Gesamtfläche. Verglichen wird der Anteil der Läsionsfläche an der Gesamtfläche in Prozent in den verschiedenen Gruppen. Signifikante Unterschiede zwischen Kontrolle/ Operation ($p=0.002$) und Kontrolle/ Bestrahlung ($p=0.038$) (siehe Tabelle 22).

Tabelle 21: Kruskal-Wallis-Test – EMH, Definiens

Definiens	p-Werte				
	Läsionen pro mm ²	Kleine Läsionen pro mm ²	Mittlere Läsionen pro mm ²	Große Läsionen pro mm ²	Läsionsfläche pro Gesamtfläche in %
Kruskal-Wallis-Test bei unabhängigen Stichproben					
Kontrolle, Bestrahlung, Operation, Unbehandelt	0.013*	0.023*	0.004*	0.005*	0.009*

* Signifikanzniveau bei 0.05.

** Signifikanzniveau bei 0.01.

Tabelle 22: Mann-Whitney-U-Test – EMH, Definiens

Definiens	p-Werte				
	Läsionen pro mm ²	Kleine Läsionen pro mm ²	Mittlere Läsionen pro mm ²	Große Läsionen pro mm ²	Läsionsfläche pro Gesamtfläche in %
Mann-Whitney-U-Test bei unabhängigen Stichproben					
Kontrolle vs. Bestrahlung	0.038*	0.114	0.01**	0.01**	0.038*
Kontrolle vs. Operation	0.004**	0.004**	0.002**	0.002**	0.002**
Bestrahlung vs. Operation	0.114	0.114	0.038*	0.476	0.114

* Signifikanzniveau bei 0.05.

** Signifikanzniveau bei 0.01.

4.8 Vergleich der Verfahren – ImageScope und Definiens

Ein weiteres Ziel dieser Arbeit war festzustellen ob die beiden verwendeten Verfahren zur Quantifizierung von EMH – ImageScope und Definiens – vergleichbare Ergebnisse hervorbringen. In Abbildung 39 bis Abbildung 43 werden die Ergebnisse aus ImageScope und Definiens in Boxplots gegenüber gestellt. Dabei ist zu sehen, dass grundsätzlich sehr ähnliche Tendenzen bei beiden Verfahren zu beobachten sind. Dennoch berechnet Definiens etwas geringere Werte in der Kontrollgruppe als ImageScope. Um die Unterschiede der Verfahren graphisch darzustellen, wurden erneut Bland-Altman-Plots erstellt, in denen die Mittelwerte der Messpaare gegen ihre Differenzen aufgetragen wurden (siehe Abbildung 44 bis Abbildung 48). Größere Werte gehen mit größeren Differenzen einher, was vor allem in der Kontrollgruppe der Fall ist. In der Kontrollgruppe erkennt man schon im Boxplot einen Unterschied zwischen ImageScope und Definiens über alle Parameter, wobei Definiens kleinere Werte ergibt. Dazu passt, dass große Differenzen und Ausreißer meist im negativen Bereich liegen. Die Verzerrung liegt bei allen Diagrammen im negativen Bereich (siehe Tabelle 33).

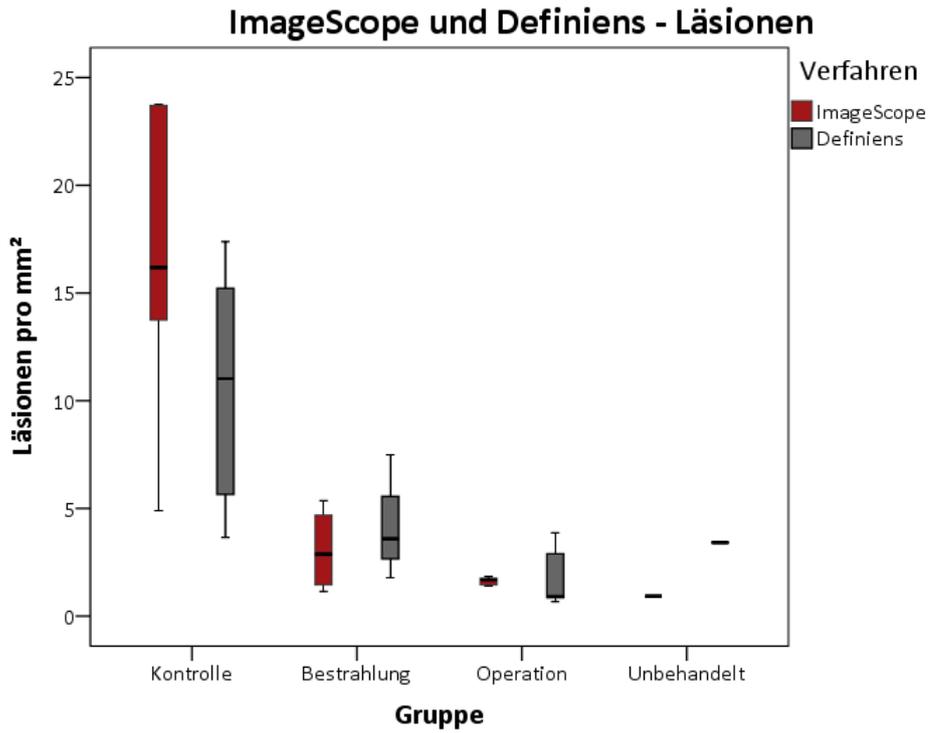


Abbildung 39: ImageScope und Definiens - Läsionen pro Fläche. Verglichen werden die beiden angewendeten Verfahren ImageScope und Definiens bezüglich der Anzahl der Läsionen pro mm² in den verschiedenen Gruppen.

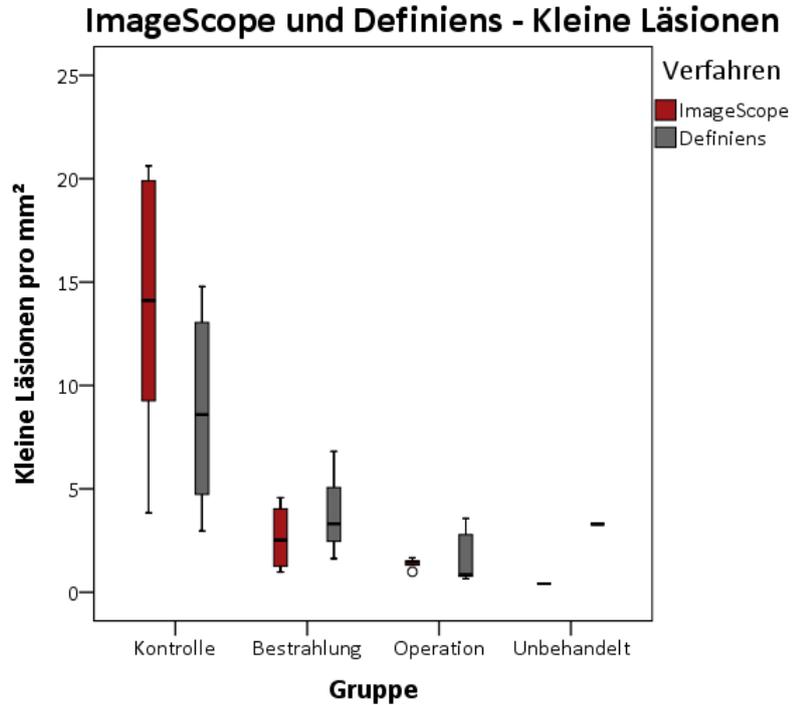


Abbildung 40: ImageScope und Definiens – Kleine Läsionen pro Fläche. Verglichen werden die beiden angewendeten Verfahren ImageScope und Definiens bezüglich der Anzahl der kleinen Läsionen pro mm² in den verschiedenen Gruppen.

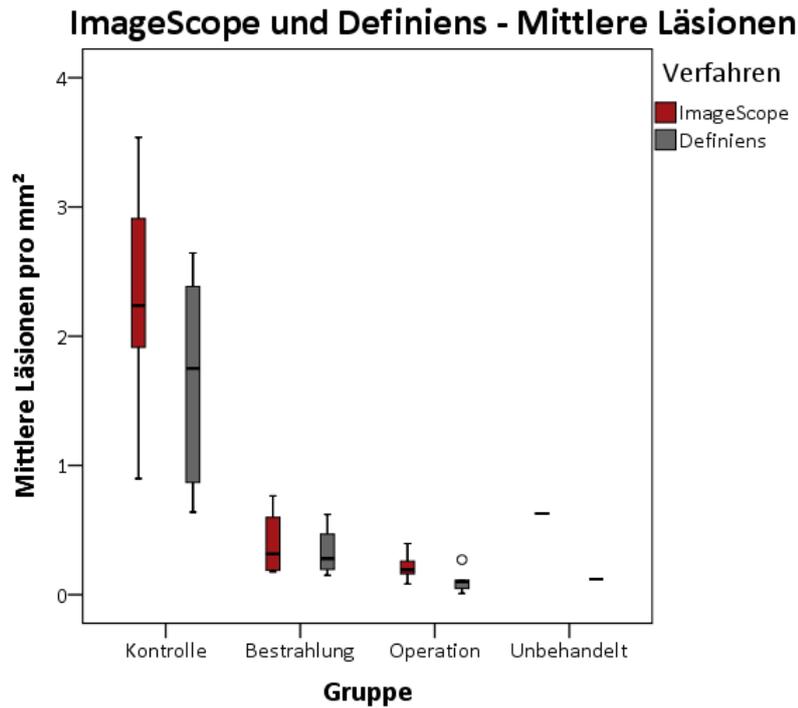


Abbildung 41: ImageScope und Definiens – Mittlere Läsionen pro Fläche. Verglichen werden die beiden angewendeten Verfahren ImageScope und Definiens bezüglich der Anzahl der mittleren Läsionen pro mm² in den verschiedenen Gruppen.

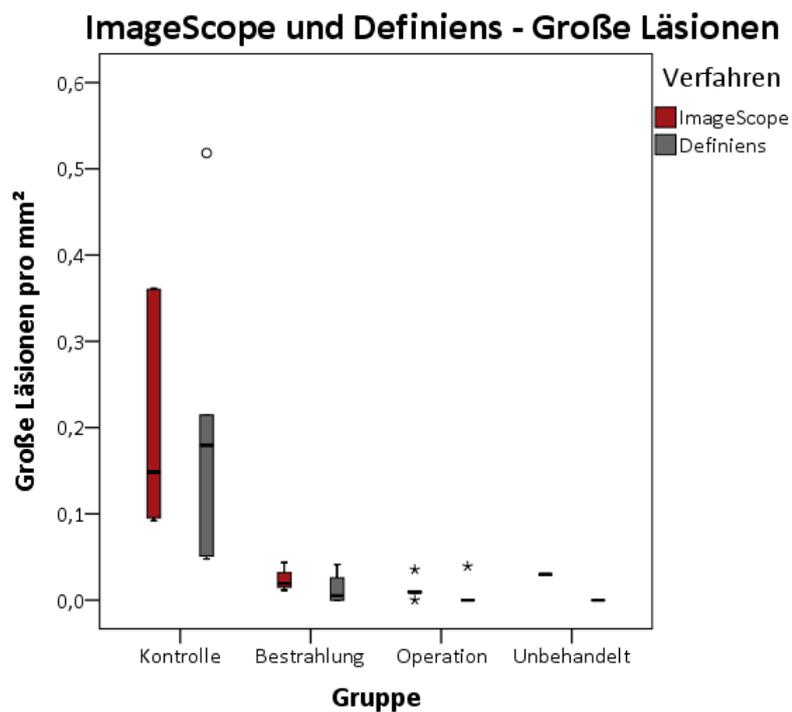


Abbildung 42: ImageScope und Definiens – Große Läsionen pro Fläche. Verglichen werden die beiden angewendeten Verfahren ImageScope und Definiens bezüglich der Anzahl der großen Läsionen pro mm² in den verschiedenen Gruppen.

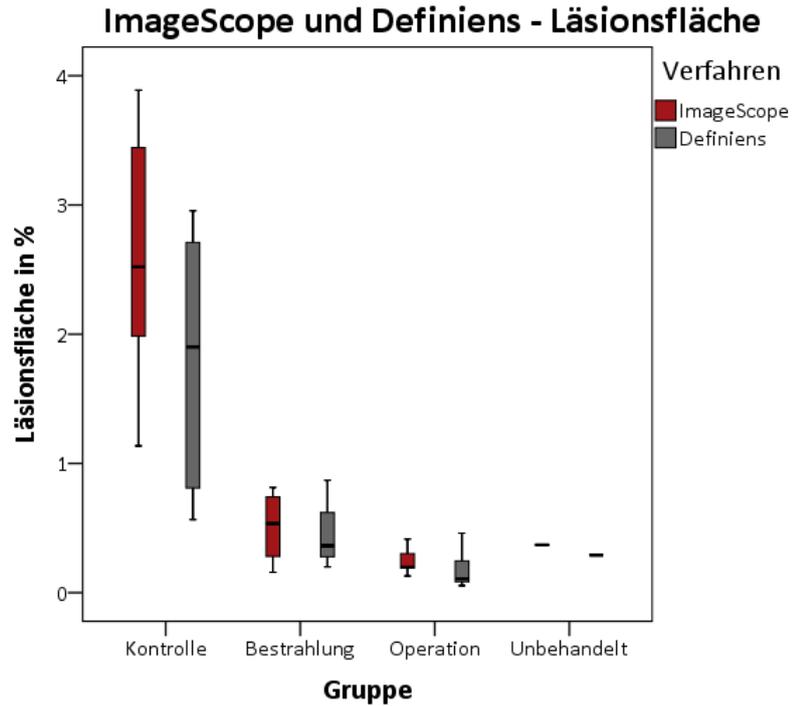


Abbildung 43: ImageScope und Definiens – Läsionsfläche pro Gesamtfläche Verglichen werden die beiden angewendeten Verfahren ImageScope und Definiens bezüglich dem Anteil der Läsionsfläche an der Gesamtfläche in Prozent in den verschiedenen Gruppen.

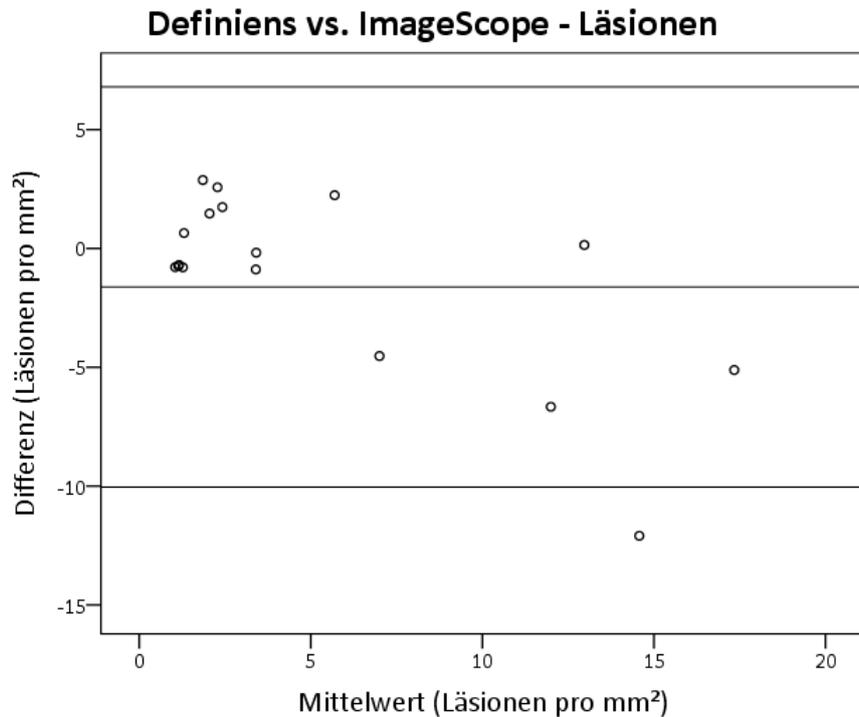


Abbildung 44: Bland-Altman-Plot Definiens vs. ImageScope – Läsionen pro Fläche. Dargestellt sind die Mittelwerte der Läsionen pro mm² aus Definiens und ImageScope aufgetragen gegen die Differenzen der Läsionen pro mm² aus Definiens und ImageScope. Außerdem dargestellt die Verzerrung als Mittelwert über alle Differenzen (mittlere Linie) sowie die Übereinstimmungsgrenzen $d \pm 2 \times s$, wobei d den Mittelwert der Differenzen und s die Standardabweichung der Differenzen bezeichnet (obere und untere Linie).

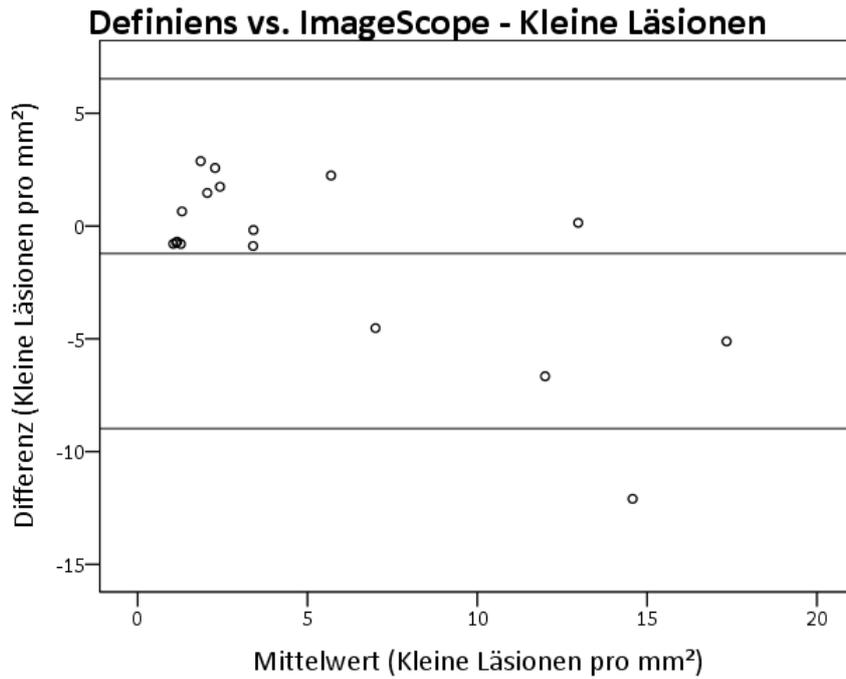


Abbildung 45: Bland-Altman-Plot Definiens vs. ImageScope – Kleine Läsionen pro Fläche. Dargestellt sind die Mittelwerte der kleinen Läsionen pro mm² aus Definiens und ImageScope aufgetragen gegen die Differenzen der kleinen Läsionen pro mm² aus Definiens und ImageScope. Außerdem dargestellt die Verzerrung als Mittelwert über alle Differenzen (mittlere Linie) sowie die Übereinstimmungsgrenzen $d \pm 2 \times s$, wobei d den Mittelwert der Differenzen und s die Standardabweichung der Differenzen bezeichnet (obere und untere Linie).

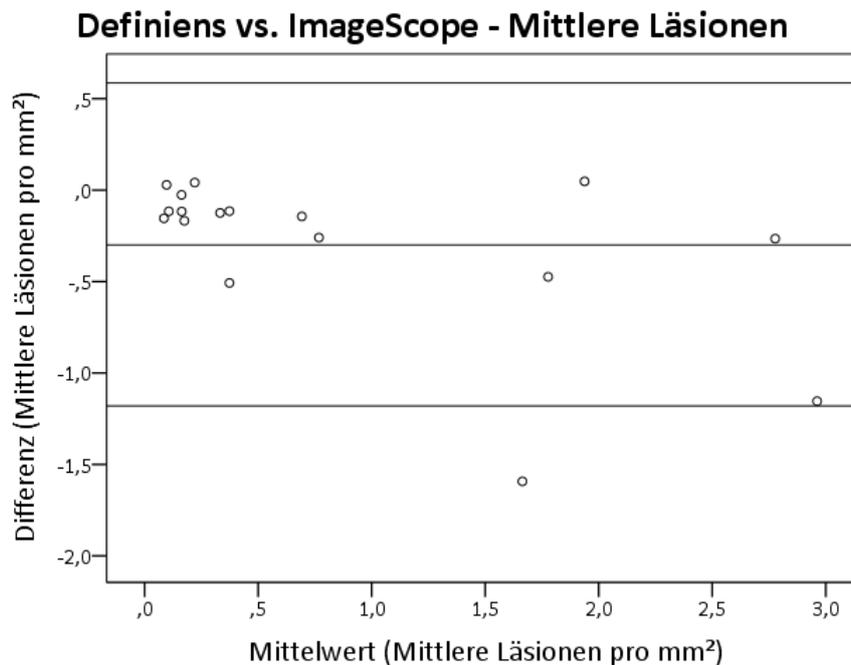


Abbildung 46: Bland-Altman-Plot Definiens vs. ImageScope – Mittlere Läsionen pro Fläche. Dargestellt sind die Mittelwerte der mittleren Läsionen pro mm² aus Definiens und ImageScope aufgetragen gegen die Differenzen der mittleren Läsionen pro mm² aus Definiens und ImageScope. Außerdem dargestellt die Verzerrung als Mittelwert über alle Differenzen (mittlere Linie) sowie die Übereinstimmungsgrenzen $d \pm 2 \times s$, wobei d den Mittelwert der Differenzen und s die Standardabweichung der Differenzen bezeichnet (obere und untere Linie).

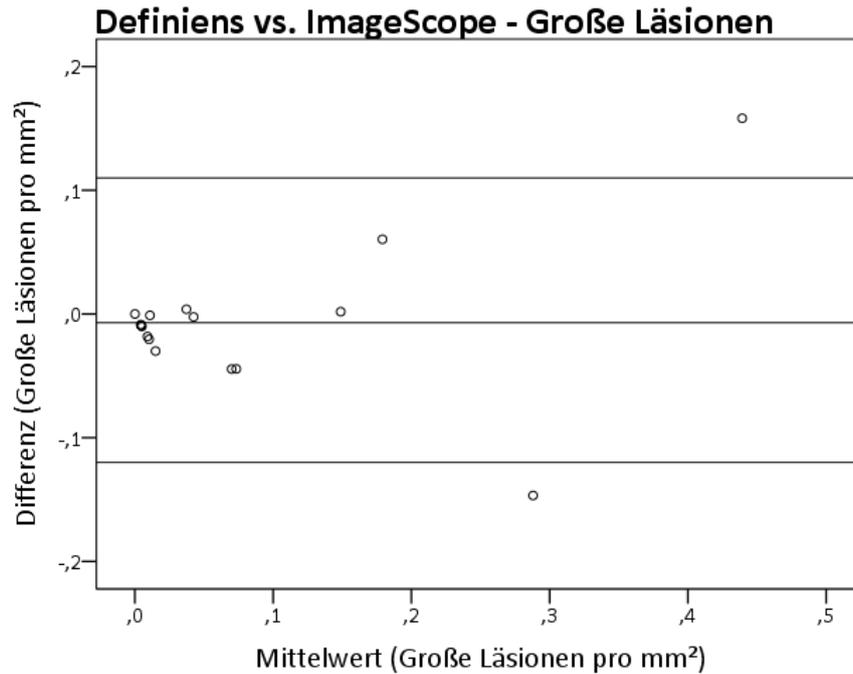


Abbildung 47: Bland-Altman-Plot Definiens vs. ImageScope – Große Läsionen pro Fläche. Dargestellt sind die Mittelwerte der großen Läsionen pro mm² aus Definiens und ImageScope aufgetragen gegen die Differenzen der großen Läsionen pro mm² aus Definiens und ImageScope. Außerdem dargestellt die Verzerrung als Mittelwert über alle Differenzen (mittlere Linie) sowie die Übereinstimmungsgrenzen $d \pm 2 \times s$, wobei d den Mittelwert der Differenzen und s die Standardabweichung der Differenzen bezeichnet (obere und untere Linie).

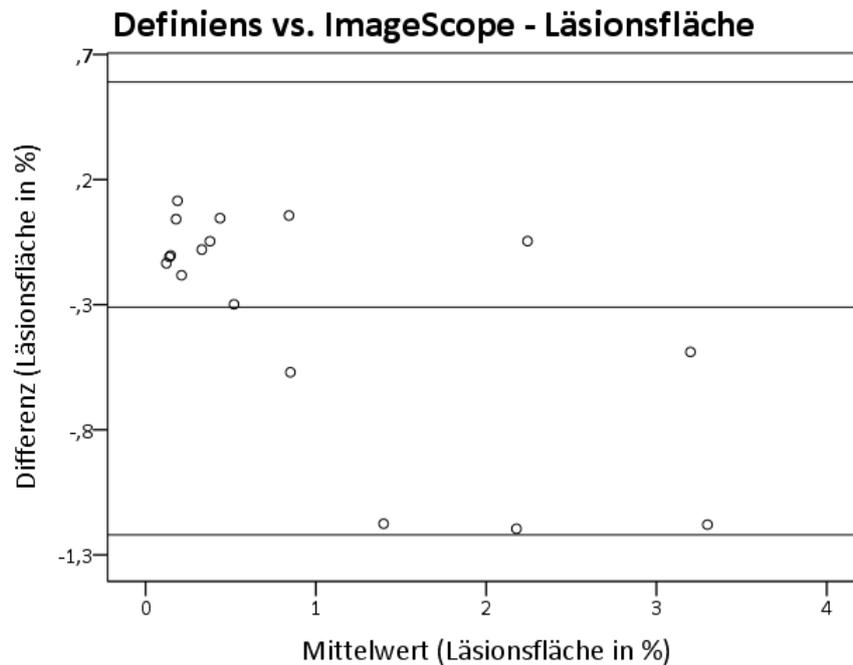


Abbildung 48: Bland-Altman-Plot Definiens vs. ImageScope – Läsionsfläche pro Gesamtfläche. Dargestellt sind die Mittelwerte der prozentualen Anteile der Läsionsfläche an der Gesamtfläche in Prozent aus Definiens und ImageScope aufgetragen gegen die Mittelwertdifferenzen der prozentualen Anteile der Läsionsfläche an der Gesamtfläche in Prozent aus Definiens und ImageScope. Außerdem dargestellt die Verzerrung als Mittelwert über alle Differenzen (mittlere Linie) sowie die Übereinstimmungsgrenzen $d \pm 2 \times s$, wobei d den Mittelwert der Differenzen und s die Standardabweichung der Differenzen bezeichnet (obere und untere Linie).

4.9 ImageScope – Korrelation der EMH Parameter mit dem Tumorzvolumen

Um den Verdacht zu bestätigen, dass das Ausmaß der EMH mit dem Tumorzvolumen korreliert, wurden die Mittelwerte der in ImageScope ermittelten EMH Parameter aus den zwei Scans jeder Maus - getrennt für die drei Gruppen - in Streudiagrammen gegen die entsprechenden Tumorzvolumina aufgetragen (siehe Abbildung 49 bis Abbildung 53) und die Korrelationskoeffizienten nach Spearman bestimmt (siehe Tabelle 23 bis Tabelle 27). Hier zeigen sich bei allen bestimmten Parametern hohe positive Korrelationskoeffizienten mit einer signifikanten Korrelation über alle drei Gruppen (siehe Tabelle 23 bis Tabelle 27). Betrachtet man die Korrelationskoeffizienten getrennt für die 3 Gruppen, so sieht man sowohl in der Kontrollgruppe wie auch in der bestrahlten Gruppe hohe positive Korrelationskoeffizienten (siehe Tabelle 23 bis Tabelle 27). Diese sind in der Kontrollgruppe bei den kleinen Läsionen pro Fläche und in der bestrahlten Gruppe bei allen Parametern statistisch signifikant (siehe Tabelle 23 bis Tabelle 27). Anders stellte es sich in der operierten Gruppe dar. Dort sind die Korrelationskoeffizienten niedrig und teilweise negativ (siehe Tabelle 23 bis Tabelle 27).

Tabelle 23: Korrelationskoeffizienten – Läsionen pro Fläche, ImageScope

ImageScope Läsionen pro Fläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.885**	0.771	1.000**	0.257
	Sig. (2- seitig)	<0.001	0.072		0.623

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

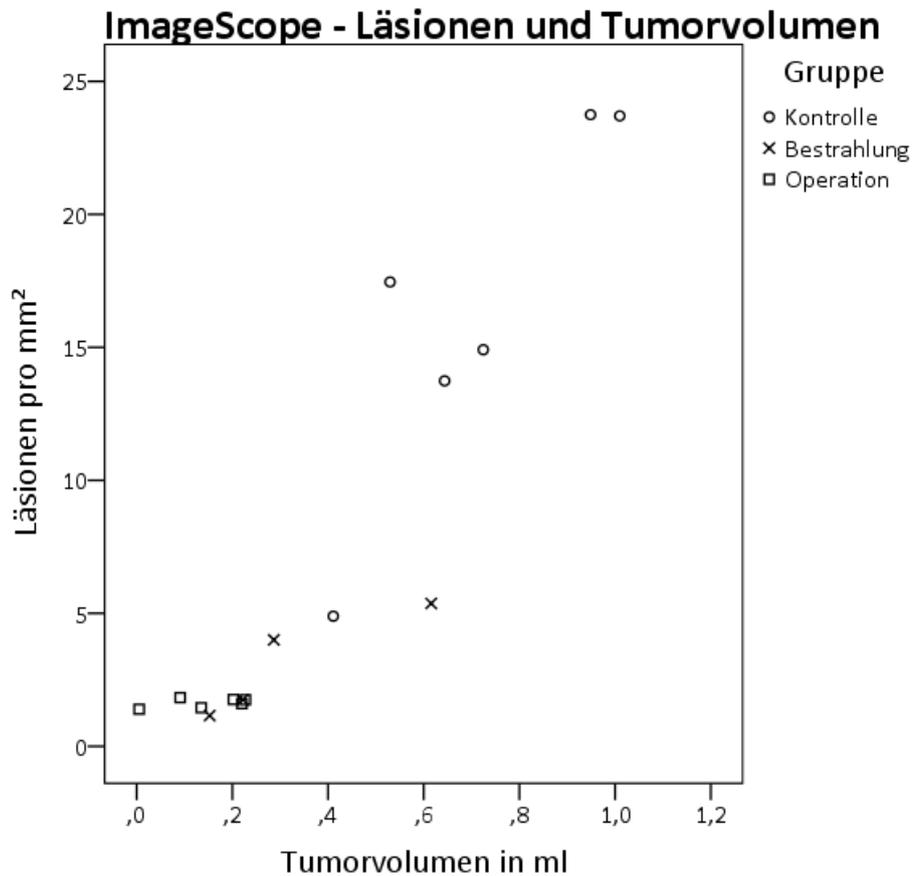


Abbildung 49: ImageScope – Korrelation Läsionen pro Fläche und Tumorzvolumen. Die Anzahl der Läsionen pro mm² der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen Anzahl der Läsionen und Tumorzvolumen über alle Gruppen. In der bestrahlten Gruppe signifikante Korrelation zwischen Anzahl der Läsionen und Tumorzvolumen (siehe Tabelle 23).

Tabelle 24: Korrelationskoeffizienten – Kleine Läsionen pro Fläche, ImageScope

ImageScope Kleine Läsionen pro Fläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.897**	0.829*	1.000**	0.257
	Sig. (2- seitig)	<0.001	0.042		0.623

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

ImageScope - Kleine Läsionen und Tumorzvolumen

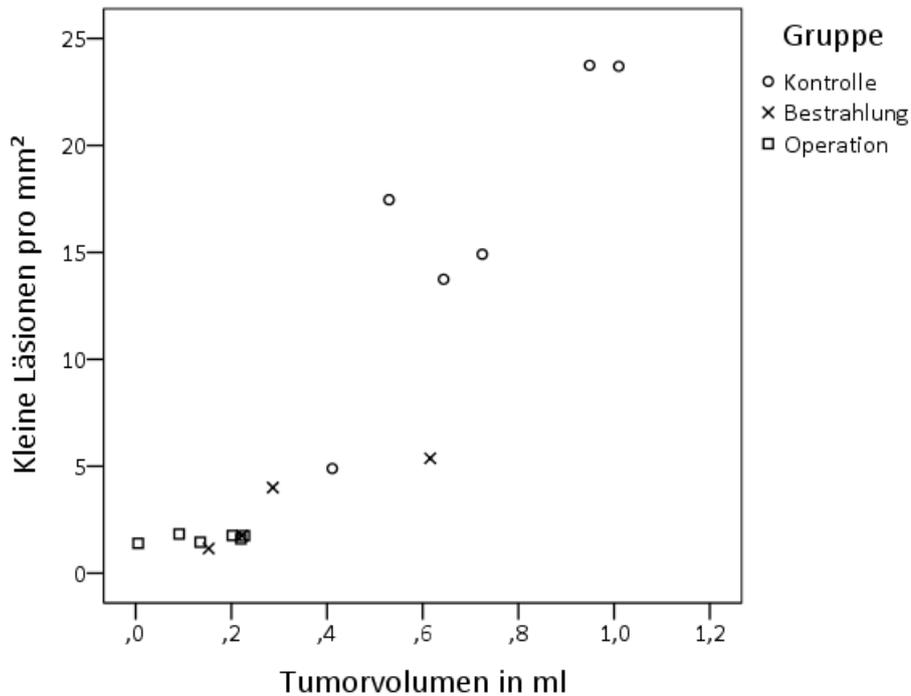


Abbildung 50: ImageScope – Korrelation kleine Läsionen pro Fläche und Tumorzvolumen. Die Anzahl der kleinen Läsionen pro mm² der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen Anzahl der kleinen Läsionen und Tumorzvolumen über alle Gruppen. In der Kontrollgruppe und in der bestrahlten Gruppe signifikante Korrelation zwischen Anzahl der kleinen Läsionen und Tumorzvolumen (siehe Tabelle 24).

Tabelle 25: Korrelationskoeffizienten – Mittlere Läsionen pro Fläche, ImageScope

Image Scope Mittlere Läsionen pro Fläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.897**	0.771	1.000**	0.314
	Sig. (2- seitig)	<0.001	0.072		0.544

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

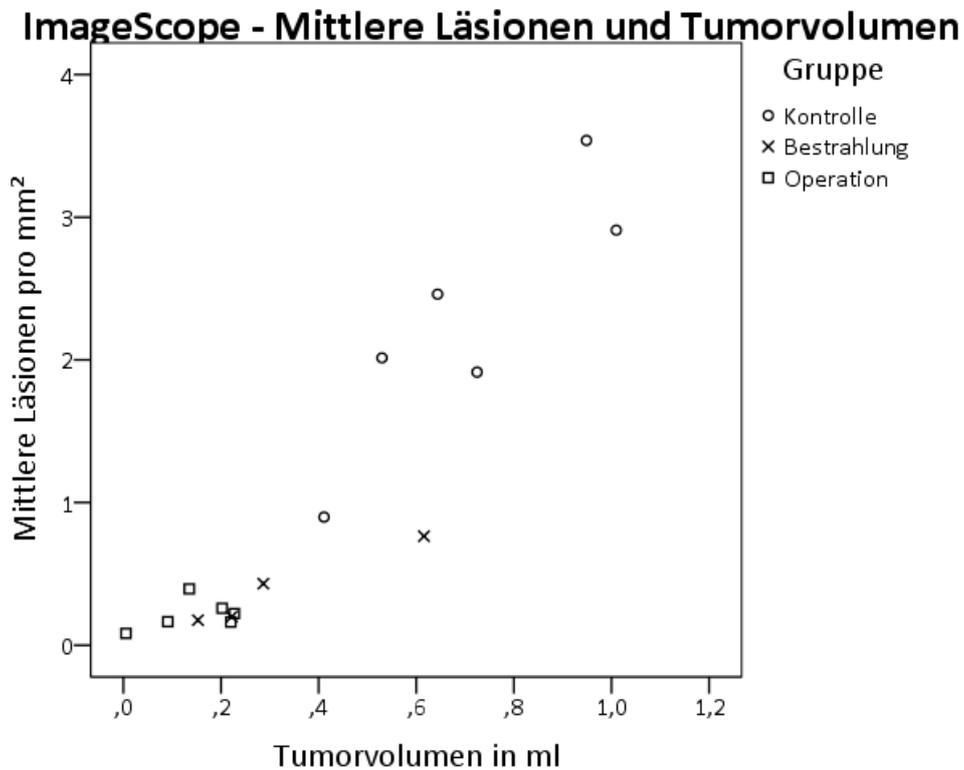


Abbildung 51: ImageScope – Korrelation mittlere Läsionen pro Fläche und Tumorzvolumen. Die Anzahl der mittleren Läsionen pro mm² der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen Anzahl der mittleren Läsionen und Tumorzvolumen über alle Gruppen. In der bestrahlten Gruppe signifikante Korrelation zwischen Anzahl der mittleren Läsionen und Tumorzvolumen (siehe Tabelle 25).

Tabelle 26: Korrelationskoeffizienten – Große Läsionen pro Fläche, ImageScope

ImageScope Große Läsionen pro Fläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.809**	0.771	0.800	-0.429
	Sig. (2- seitig)	<0.001	0.072	0.200	0.397

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

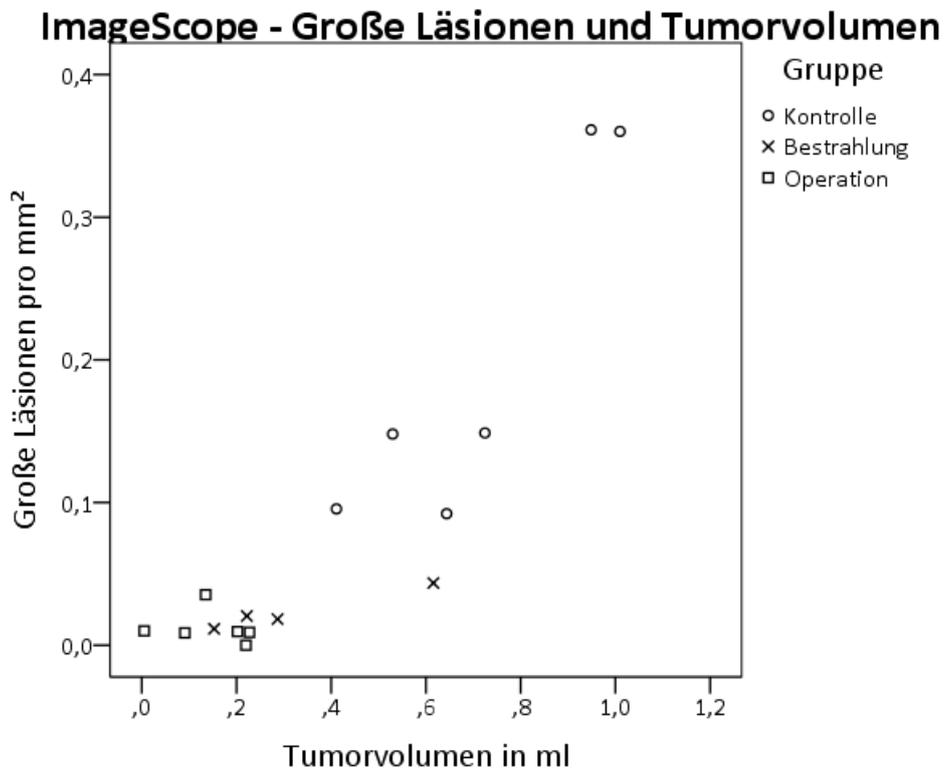


Abbildung 52: ImageScope – Korrelation große Läsionen pro Fläche und Tumorzvolumen. Die Anzahl der großen Läsionen pro mm² der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen Anzahl der großen Läsionen und Tumorzvolumen über alle Gruppen (siehe Tabelle 26).

Tabelle 27: Korrelationskoeffizienten – Läsionsfläche pro Gesamtfläche, ImageScope

ImageScope Läsionsfläche pro Gesamtfläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.906**	0.771	1.000**	0.429
	Sig. (2- seitig)	<0.001	0.072		0.397

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

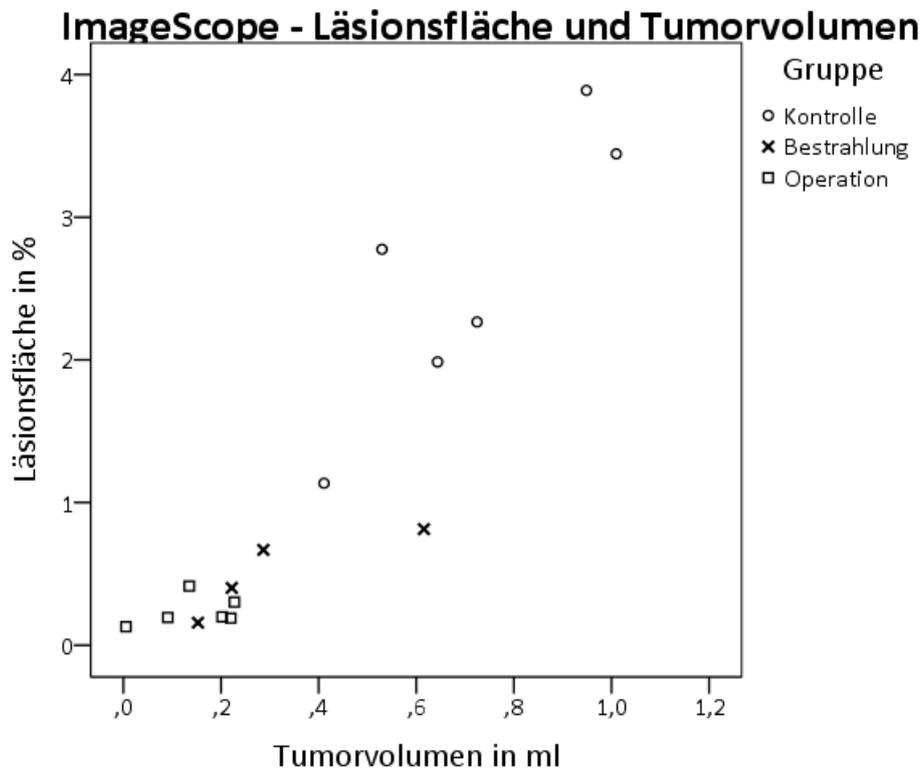


Abbildung 53: ImageScope – Korrelation Läsionsfläche pro Gesamtfläche und Tumorzvolumen. Der prozentuale Anteil der Läsionsfläche an der Gesamtfläche der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen dem prozentualen Anteil der Läsionsfläche an der Gesamtfläche und dem Tumorzvolumen über alle Gruppen. In der bestrahlten Gruppe signifikante Korrelation zwischen dem prozentualen Anteil der Läsionsfläche an der Gesamtfläche und dem Tumorzvolumen (siehe Tabelle 27).

4.10 Definiens – Korrelation der EMH Parameter mit dem Tumorzvolumen

Entsprechend dem Vorgehen bei ImageScope, sind in Abbildung 54 bis Abbildung 58 für alle drei Gruppen die Mittelwerte der durch Definiens ermittelten EMH Parameter aus zwei CD45 spezifisch gefärbten Scans jeder Maus in Streudiagrammen gegen die entsprechenden Tumorzvolumina aufgetragen. Genau wie bei den Streudiagrammen der Ergebnisse aus ImageScope zeigen sich bei allen bestimmten Parametern hohe positive Korrelationskoeffizienten mit einer signifikanten Korrelation über alle drei Gruppen (siehe Tabelle 28 bis Tabelle 32). Betrachtet man die Korrelationskoeffizienten getrennt für die 3 Gruppen, so sieht man sowohl in der Kontrollgruppe wie auch in der bestrahlten Gruppe hohe, oft signifikante positive Korrelationskoeffizienten (siehe Tabelle 28 bis Tabelle 32). Eine Ausnahme bildet die Gruppe der kleinen Läsionen pro Fläche. Dort ist der Korrelationskoeffizient in der Kontrollgruppe nur 0.543 (siehe Tabelle 29). In der operierten Gruppe waren die Korrelationskoeffizienten wiederum niedrig und teilweise negativ (siehe Tabelle 28 bis Tabelle 32).

Tabelle 28: Korrelationskoeffizienten – Läsionen pro Fläche, Definiens

Definiens Läsionen pro Fläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.774**	0.771	1.000**	-0.657
	Sig. (2- seitig)	<0.001	0.072		0.156

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

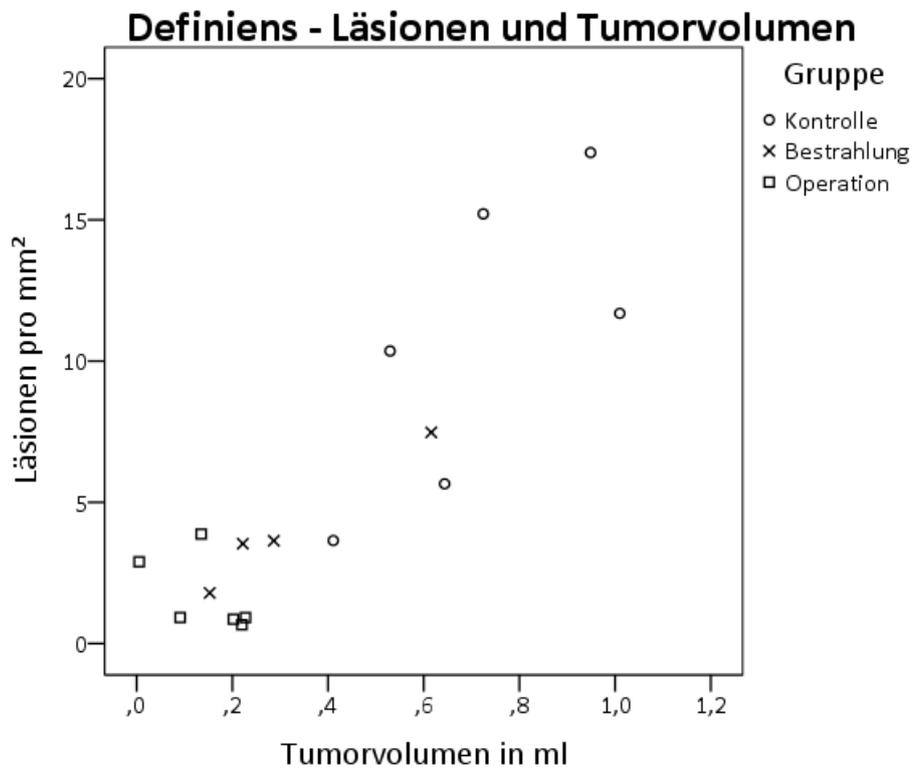


Abbildung 54: Definiens – Korrelation Läsionen pro Fläche und Tumorzvolumen. Die Anzahl der Läsionen pro mm² der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen Anzahl der Läsionen und Tumorzvolumen über alle Gruppen. In der bestrahlten Gruppe signifikante Korrelation zwischen Anzahl der Läsionen und Tumorzvolumen (siehe Tabelle 28).

Tabelle 29: Korrelationskoeffizienten – Kleine Läsionen pro Fläche, Definiens

Definiens Kleine Läsionen pro Fläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.747**	0.543	1.000**	-0.657
	Sig. (2- seitig)	0.001	0.266		0.156

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

Definiens - Kleine Läsionen und Tumorzvolumen

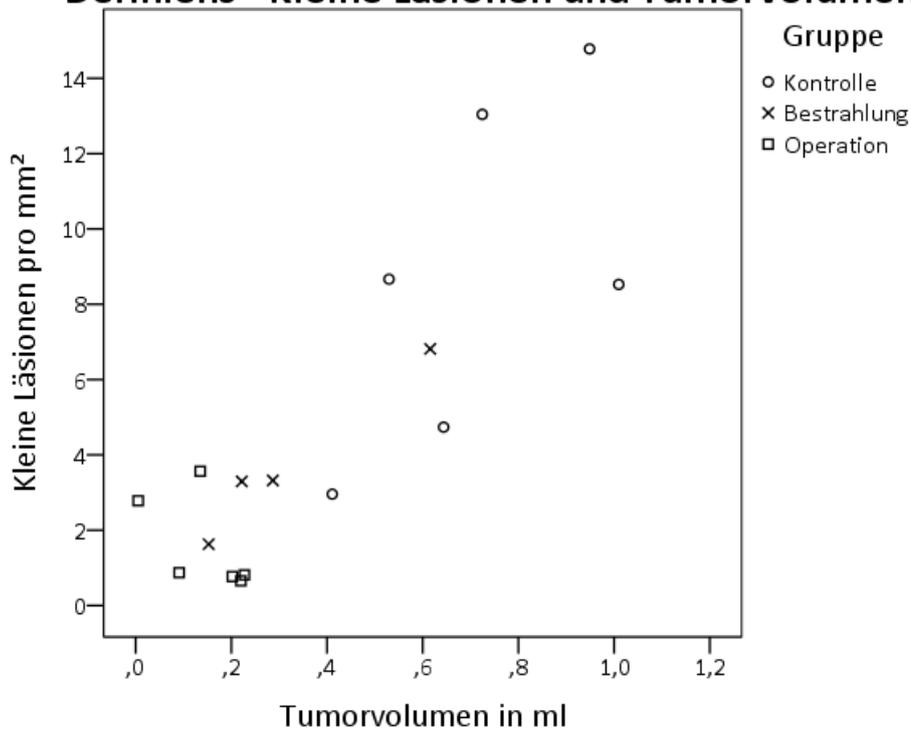


Abbildung 55: Definiens – Korrelation kleine Läsionen pro Fläche und Tumorzvolumen. Die Anzahl der kleinen Läsionen pro mm² der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen Anzahl der kleinen Läsionen und Tumorzvolumen über alle Gruppen. In der bestrahlten Gruppe signifikante Korrelation zwischen Anzahl der kleinen Läsionen und Tumorzvolumen (siehe Tabelle 29).

Tabelle 30: Korrelationskoeffizienten – Mittlere Läsionen pro Fläche, Definiens

Definiens Mittlere Läsionen pro Fläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.853**	0.943**	1.000**	-0.314
	Sig. (2- seitig)	<0.001	0.005		0.544

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

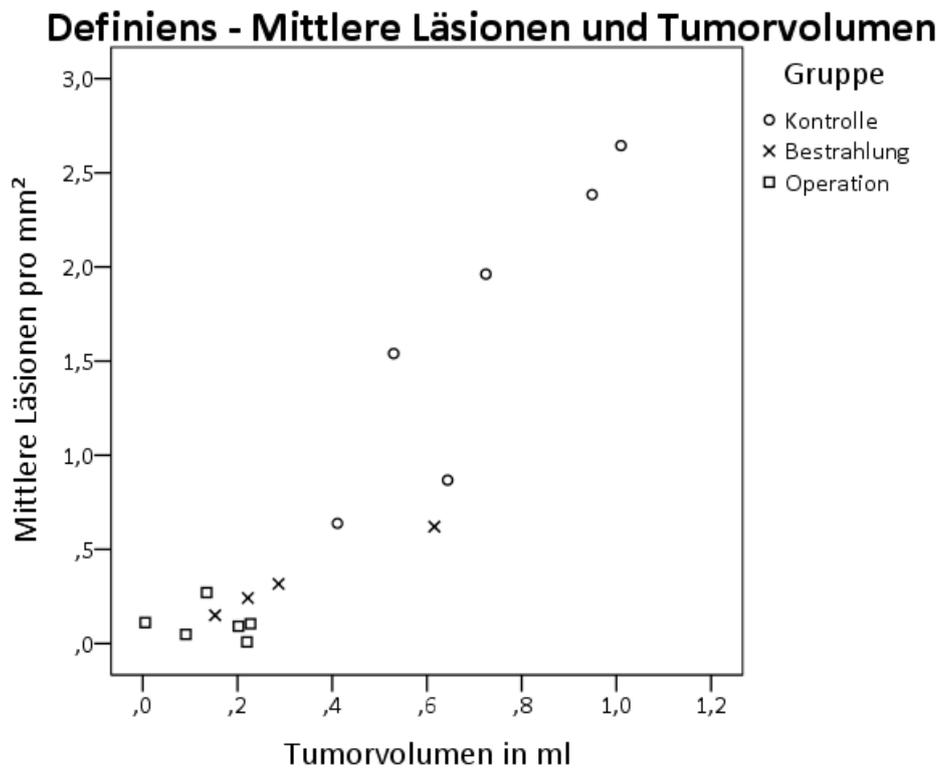


Abbildung 56: Definiens – Korrelation mittlere Läsionen pro Fläche und Tumorzvolumen. Die Anzahl der mittleren Läsionen pro mm² der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen Anzahl der mittleren Läsionen und Tumorzvolumen über alle Gruppen. In der Kontrollgruppe und in der bestrahlten Gruppe signifikante Korrelation zwischen Anzahl der mittleren Läsionen und Tumorzvolumen (siehe Tabelle 30).

Tabelle 31: Korrelationskoeffizienten – Große Läsionen pro Fläche, Definiens

Definiens Große Läsionen pro Fläche		Korrelationen			
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.792**	0.829*	0.316	-0.131
	Sig. (2- seitig)	<0.001	0.042	0.684	0.805

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

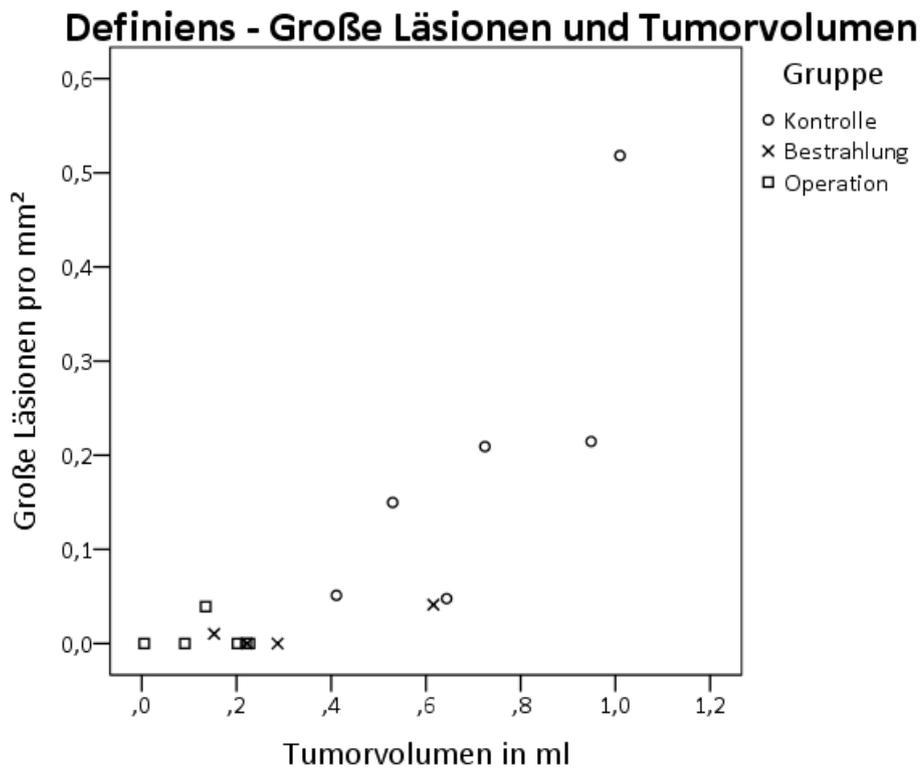


Abbildung 57: Definiens – Korrelation große Läsionen pro Fläche und Tumorzvolumen. Die Anzahl der großen Läsionen pro mm² der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen Anzahl der großen Läsionen und Tumorzvolumen über alle Gruppen. In der Kontrollgruppe signifikante Korrelation zwischen Anzahl der großen Läsionen und Tumorzvolumen (siehe Tabelle 31).

Tabelle 32: Korrelationskoeffizienten – Läsionsfläche pro Gesamtfläche, Definiens

Definiens Läsionsfläche pro Gesamtfläche	Korrelationen				
		Über alle Gruppen	Kontrolle	Bestrahlung	Operation
Tumorzvolumen	Spearman- Rho	0.829**	0.943**	1.000**	-0.314
	Sig. (2- seitig)	<0.001	0.005		0.544

* Korrelation bei Niveau 0.05 signifikant (zweiseitig).

** Korrelation bei Niveau 0.01 signifikant (zweiseitig).

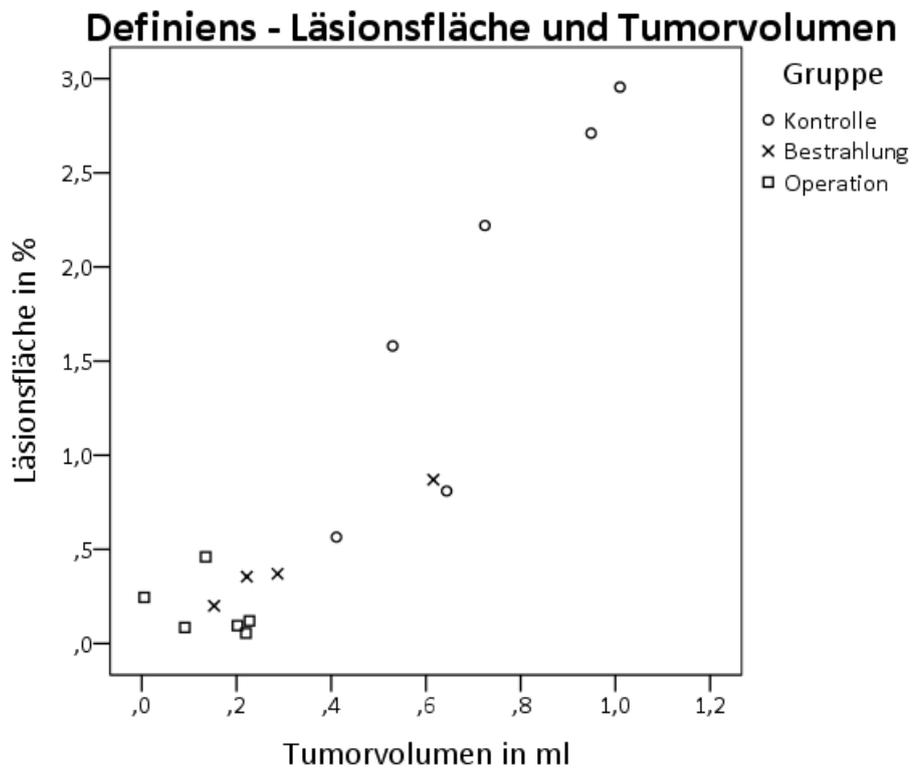


Abbildung 58: Definiens – Korrelation Läsionsfläche pro Gesamtfläche und Tumorzvolumen. Der prozentuale Anteil der Läsionsfläche an der Gesamtfläche der verschiedenen Gruppen aufgetragen gegen das zugehörige Tumorzvolumen. Signifikante Korrelation zwischen prozentualem Anteil der Läsionsfläche an der Gesamtfläche und Tumorzvolumen über alle Gruppen. In der Kontrollgruppe und in der bestrahlten Gruppe signifikante Korrelation zwischen prozentualem Anteil der Läsionsfläche an der Gesamtfläche und Tumorzvolumen (siehe Tabelle 32).

5 Diskussion

5.1 Scans der H.E. gefärbten Leberschnitte

Wie unter 2.4 erläutert, wurde die Ausbildung von EMH im 4T1 Mammakarzinommodell beschrieben (Tao et al., 2008) und diese als Quelle von MDSCs, also immunsuppressiver myeloischer Zellen, gesehen (Ilkovitch and Lopez, 2009). Bei den Läsionen in den H.E. Scans handelt es sich um EMH mit Bildung myeloischer Zellen, die viele morphologische Eigenschaften neutrophiler Granulozyten tragen (siehe 2.3.1). Dazu gehört der schmale Zytoplasmasaum sowie die Morphologie des Zellkerns, der fast das gesamte Zytoplasma einnimmt und teils gelappt und teils rundlich ist (vgl. Abbildung 7). Bei den Zellen, die einen rundlichen Zellkern haben, vermuten wir, dass es sich um unreife myeloische Zellen handelt.

5.2 Scans der immunhistochemisch gefärbten Leberschnitte

Dass fast alle Zellen in den Läsionen Ly6G/Ly6C positiv sind, bestätigte, dass es sich bei den EMH Herden in der Leber um MDSCs handelt. Wie unter Kapitel 2.3.2 erläutert, handelt es sich bei ihnen um Gr-1+CD11b+ Zellen (Gabilovich et al., 2007), die nochmals in Zellen mit einem granulozytären Phänotyp, die Ly6G Marker exprimieren und einen monozytären Phänotyp, die Ly6C Marker exprimieren, unterteilt werden können (Movahedi et al., 2008, Youn et al., 2008).

Ribechini et al. (2009) schrieben über Gr-1: Gr-1 besteht aus den Subtypen Ly6G und Ly6C und wird von frühen myeloischen Zellstufen exprimiert. Während der weiteren Differenzierung kommt es zu einer Veränderung des Expressionsmusters. Ly6G, aber nicht Ly6C werden in der weiteren Entwicklung der neutrophilen Granulozyten hoch reguliert, Ly6G und Ly6C dagegen bei Makrophagen und dendritischen Zellen herunter reguliert. Ly6C zeigt die stärkste Expression bei Monozyten im KM und eine mittlere Expression bei Granulozyten im KM. Ly6G/Ly6C (Gr-1) bindet mit hoher Affinität Ly6G und mit etwas niedrigerer Affinität Ly6C.

Aufgrund der Morphologie der Zellen in den Ansammlungen in der Leber liegt der Verdacht nahe, dass es sich hierbei hauptsächlich um den granulozytären Phänotyp handelt. Fast alle Zellen in den Läsionen sind Ly6G/Ly6C positiv (siehe Abbildung 12 bis Abbildung 14) und enthalten wohl myeloische Zellen aller Reifungsstufen. Worum es sich bei den Ly6G/Ly6C negativen Zellen handelt, bleibt ungeklärt. Möglicherweise sind es Ly6C positive monozytäre Zellen, an die der Antikörper mit unzureichender Affinität gebunden hat oder ausgereifte Makrophagen bzw. dendritische Zellen, die Ly6G/Ly6C nicht mehr exprimieren. Makrophagen müssten aber wiederum F4/80 positiv sein (Ribechini et al., 2009). Es gibt zwar einige F4/80 positive Zellen in den Läsionen, die sich aber eher zwischen den anderen hämatopoetischen Zellen befinden (siehe Abbildung 16). Diese könnten aus

MDSCs hervorgegangene Makrophagen sein, da MDSCs den immunsuppressiven M2 Typ bei Makrophagen induzieren können (Sinha et al., 2007). Alternativ könnte es sich bei den F4/80 positiven Zellen in den Läsionen um eine Interaktion zwischen MDSCs und ortsansässigen Makrophagen, sogenannten Kupffer'schen Sternzellen, handeln. Wie unter 2.3.2 erläutert, wurde gezeigt, dass Kupffer'sche Sternzellen durch vermehrte Expression des T-Zell supprimierenden Oberflächenmoleküls PD-L1 die T-Zell Antwort unterdrücken und dass die hepatischen MDSCs diesen Mechanismus unterstützen (Ilkovitch and Lopez, 2009).

Bei den F4/80 positiven Zellen, die die gesamte Leber durchziehen (siehe Abbildung 15), handelt es sich wahrscheinlich um Lebermakrophagen (Ilkovitch and Lopez, 2009, Lüllmann-Rauch, Leber, 2006).

Warum sich NK- und T-Zellen in den EMH Herden finden lassen ist unklar. Wie bereits erwähnt, können MDSCs die Toxizität der NK-Zellen herunter regulieren (Liu et al., 2007) und T-Zellen hemmen (Talmadge, 2007) bzw. immunsuppressive regulatorische T-Zellen induzieren (Huang et al., 2006). Dass sich Makrophagen, NK-Zellen und T-Zellen in den EMH-Läsionen finden lassen, deutet darauf hin, dass dort eine Interaktion zwischen MDSCs und reifen Immunzellen stattfindet.

Es stellt sich nun die Frage, warum es bei Tumor tragenden Mäusen, aber nicht bei an Krebs erkrankten Menschen zur Ausbildung von EMH in der Leber kommt. Ein Grund ist, dass das Knochenmark von Mäusen schon im gesunden Zustand fast vollständig mit blutbildendem rotem Knochenmark ausgefüllt ist und es bei der Maus bei Stress, beispielsweise durch Hämolyse, viel schneller zu EMH mit Erythropoese in Leber und Milz kommt (Everds, 2004). Dasselbe Prinzip könnte auch für die durch den Primärtumor hervorgerufenen Myelopoese gelten. Desweiteren ist im Mausversuch der Tumor im Verhältnis zum Gesamtorganismus viel größer, als das beim Menschen aufgrund der Früherkennung meist der Fall ist (Diaz-Montero et al., 2009). Bestätigt wird das durch die Beobachtung, dass im Vergleich zum Mausmodell die relative Anzahl der MDSCs an den Immunzellen bei Menschen deutlich geringer ist (Diaz-Montero et al., 2009). Wahrscheinlich findet also die Bildung von MDSCs beim Menschen ausschließlich im Knochenmark statt.

5.3 Scans der H.E. und immunhistochemisch gefärbten Tumorschnitte

Bei der Betrachtung der H.E. gefärbten Scans der Primärtumore fällt eine ausgeprägte Nekrose vor allem in den zentralen Tumorarealen, teilweise aber auch peripher auf (siehe Abbildung 9). Tao et al. (2008) beschrieben, dass diese charakteristisch für das 4T1 Modell und für die Metastasierung förderlich ist. Desweiteren beobachteten sie eine ausgeprägte Entzündungsreaktion mit Infiltration von Immunzellen in den Tumoren. Das ließ sich in meinen immunhistochemischen Färbungen bestätigen. In diesen sehen wir vor allem eine massive Zahl CD45 sowie Ly6G/Ly6C positiver Entzündungszellen im gesamten Tumor und im umgebenden Bindegewebe (siehe Abbildung 19 bis

Abbildung 22). Tao et al. (2008) nannten als mögliche Ursache dafür die ausgeprägte Nekrose, die eine Entzündungsreaktion nach sich zieht. Zusätzlich sehen sie die Infiltration des Tumors durch Entzündungszellen als Folge der Stimulation durch sezernierte Chemokine der Tumorzellen. Neben den Neutrophilen lassen sich auch zahlreiche Makrophagen, NK-Zellen und T-Zellen in den Tumoren (siehe Abbildung 23 bis Abbildung 26) nachweisen. Bei den Makrophagen könnte es sich um die unter 2.3.2 beschriebenen TAMs handeln.

5.4 Vergleich des Ausmaßes an EMH in den verschiedenen Gruppen

Ein zentraler Punkt dieser Arbeit ist das Ausmaß an EMH in den verschiedenen Behandlungsgruppen zu vergleichen. Dazu wurden zwei verschiedene Verfahren, ImageScope und Definiens, verwendet.

Kernaussage beim Vergleich des Ausmaßes der EMH in den verschiedenen Gruppen ist, dass die EMH in der Kontrollgruppe am stärksten ausgeprägt ist. Fast immer und in beiden angewandten Verfahren lassen sich statistisch signifikante Unterschiede zwischen Kontrollgruppe und bestrahlter Gruppe und Kontrollgruppe und operierter Gruppe finden. Dies gilt nicht für bestrahlte und operierte Gruppe. Des Weiteren kann man sagen, dass es sich bei der EMH in der Leber vor allem um kleine Läsionen $<2.000 \mu\text{m}^2$ handelt.

Bei der Quantifizierung von EMH mittels ImageScope sollte sichergestellt werden, dass die Selektion von bestimmten Bereichen in Form von Rechtecken zu keinem Bias führt. Deshalb wurde an je einem Scan pro Behandlungsgruppe eine Quantifizierung der EMH anhand von Ellipsen durchgeführt. In einem Bland-Altman-Plot wurden die Mittelwerte der Messwerte Ellipse und Rechteck gegen die Differenz aus Ellipse und Rechteck aufgetragen. Die Werte aus der Methode Ellipse ergeben in 95% der Fälle einen Wert, der um 3.6 größer bzw. um 4.3 kleiner ist als der Wert aus der Methode Rechteck. Dabei muss man bedenken, dass in unserem Fall unterschiedlich große Parameter in demselben Diagramm dargestellt wurden. So sind die Werte Läsionen pro mm^2 und kleine Läsionen pro mm^2 in der Kontrollgruppe viel höher als die anderen und werden durch die beiden Punkte ganz rechts im Diagramm dargestellt. Man sieht, dass es bei diesen hohen Werten tendenziell auch höhere Differenzen gibt, was Auswirkungen auf die Übereinstimmungsgrenzen hat. So wären Unterschiede von 3.6 bzw. 4.3 bei den anderen Ergebnissen der Kontrollgruppe oder bei allen Ergebnissen der bestrahlten bzw. operierten Gruppe nicht akzeptabel. Betrachtet man das Diagramm (Abbildung 33), so sieht man, dass das nicht der Fall ist und es bei den niedrigen Werten auch niedrigere Differenzen gibt. Außerdem lässt sich feststellen, dass sich der Mittelwert über alle Differenzen nahe der 0 bewegt, was daran liegt, dass keines der beiden Verfahren immer höhere Werte ergibt als das andere sondern dass die Differenzen abwechselnd positiv und negativ sind. Zusammenfassend lässt sich also sagen, dass die Ermittlung der EMH in den H.E. gefärbten Leberscans durch die Betrachtung von rechteckigen Ausschnitten sich nicht gravierend von der

Betrachtung elliptischer Ausschnitte unterscheidet und man vermuten kann, dass keine relevanten Bereiche im Scan ausgespart werden.

Auch in der Leber der gesunden Maus gibt es relevante EMH (siehe Abbildung 28 bis Abbildung 32 sowie Abbildung 34 bis Abbildung 38), was so nicht erwartet wurde. Youn et al. (2008) schrieben, wie unter 2.3.2 bereits angeführt, dass bei nicht-tumortragenden Mäusen MDSCs nur in geringem Maße im Knochenmark und in der Milz zu finden sind und dort nicht immunsuppressiv wirken. Warum es in unserem Fall zu EMH in der Leber kam, ist nicht klar. Möglicherweise wurde das Immunsystem der Maus durch andere Stressoren wie z.B. eine Infektion stimuliert.

5.5 Vergleich der angewandten Verfahren ImageScope und Definiens

Eine weitere Zielsetzung meiner Arbeit war herausfinden, ob die beiden angewendeten Verfahren, ImageScope und Definiens, vergleichbare Ergebnisse lieferten.

Wie unter 4.8 beschrieben, sind ähnliche Tendenzen bei beiden Verfahren zu beobachten. In Bland-Altman-Plots wurden die Mittelwerte der Messpaare gegen ihre Differenzen aufgetragen (siehe Abbildung 44 bis Abbildung 48). Wiederum sieht man, dass größere Werte mit größeren Differenzen einhergehen, was den Erwartungen entspricht. Diese liegen vor allem in der Kontrollgruppe vor und entsprechen dem schon im Boxplot zu sehenden Unterschied zwischen ImageScope und Definiens, wobei Definiens kleinere Werte ergibt. Die Verzerrung liegt bei allen Diagrammen im negativen Bereich (siehe Tabelle 33), was aber zumindest teilweise durch die hohen negativen Differenzen bei großen Werten erklärt werden kann. Diese beeinflussen den Mittelwert der Differenzen besonders und machen ihn negativer. Kleine Werte liegen eher oberhalb der Verzerrungslinie und teilweise auch im positiven Bereich. Vor allem im kleineren Bereich ergibt Definiens auch größere Werte als ImageScope.

Tabelle 33: Vergleich von Definiens und ImageScope – Parameter aus Bland-Altman-Plot

Übereinstimmungsgrenzen	Untere	Obere	Mittelwert der Differenzen
Anzahl der Läsionen pro Fläche	-10.037	6.796	-1.621
Anzahl der kleinen Läsionen pro Fläche	-8.977	6.534	-1.221
Anzahl der mittleren Läsionen pro Fläche	-1.179	0.580	-0.300
Anzahl der großen Läsionen pro Fläche	-0.125	0.110	-0.007
Läsionsfläche pro Gesamtfläche in %	-1.220	0.590	-0.315

Möglicherweise führen auch schwach CD45 positive Läsionen, die von Definiens nicht als positiv gewertet wurden, zu den oben beschriebenen Tendenzen (vergleiche Abbildung 11). Betrachtet man die Übereinstimmungsgrenzen in Tabelle 33, so lässt sich sagen, dass die Unterschiede in den beiden Verfahren akzeptabel sind. Sowohl ImageScope als auch Definiens sind somit geeignete Verfahren zur Quantifizierung von EMH in der Leber. Dabei haben beide Vor- und Nachteile.

Gegen die Methode nach Definiens spricht, dass die Durchführung immunhistochemischer Färbungen sehr zeitaufwendig, fehleranfällig und teuer ist. An diesem Punkt sollte auch erwähnt werden, dass der Mensch viel besser und zuverlässiger in der Lage ist bestimmte histologische Eigenschaften zu erkennen und zu bewerten als ein Algorithmus. Auch wenn Definiens ein entwicklungstechnisch fortschrittliches Programm ist, so ist man als Nutzer immer gezwungen Kompromisse einzugehen. Es werden beispielsweise bestimmte Bereiche, die man selbst als positiv gewertet hätte, nicht erkannt. Will man, dass diese positiv gewertet werden, so muss man in Kauf nehmen, dass auch viele unspezifisch gefärbte Bereiche in die Ergebnisse einfließen. Die Analyse selbst ist, wenn man sich mit dem Programm etwas vertraut gemacht hat, weniger zeitaufwendig als mit ImageScope. Ein weiterer Vorteil von Definiens ist, dass der gesamte Scan beurteilt wird. Die manuelle Markierung einzelner Läsionen in ImageScope scheint die zuverlässigere Variante zu sein, da der Nutzer jede Läsion einzeln als solche erkennt und markiert. Ein weiterer Vorteil bei dieser Methode ist, dass man die H.E. Färbung günstig und schnell durchführen kann und sie wenig fehleranfällig ist. Dafür ist die Untersuchung selbst umso zeitaufwendiger und in der Praxis nicht realisierbar. Hinzu kommt, dass dadurch, dass nur bestimmte Bereiche der Scans betrachtet wurden, möglicherweise Fehler entstanden sind.

Elegant ist also eine Möglichkeit H.E. gefärbte Scans automatisiert zu analysieren, da Definiens für die Analyse H.E. gefärbter Scans nicht geeignet ist. Peter et al. (2016, 2014) fanden anhand meiner H.E. gefärbten Leberscans einen Algorithmus, der eine halbautomatische Analyse ermöglicht. Zu Beginn erlernt das Programm anhand von Labels das Aussehen der Läsionen. Das ist die Voraussetzung dafür, dass es später dem Nutzer mögliche interessante Bereiche aus dem Scan vorschlägt. Der Nutzer kann entscheiden, ob diese Bereiche gezählt werden sollen oder nicht. Sind diese Vorschläge falsch, kann das Programm in Verlauf sogar hinzulernen und die Treffsicherheit erhöhen. Der große Vorteil dieser Methode ist, dass der gesamte Scan trotz seiner Größe in relativ kurzer Zeit durchsucht werden kann und der Nutzer entscheidet, ob ein Bereich als positiv gewertet werden soll oder nicht.

5.6 Zusammenhang zwischen EMH, Therapieform und Tumorzvolumen

Wie oben bereits dargestellt, fanden wir signifikante Unterschiede im Ausmaß der EMH jeweils zwischen Kontrollgruppe und bestrahlter Gruppe bzw. operierter Gruppe für fast alle ermittelten Parameter. Es ist bekannt, dass systemische Level von MDSCs und Tumormasse einen positiven Zusammenhang haben, was wohl daher kommt, dass Tumore umso mehr stimulierende Faktoren sezernieren, desto größer sie sind (Ilkovitch and Lopez, 2009). Diaz-Montero et al. (2009) untersuchten den Zusammenhang zwischen Tumorstadium und zirkulierenden MDSCs beim Menschen und fanden Hinweise auf eine positive Korrelation zwischen der Höhe der MDSCs und dem Tumorstadium bei metastasierten Tumoren. Daraus folgerten wir, dass auch bei unseren Versuchen das Ausmaß der EMH wahrscheinlich vom Volumen des Primärtumors abhängen würde. Dazu haben wir Streudiagramme erstellt, in denen die EMH Parameter gegen das Tumorzvolumen des jeweiligen Primärtumors aufgetragen wurden. Um die Stärke ihres Zusammenhangs zu ermitteln, wurden die Korrelationskoeffizienten nach Spearman bestimmt und auf Signifikanz getestet. Daraus sollten Hinweise dafür gefunden werden, ob man das Ausmaß der EMH in einer Maus anhand ihres Tumorzvolumens vorhersagen kann.

Die hohen positiven Korrelationskoeffizienten in der Kontrollgruppe wie auch in der bestrahlten Gruppe bei ImageScope und Definiens bestätigen die Beobachtung von Ilkovitch and Lopez (2009), dass das Ausmaß der EMH vom Primärtumorzvolumen abhängt. Aus den beschriebenen Ergebnissen ergibt sich die Annahme, dass es in der Kontrollgruppe und der bestrahlten Gruppe, nicht aber in der operierten Gruppe einen Zusammenhang zwischen Tumorzvolumen und Ausmaß der EMH in der Leber gibt. Dass bei den Ergebnissen aus Definiens in der Gruppe der kleinen Läsionen pro Fläche der Korrelationskoeffizient in der Kontrollgruppe nur 0.543 ist, könnte daher kommen, dass unspezifisch gefärbte Bereiche ungewollt in die Analyse eingeflossen sind. Es war vor allem bei den kleinen Läsionen schwierig, eine gute Trennung zwischen Artefakten und Läsionen zu erlangen.

Zusammenfassend kann man sagen, dass alle Ergebnisse aus ImageScope und Definiens in der Kontrollgruppe und der bestrahlten Gruppe einen positiven Zusammenhang zwischen dem Primärtumorzvolumen und dem Ausmaß an EMH zeigen. Das heißt je größer der Tumor ist, desto mehr EMH findet statt. Das lässt sich wahrscheinlich dadurch erklären, dass ein Tumor umso mehr das Immunsystem stimulierende Faktoren sezerniert desto größer er ist. Bei den operierten Mäusen gibt es keinen Anhaltspunkt dafür, dass es einen Zusammenhang zwischen Tumorzvolumen und Anteil an EMH am Lebergewebe gibt. Das lässt sich wahrscheinlich dadurch erklären, dass die Organe dieser Mäuse erst zwei Wochen nach Tumorentfernung entnommen wurden. In dieser Gruppe ist der stimulierende Effekt des Primärtumors also seit einiger Zeit ausgeblieben. Ob sich die EMH wieder zurückgebildet hat oder auf dem Niveau geblieben ist, auf dem sie zwei Wochen nach

Zellimplantation war, lässt sich im Nachhinein nicht mehr feststellen, da keine Lebern zu diesem Zeitpunkt entnommen wurden.

5.7 Ausblick

Eine zentrale Frage, die es noch zu beantworten gilt, ist, ob neben dem Tumolvolumen noch andere Faktoren das Ausmaß der EMH in der Leber beeinflussen. Es wäre beispielsweise denkbar, dass bestrahlte Tumore bei gleichem Volumen wie unbehandelte Tumore weniger EMH auslösen, da sie nach der Bestrahlung sekretorisch weniger aktiv sind. Um das herauszufinden, haben wir in den Streudiagrammen Ausgleichsgeraden durch die verschiedenen Gruppen gelegt (Daten nicht gezeigt). Dabei ergaben sich in der bestrahlten Gruppe tatsächlich geringere Steigungen. Da aber die Kontrolltumore in den Diagrammen rechts und die bestrahlten Tumore links liegen (vergleiche Abbildung 49 bis Abbildung 58) und es fast keine Überschneidungen ihrer Bereiche im Diagramm gibt und generell nur eine sehr geringe Fallzahl vorliegt, sind diese Daten statistisch nicht zu verwerten. Sie geben aber Anlass für zukünftige Untersuchungen. Barcellos-Hoff et al. (2005) schrieben bezüglich des Einflusses von Bestrahlung auf das Tumormilieu, dass sie zu akuten und andauernden Veränderungen der löslichen und nicht-löslichen Bestandteile im Gewebe führe. Das hat nicht nur positive Effekte. Inflammation als Folge von Bestrahlung spielt bei der Strahlenempfindlichkeit des Tumors und des Mikromilieus eine zentrale Rolle und kann zu Strahlenresistenz und Tumorprogression führen (Multhoff and Radons, 2012).

Interessante Ansätze für zukünftige Arbeiten sind beispielsweise die Erforschung des für EMH und Metastasierung verantwortlichen Mikromilieus des Tumors und inwiefern sich dieses beispielsweise nach Bestrahlung verändert. Tao et al. (2008) nannten diesbezüglich einige Zytokine und Botenstoffe wie CSF, Angiogenese Stimulatoren, Chemokine, Proteasen und akute Phase Proteine, die EMH und Metastasierung begünstigen. Deren Zusammensetzung und Menge könnten weitere Erkenntnisse für das Zustandekommen der EMH in der Leber bringen.

6 Zusammenfassung

Fokus dieser Arbeit war die EMH als Quelle immunsuppressiver myeloischer Zellen in der Leber tumortragender Mäuse. Wir bestätigten die Beobachtungen vieler anderer Arbeitsgruppen, dass es sich dabei vor allem um Ly6G/Ly6C (Gr-1) positive Zellen, also Zellen der myeloischen Reihe, handelt. Außerdem konnten wir zeigen, dass auch andere Immunzellen wie Makrophagen, T-Zellen und NK-Zellen in den Läsionen vertreten sind. Zwei verschiedene Verfahren zur Quantifizierung EMH in der Leber wurden angewendet: Zum einen die manuelle Markierung von Läsionen in H.E. gefärbten Scans in Aperio ImageScope, zum anderen die automatisierte Quantifizierung immunhistochemisch

CD45 spezifisch gefärbter Scans in Definiens. Die beiden Verfahren wurden verglichen und ihre Vor- und Nachteile beleuchtet. Eine Methode zur Quantifizierung von EMH in der Leber anhand meiner H.E. Scans wurde in Aussicht gestellt (Peter et al., 2017, Peter et al., 2014). Es wurde ein positiver Zusammenhang zwischen dem Tumorzvolumen und der EMH in der Kontrollgruppe und in der bestrahlten Gruppe gezeigt. Außerdem ließen sich Hinweise dafür finden, dass es in der Kontrollgruppe zu besonders extensiver EMH kam. Möglicherweise ist ein hohes Maß an stimulierenden Faktoren nötig, um solch große Bereiche an EMH zu bilden, was nur bei den Kontrollmäusen der Fall war. In der operierten Gruppe scheint es keinen Zusammenhang zwischen EMH und Primärtumorzvolumen zu geben und es gibt Hinweise darauf, dass sich die EMH in Abwesenheit des Tumors zurückbilden kann.

Literaturverzeichnis

- ADAMS, J. M. 2003. Ways of dying: multiple pathways to apoptosis. *Genes Dev*, 17, 2481-2495.
- ADAMS, J. M. & CORY, S. 2007. The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene*, 26, 1324-1337.
- AGUIRRE-GHISO, J. A. 2007. Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer*, 7, 834-846.
- ARTANDI, S. E. & DEPINHO, R. A. 2010. Telomeres and telomerase in cancer. *Carcinogenesis*, 31, 9-18.
- BARCELLOS-HOFF, M. H., PARK, C. & WRIGHT, E. G. 2005. Radiation and the microenvironment - tumorigenesis and therapy. *Nat Rev Cancer*, 5, 867-875.
- BERDASCO, M. & ESTELLER, M. 2010. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev Cell*, 19, 698-711.
- BERGERS, G. & BENJAMIN, L. E. 2003. Tumorigenesis and the angiogenic switch. *Nat Rev Cancer*, 3, 401-410.
- BERGERS, G. & SONG, S. 2005. The role of pericytes in blood-vessel formation and maintenance. *Neuro Oncol*, 7, 452-464.
- BERNIER, J., HALL, E. J. & GIACCIA, A. 2004. Radiation oncology: a century of achievements. *Nat Rev Cancer*, 4, 737-747.
- BERX, G. & VAN ROY, F. 2009. Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harb Perspect Biol*, 1, a003129.
- BHOWMICK, N. A., NEILSON, E. G. & MOSES, H. L. 2004. Stromal fibroblasts in cancer initiation and progression. *Nature*, 432, 332-337.
- BIGANZOLI, E., CORADINI, D., AMBROGI, F., GARIBALDI, J. M., LISBOA, P., SORIA, D., GREEN, A. R., PEDRIALI, M., PIANTELLI, M., QUERZOLI, P., DEMICHELI, R., BORACCHI, P., NENCI, I., ELLIS, I. O. & ALBERTI, S. 2011. p53 status identifies two subgroups of triple-negative breast cancers with distinct biological features. *Jpn J Clin Oncol*, 41, 172-179.
- BINGLE, L., BROWN, N. J. & LEWIS, C. E. 2002. The role of tumour-associated macrophages in tumour progression: implications for new anticancer therapies. *J Pathol*, 196, 254-265.
- BISWAS, S. K., GANGI, L., PAUL, S., SCHIOPPA, T., SACCANI, A., SIRONI, M., BOTTAZZI, B., DONI, A., VINCENZO, B., PASQUALINI, F., VAGO, L., NEBULONI, M., MANTOVANI, A. & SICA, A. 2006. A distinct and unique transcriptional program expressed by tumor-associated macrophages (defective NF-kappaB and enhanced IRF-3/STAT1 activation). *Blood*, 107, 2112-2122.
- BLASCO, M. A. 2005. Telomeres and human disease: ageing, cancer and beyond. *Nat Rev Genet*, 6, 611-622.
- BÖCKER, D., HEITZ, MOCH 2008. Repetitorium Pathologie.
- BOEHM, U., KLAMP, T., GROOT, M. & HOWARD, J. C. 1997. Cellular responses to interferon-gamma. *Annu Rev Immunol*, 15, 749-795.
- BOEKHOUT, A. H., BEIJNEN, J. H. & SCHELLENS, J. H. M. 2011. Trastuzumab. *Oncologist*, 16, 800-810.
- CAVALLARO, U. & CHRISTOFORI, G. 2004. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nat Rev Cancer*, 4, 118-132.
- CHAE, B. J., BAE, J. S., LEE, A., PARK, W. C., SEO, Y. J., SONG, B. J., KIM, J. S. & JUNG, S. S. 2009. p53 as a specific prognostic factor in triple-negative breast cancer. *Jpn J Clin Oncol*, 39, 217-224.
- CHAFFER, C. L., THOMPSON, E. W. & WILLIAMS, E. D. 2007. Mesenchymal to epithelial transition in development and disease. *Cells Tissues Organs*, 185, 7-19.
- COGHLIN, C. & MURRAY, G. I. 2010. Current and emerging concepts in tumour metastasis. *J Pathol*, 222, 1-15.
- COLLADO, M. & SERRANO, M. 2010. Senescence in tumours: evidence from mice and humans. *Nat Rev Cancer*, 10, 51-57.
- COLOTTA, F., RE, F., POLENTARUTTI, N., SOZZANI, S. & MANTOVANI, A. 1992. Modulation of granulocyte survival and programmed cell death by cytokines and bacterial products. *Blood*, 80, 2012-2020.

- CONDEELIS, J. & POLLARD, J. W. 2006. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell*, 124, 263-266.
- CUMANO, A. & GODIN, I. 2007. Ontogeny of the hematopoietic system. *Annu Rev Immunol*, 25, 745-785.
- DALERBA, P., CHO, R. W. & CLARKE, M. F. 2007. Cancer stem cells: models and concepts. *Annu Rev Med*, 58, 267-284.
- DE VISSER, K. E., EICHTEN, A. & COUSSENS, L. M. 2006. Paradoxical roles of the immune system during cancer development. *Nat Rev Cancer*, 6, 24-37.
- DELVES, P. J. & ROITT, I. M. 2000. The immune system. First of two parts. *N Engl J Med*, 343, 37-49.
- DEMICHELI, R., RETSKY, M. W., HRUSHESKY, W. J., BAUM, M. & GUKAS, I. D. 2008. The effects of surgery on tumor growth: a century of investigations. *Ann Oncol*, 19, 1821-1828.
- DENARDO, D. G., BARRETO, J. B., ANDREU, P., VASQUEZ, L., TAWFIK, D., KOLHATKAR, N. & COUSSENS, L. M. 2009. CD4(+) T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages. *Cancer Cell*, 16, 91-102.
- DENZEL, A., MAUS, U. A., RODRIGUEZ GOMEZ, M., MOLL, C., NIEDERMEIER, M., WINTER, C., MAUS, R., HOLLINGSHEAD, S., BRILES, D. E., KUNZ-SCHUGHART, L. A., TALKE, Y. & MACK, M. 2008. Basophils enhance immunological memory responses. *Nat Immunol*, 9, 733-742.
- DIAZ-MONTERO, C. M., SALEM, M. L., NISHIMURA, M. I., GARRETT-MAYER, E., COLE, D. J. & MONTERO, A. J. 2009. Increased circulating myeloid-derived suppressor cells correlate with clinical cancer stage, metastatic tumor burden, and doxorubicin-cyclophosphamide chemotherapy. *Cancer Immunol Immunother*, 58, 49-59.
- ERIKSSON, D. & STIGBRAND, T. 2010. Radiation-induced cell death mechanisms. *Tumour Biol*, 31, 363-372.
- EUROPEAN MEDICINES AGENCY. 2016. European public assessment report (EPAR) for Avastin. Available: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/000582/WC500029260.pdf [Accessed 01/08/2016].
- EVERDS, N. 2004. Hematology of the Mouse. In: BULLOCK, P. H. J. H. G. (ed.) *The Laboratory Mouse*. London: Academic Press.
- FAVERLY, D. R., BURGERS, L., BULT, P. & HOLLAND, R. 1994. Three dimensional imaging of mammary ductal carcinoma in situ: clinical implications. *Semin Diagn Pathol*, 11, 193-198.
- FOOD AND DRUG ASSOCIATION 2011. Proposal to Withdraw Approval for the Breast Cancer Indication for Avastin (Bevacizumab). Rockville.
- FRIDLENDER, Z. G., SUN, J., KIM, S., KAPOOR, V., CHENG, G., LING, L., WORTHEN, G. S. & ALBELDA, S. M. 2009. Polarization of tumor-associated neutrophil phenotype by TGF-beta: "N1" versus "N2" TAN. *Cancer Cell*, 16, 183-194.
- FRIEDL, P. & WOLF, K. 2010. Plasticity of cell migration: a multiscale tuning model. *J Cell Biol*, 188, 11-19.
- GABRILOVICH, D. I., BRONTE, V., CHEN, S. H., COLOMBO, M. P., OCHOA, A., OSTRAND-ROSENBERG, S. & SCHREIBER, H. 2007. The terminology issue for myeloid-derived suppressor cells. *Cancer Res*, 67, 425; author reply 426.
- GRIVENNIKOV, S. I., GRETEN, F. R. & KARIN, M. 2010. Immunity, inflammation, and cancer. *Cell*, 140, 883-899.
- GROUVEN, U., BENDER, R., ZIEGLER, A. & LANGE, S. 2007. Vergleich von Messmethoden. *Deutsche Medizinische Wochenschrift* [Online], 132, e69-e73. Available: <https://www.thieme-connect.de/products/ejournals/html/10.1055/s-2007-959047> [Accessed 03/12/2014].
- GUPTA, P. B., CHAFFER, C. L. & WEINBERG, R. A. 2009. Cancer stem cells: mirage or reality? *Nat Med*, 15, 1010-1012.
- HANAHAN, D. & FOLKMAN, J. 1996. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell*, 86, 353-364.
- HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *Cell*, 100, 57-70.
- HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell*, 144, 646-674.

- HARBECK, N. & HEYWANG-KÖBRUNNER, S. 2011. Chapter 26.8 - Veränderungen und Tumoren der Mamma. In: KIECHLE, M. (ed.) *Gynäkologie und Geburtshilfe*. 2. Auflage ed. München: Elsevier.
- HASHMI, M. A., GUHA, S., SENGUPTA, P., BASU, D., BABOO, S. & NEHA 2014. Thoracic cord compression by extramedullary hematopoiesis in thalassemia. *Asian J Neurosurg*, 9, 102-104.
- HOLYSZ, H., LIPINSKA, N., PASZEL-JAWORSKA, A. & RUBIS, B. 2013. Telomerase as a useful target in cancer fighting-the breast cancer case. *Tumour Biol*, 34, 1371-1380.
- HOUSMAN, G., BYLER, S., HEERBOTH, S., LAPINSKA, K., LONGACRE, M., SNYDER, N. & SARKAR, S. 2014. Drug resistance in cancer: an overview. *Cancers (Basel)*, 6, 1769-1792.
- HUANG, B., PAN, P. Y., LI, Q., SATO, A. I., LEVY, D. E., BROMBERG, J., DIVINO, C. M. & CHEN, S. H. 2006. Gr-1+CD115+ immature myeloid suppressor cells mediate the development of tumor-induced T regulatory cells and T-cell anergy in tumor-bearing host. *Cancer Res*, 66, 1123-1131.
- HUGO, H., ACKLAND, M. L., BLICK, T., LAWRENCE, M. G., CLEMENTS, J. A., WILLIAMS, E. D. & THOMPSON, E. W. 2007. Epithelial--mesenchymal and mesenchymal--epithelial transitions in carcinoma progression. *J Cell Physiol*, 213, 374-383.
- HUSEMANN, Y., GEIGL, J. B., SCHUBERT, F., MUSIANI, P., MEYER, M., BURGHART, E., FORNI, G., EILS, R., FEHM, T., RIETHMULLER, G. & KLEIN, C. A. 2008. Systemic spread is an early step in breast cancer. *Cancer Cell*, 13, 58-68.
- HYNES, N. E. & MACDONALD, G. 2009. ErbB receptors and signaling pathways in cancer. *Curr Opin Cell Biol*, 21, 177-184.
- ILKOVITCH, D. & LOPEZ, D. M. 2009. The liver is a site for tumor-induced myeloid-derived suppressor cell accumulation and immunosuppression. *Cancer Res*, 69, 5514-5521.
- JUNTTILA, M. R. & EVAN, G. I. 2009. p53--a Jack of all trades but master of none. *Nat Rev Cancer*, 9, 821-829.
- KARNOUB, A. E., DASH, A. B., VO, A. P., SULLIVAN, A., BROOKS, M. W., BELL, G. W., RICHARDSON, A. L., POLYAK, K., TUBO, R. & WEINBERG, R. A. 2007. Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. *Nature*, 449, 557-563.
- KENIFIC, C. M., THORBURN, A. & DEBNATH, J. 2010. Autophagy and metastasis: another double-edged sword. *Curr Opin Cell Biol*, 22, 241-245.
- KENNEDY, K. M. & DEWHIRST, M. W. 2010. Tumor metabolism of lactate: the influence and therapeutic potential for MCT and CD147 regulation. *Future Oncol*, 6, 127-148.
- KIM, R., EMI, M. & TANABE, K. 2007. Cancer immunoediting from immune surveillance to immune escape. *Immunology*, 121, 1-14.
- KRAMS, M., FRAHM, S. O., KELLNER, U. & MAWRIN, C. 2013. Blut und hämatopoetisches System. *Kurzlehrbuch Pathologie*. 2. Auflage ed. Stuttgart: Thieme.
- KUMLER, I., CHRISTIANSEN, O. G. & NIELSEN, D. L. 2014. A systematic review of bevacizumab efficacy in breast cancer. *Cancer Treat Rev*, 40, 960-973.
- KUSMARTSEV, S. & GABRILOVICH, D. I. 2005. STAT1 signaling regulates tumor-associated macrophage-mediated T cell deletion. *J Immunol*, 174, 4880-4891.
- LAI, L., ALAVERDI, N., MALTAIS, L. & MORSE, H. C., 3RD 1998. Mouse cell surface antigens: nomenclature and immunophenotyping. *J Immunol*, 160, 3861-3868.
- LAMMERMANN, T. & SIXT, M. 2009. Mechanical modes of 'amoeboid' cell migration. *Curr Opin Cell Biol*, 21, 636-644.
- LEITLINIENPROGRAMM ONKOLOGIE. 2012. Interdisziplinäre S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms. Langversion. Available: http://www.awmf.org/uploads/tx_szleitlinien/032-045OL_I_S3__Brustkrebs_Mammakarzinom_Diagnostik_Therapie_Nachsorge_2012-07.pdf [Accessed 10/08/2016].
- LEMMON, M. A. & SCHLESSINGER, J. 2010. Cell signaling by receptor tyrosine kinases. *Cell*, 141, 1117-1134.
- LEVINE, B. 2007. Cell biology: autophagy and cancer. *Nature*, 446, 745-747.

- LEVINE, B. & KROEMER, G. 2008. Autophagy in the pathogenesis of disease. *Cell*, 132, 27-42.
- LI, H., HAN, Y., GUO, Q., ZHANG, M. & CAO, X. 2009. Cancer-expanded myeloid-derived suppressor cells induce anergy of NK cells through membrane-bound TGF-beta 1. *J Immunol*, 182, 240-249.
- LIU, C., YU, S., KAPPES, J., WANG, J., GRIZZLE, W. E., ZINN, K. R. & ZHANG, H. G. 2007. Expansion of spleen myeloid suppressor cells represses NK cell cytotoxicity in tumor-bearing host. *Blood*, 109, 4336-4342.
- LOBO, N. A., SHIMONO, Y., QIAN, D. & CLARKE, M. F. 2007. The biology of cancer stem cells. *Annu Rev Cell Dev Biol*, 23, 675-699.
- LOWE, S. W., CEPERO, E. & EVAN, G. 2004. Intrinsic tumour suppression. *Nature*, 432, 307-315.
- LÜLLMANN-RAUCH, R. 2006. Blut und Blutbildung. *Taschenlehrbuch Histologie*. Stuttgart: Thieme.
- LÜLLMANN-RAUCH, R. 2006. Das Immunsystem. *Taschenlehrbuch Histologie*. Stuttgart: Thieme.
- LÜLLMANN-RAUCH, R. 2006. Leber. *Taschenlehrbuch Histologie*. 2. Auflage ed. Stuttgart: Thieme.
- MANTOVANI, A., ALLAVENA, P., SICA, A. & BALKWILL, F. 2008. Cancer-related inflammation. *Nature*, 454, 436-444.
- MEDVINSKY, A., RYBTSOV, S. & TAOUDI, S. 2011. Embryonic origin of the adult hematopoietic system: advances and questions. *Development*, 138, 1017-1031.
- MERLUZZI, S., FROSSI, B., GRI, G., PARUSSO, S., TRIPODO, C. & PUCILLO, C. 2010. Mast cells enhance proliferation of B lymphocytes and drive their differentiation toward IgA-secreting plasma cells. *Blood*, 115, 2810-2817.
- MOSES, H. & BARCELLOS-HOFF, M. H. 2011. TGF- β Biology in Mammary Development and Breast Cancer. *Cold Spring Harb Perspect Biol*, 3.
- MOVAHEDI, K., GUILLIAMS, M., VAN DEN BOSSCHE, J., VAN DEN BERGH, R., GYSEMANS, C., BESCHIN, A., DE BAETSELIER, P. & VAN GINDERACHTER, J. A. 2008. Identification of discrete tumor-induced myeloid-derived suppressor cell subpopulations with distinct T cell-suppressive activity. *Blood*, 111, 4233-4244.
- MULTHOFF, G. & RADONS, J. 2012. Radiation, inflammation, and immune responses in cancer. *Front Oncol*, 2, 58.
- NAGY, J. A., CHANG, S. H., SHIH, S. C., DVORAK, A. M. & DVORAK, H. F. 2010. Heterogeneity of the tumor vasculature. *Semin Thromb Hemost*, 36, 321-331.
- NICKOLOFF, B. J., BEN-NERIAH, Y. & PIKARSKY, E. 2005. Inflammation and cancer: is the link as simple as we think? *J Invest Dermatol*, 124, x-xiv.
- NISHIKAWA, H. & SAKAGUCHI, S. 2010. Regulatory T cells in tumor immunity. *Int J Cancer*, 127, 759-767.
- ORKIN, S. H. 2000. Diversification of haematopoietic stem cells to specific lineages. *Nat Rev Genet*, 1, 57-64.
- PAGES, F., GALON, J., DIEU-NOSJEAN, M. C., TARTOUR, E., SAUTES-FRIDMAN, C. & FRIDMAN, W. H. 2010. Immune infiltration in human tumors: a prognostic factor that should not be ignored. *Oncogene*, 29, 1093-1102.
- PARKER, K. H., BEURY, D. W. & OSTRAND-ROSENBERG, S. 2015. Myeloid-Derived Suppressor Cells: Critical Cells Driving Immune Suppression in the Tumor Microenvironment. *Adv Cancer Res*, 128, 95-139.
- PARKIN, J. & COHEN, B. 2001. An overview of the immune system. *Lancet*, 357, 1777-1789.
- PAWANKAR, R., OKUDA, M., YSSEL, H., OKUMURA, K. & RA, C. 1997. Nasal mast cells in perennial allergic rhinitis exhibit increased expression of the Fc epsilonRI, CD40L, IL-4, and IL-13, and can induce IgE synthesis in B cells. *J Clin Invest*, 99, 1492-1499.
- PETER, L., MATEUS, D., CHATELAIN, P., DECLARA, D., SCHWORM, N., STANGL, S., MULTHOFF, G. & NAVAB, N. 2017. Assisting the Examination of Large Histopathological Slides with Adaptive Forests. *Medical Image Analysis*.
- PETER, L., MATEUS, D., CHATELAIN, P., SCHWORM, N., STANGL, S., MULTHOFF, G. & NAVAB, N. 2014. Leveraging Random Forests for Interactive Exploration of Large Histological Images.

Proceedings of the 17th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), September 2014.

- PHAN, L. M., YEUNG, S. C. & LEE, M. H. 2014. Cancer metabolic reprogramming: importance, main features, and potentials for precise targeted anti-cancer therapies. *Cancer Biol Med*, 11, 1-19.
- PHAROAH, P. D., DAY, N. E. & CALDAS, C. 1999. Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *Br J Cancer*, 80, 1968-1973.
- PRUD'HOMME, G. J. 2007. Pathobiology of transforming growth factor beta in cancer, fibrosis and immunologic disease, and therapeutic considerations. *Lab Invest*, 87, 1077-1091.
- PUNTURIERI, A., SZABO, E., CROXTON, T. L., SHAPIRO, S. D. & DUBINETT, S. M. 2009. Lung cancer and chronic obstructive pulmonary disease: needs and opportunities for integrated research. *J Natl Cancer Inst*, 101, 554-559.
- RAZA, A., FRANKLIN, M. J. & DUDEK, A. Z. 2010. Pericytes and vessel maturation during tumor angiogenesis and metastasis. *Am J Hematol*, 85, 593-598.
- RIBECHINI, E., LEENEN, P. J. & LUTZ, M. B. 2009. Gr-1 antibody induces STAT signaling, macrophage marker expression and abrogation of myeloid-derived suppressor cell activity in BM cells. *Eur J Immunol*, 39, 3538-3551.
- ROBERT-KOCH-INSTITUT. 2010. Verbreitung von Krebserkrankungen in Deutschland. Entwicklung der Prävalenzen zwischen 1990 und 2010. *Beiträge zur Gesundheitsberichterstattung des Bundes*. [Online], 77-82. Available: http://edoc.rki.de/documents/rki_fv/re2vZ2t28lr8Y/PDF/23GSS31yB0GKUuU.pdf [Accessed 15/08/2016].
- ROBERT-KOCH-INSTITUT. 2015. Krebs in Deutschland 2011/2012. Available: http://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/kid_2015/krebs_in_deutschland_2015.pdf;jsessionid=C4E6B4E764C4B15236ABE7C317E93698.2_cid290?__blob=publicationFile [Accessed 15/08/2015].
- ROBERT-KOCH-INSTITUT & STATISTISCHES BUNDESAMT. 2005. Brustkrebs. *Gesundheitsberichterstattung des Bundes* [Online], 25, 1-37. Available: http://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsT/brustkrebs.html;jsessionid=4263D31B23EAFD21DBC3F9D0CA5BB712.2_cid363?n=2370692 [Accessed 15/08/2016].
- RODRIGUEZ GOMEZ, M., TALKE, Y., GOEBEL, N., HERMANN, F., REICH, B. & MACK, M. 2010. Basophils support the survival of plasma cells in mice. *J Immunol*, 185, 7180-7185.
- RUOSLAHTI, E. 2002. Specialization of tumour vasculature. *Nat Rev Cancer*, 2, 83-90.
- SAUER, R. 2010. Biologische Grundlagen der Strahlentherapie von Tumoren. In: SAUER, R. (ed.) *Strahlentherapie und Onkologie*. München: Elsevier.
- SAUER, R. 2010. Mammakarzinom. In: SAUER, R. (ed.) *Strahlentherapie und Onkologie*. München: Elsevier.
- SAUER, R. 2010. Tumorpathologie. In: SAUER, R. (ed.) *Strahlentherapie und Onkologie*. München: Thieme.
- SHAWKER, T. H., HILL, M., HILL, S. & GARRA, B. 1987. Ultrasound appearance of extramedullary hematopoiesis. *Journal of Ultrasound in Medicine*, 6, 283-290.
- SHERR, C. J. & MCCORMICK, F. 2002. The RB and p53 pathways in cancer. *Cancer Cell*, 2, 103-112.
- SHIZURU, J. A., NEGRIN, R. S. & WEISSMAN, I. L. 2005. Hematopoietic stem and progenitor cells: clinical and preclinical regeneration of the hematolymphoid system. *Annu Rev Med*, 56, 509-538.
- SICA, A., ALLAVENA, P. & MANTOVANI, A. 2008. Cancer related inflammation: the macrophage connection. *Cancer Lett*, 267, 204-215.
- SINGH, A. & SETTLEMAN, J. 2010. EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene*, 29, 4741-4751.
- SINHA, P., CLEMENTS, V. K., BUNT, S. K., ALBELDA, S. M. & OSTRAND-ROSENBERG, S. 2007. Cross-talk between myeloid-derived suppressor cells and macrophages subverts tumor immunity toward a type 2 response. *J Immunol*, 179, 977-983.

- SLAMON, D. J., GODOLPHIN, W., JONES, L. A., HOLT, J. A., WONG, S. G., KEITH, D. E., LEVIN, W. J., STUART, S. G., UDOVE, J., ULLRICH, A. & ET AL. 1989. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science*, 244, 707-712.
- SOUCEK, L., LAWLOR, E. R., SOTO, D., SHCHORS, K., SWIGART, L. B. & EVAN, G. I. 2007. Mast cells are required for angiogenesis and macroscopic expansion of Myc-induced pancreatic islet tumors. *Nat Med*, 13, 1211-1218.
- SPARMANN, A. & BAR-SAGI, D. 2004. Ras-induced interleukin-8 expression plays a critical role in tumor growth and angiogenesis. *Cancer Cell*, 6, 447-458.
- STANGL, S., GEHRMANN, M., RIEGGER, J., KUHS, K., RIEDERER, I., SIEVERT, W., HUBE, K., MOCIKAT, R., DRESSEL, R., KREMMER, E., POCKLEY, A. G., FRIEDRICH, L., VIGH, L., SKERRA, A. & MULTHOFF, G. 2011. Targeting membrane heat-shock protein 70 (Hsp70) on tumors by cmHsp70.1 antibody. *Proc Natl Acad Sci U S A*, 108, 733-738.
- STRASSER, A., HARRIS, A. W., HUANG, D. C., KRAMMER, P. H. & CORY, S. 1995. Bcl-2 and Fas/APO-1 regulate distinct pathways to lymphocyte apoptosis. *Embo j*, 14, 6136-6147.
- STRASSER, A., O'CONNOR, L. & DIXIT, V. M. 2000. Apoptosis signaling. *Annu Rev Biochem*, 69, 217-245.
- TALMADGE, J. E. 2007. Pathways mediating the expansion and immunosuppressive activity of myeloid-derived suppressor cells and their relevance to cancer therapy. *Clin Cancer Res*, 13, 5243-5248.
- TALMADGE, J. E. & FIDLER, I. J. 2010. AACR centennial series: the biology of cancer metastasis: historical perspective. *Cancer Res*, 70, 5649-5669.
- TAO, K., FANG, M., ALROY, J. & SAHAGIAN, G. G. 2008. Imagable 4T1 model for the study of late stage breast cancer. *BMC Cancer*, 8, 228.
- VAJDIC, C. M. & VAN LEEUWEN, M. T. 2009. Cancer incidence and risk factors after solid organ transplantation. *Int J Cancer*, 125, 1747-1754.
- VAN ROY, F. & BERTX, G. 2008. The cell-cell adhesion molecule E-cadherin. *Cell Mol Life Sci*, 65, 3756-3788.
- VANDER HEIDEN, M. G., CANTLEY, L. C. & THOMPSON, C. B. 2009. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science*, 324, 1029-1033.
- VOSKOBOINIK, I., WHISSTOCK, J. C. & TRAPANI, J. A. 2015. Perforin and granzymes: function, dysfunction and human pathology. *Nat Rev Immunol*, 15, 388-400.
- WARBURG, O. 1930. The Metabolism of Tumours: Investigations from the Kaiser Wilhelm Institute for Biology.
- YADAV, B. S., CHANANA, P. & JHAMB, S. 2015. Biomarkers in triple negative breast cancer: A review. *World J Clin Oncol*, 6, 252-263.
- YANG, L., DEBUSK, L. M., FUKUDA, K., FINGLETON, B., GREEN-JARVIS, B., SHYR, Y., MATRISIAN, L. M., CARBONE, D. P. & LIN, P. C. 2004. Expansion of myeloid immune suppressor Gr⁺CD11b⁺ cells in tumor-bearing host directly promotes tumor angiogenesis. *Cancer Cell*, 6, 409-421.
- YOUN, J. I., NAGARAJ, S., COLLAZO, M. & GABRILOVICH, D. I. 2008. Subsets of myeloid-derived suppressor cells in tumor-bearing mice. *J Immunol*, 181, 5791-5802.
- ZONG, W. X. & THOMPSON, C. B. 2006. Necrotic death as a cell fate. *Genes Dev*, 20, 1-15.

Danksagung

Zunächst möchte ich mich bei dem gesamten Team der Abteilung für Experimentelle Radioonkologie und Strahlenbiologie des Klinikums rechts der Isar bedanken, das mir diese Arbeit ermöglicht hat.

Mein besonderer Dank gilt meiner Doktormutter Frau Professor Doktor Gabriele Multhoff, die mir über die Jahre hinweg mit ihrem umfangreichen Fachwissen mit Rat und Tat zur Seite stand. Auf konstruktive Kritik und ein offenes Ohr zu jeder Zeit konnte ich mich stets verlassen.

Weiterhin möchte ich mich besonders bei meinem Betreuer Doktor Stefan Stangl bedanken, der mir mit seiner fachlichen und praktischen Kompetenz stets eine große Hilfe war. Er hat sowohl den experimentellen als auch den theoretischen Teil dieser Arbeit angeleitet und möglich gemacht.

Ein Dank geht auch an Doktor Katja Steiger, die mir in der histologischen Beurteilung der Leberschnitte beratend zur Seite stand sowie Doktor Annette Feuchtinger, die mir half Definiens für meine Scans bestmöglich zu nutzen.

An dieser Stelle möchte ich mich auch bei Loïc Peter für die gute wissenschaftliche Zusammenarbeit bedanken. Ich schätze seine Kompetenz in Fragen der Informatik sowie der Analyse großer Bilddateien sehr und bedanke mich für seine stets hilfreichen Ratschläge.

Danke auch an Doktor Bernhard Haller für die kompetente und fallbezogene statistische Beratung.

Natürlich danke ich auch meinen Eltern Birgit und Ernst Schworm sowie meinen Geschwistern Lena Müller-Wirthmann und Felix Schworm, die mich über Jahre hinweg unterstützt und das Gelingen dieser Arbeit nie bezweifelt haben.

Desweiteren möchte ich Sinja Heger, Lena Sperl und Anna Schwarz dafür danken, dass sie nie müde wurden, mit mir fachliche sowie organisatorische Fragen zu klären.

Ich danke Emanuel Scholz, der mir in jeder Phase dieser Arbeit eine große Stütze war und ohne dessen Kompetenz ich an vielen Stellen schlicht gescheitert wäre.

Anhang

1. Peter, L., Mateus, D., Chatelain, P., Declara, D., Schworm, N., Stangl, S., Multhoff, G., Navab, N. - **Assisting the Examination of Large Histopathological Slides with Adaptive Forests** - *Medical Image Analysis, 2017*
2. Peter, L., Mateus, D., Chatelain, P., Schworm, N., Stangl, S., Multhoff, G., Navab, N. - **Leveraging Random Forests for Interactive Exploration of Large Histological Images** - *Proceedings of the 17th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2014*

Assisting the Examination of Large Histopathological Slides with Adaptive Forests

Loïc Peter^a, Diana Mateus^{a,b}, Pierre Chatelain^{a,c}, Denis Declara^a, Noemi Schworm^d, Stefan Stangl^d,
Gabriele Multhoff^{d,e}, Nassir Navab^{a,f}

^aComputer Aided Medical Procedures, Technische Universität München, Germany

^bInstitute of Computational Biology, Helmholtz Zentrum München, Germany

^cUniversité de Rennes 1, IRISA, France

^dDepartment of Radiation Oncology, Technische Universität München, Germany

^eInstitute of Innovative Radiotherapy (iRT), Department of Radiation Sciences, Helmholtz Zentrum München, Germany

^fComputer Aided Medical Procedures, Johns Hopkins University, USA

Abstract

The examination of biopsy samples plays a central role in the diagnosis and staging of numerous diseases, including most cancer types. However, because of the large size of the acquired images, the localization and quantification of diseased portions of a tissue is usually time-consuming, as pathologists must scroll through the whole slide to look for objects of interest which are often only scarcely distributed. In this work, we introduce an approach to facilitate the visual inspection of large digital histopathological slides. Our method builds on a random forest classifier trained to segment the structures sought by the pathologist. However, moving beyond the pixelwise segmentation task, our main contribution is an interactive exploration framework including: (i) a region scoring function which is used to rank and sequentially display regions of interest to the user, and (ii) a relevance feedback capability which leverages human annotations collected on each suggested region. Thereby, an online domain adaptation of the learned pixelwise segmentation model is performed, so that the region scores adapt on-the-fly to possible discrepancies between the original training data and the slide at hand. Three real-time update strategies are compared, including a novel approach based on online gradient descent which supports faster user interaction than an accurate delineation of objects. Our method is evaluated on the task of extramedullary hematopoiesis quantification within mouse liver slides. We assess quantitatively the retrieval abilities of our approach and the benefit of the interactive adaptation scheme. Moreover, we demonstrate the possibility of extrapolating, after a partial exploration of the slide, the surface covered by hematopoietic cells within the whole tissue.

Keywords: Random Forests, Histopathology, Online Learning, Active Learning, Domain Adaptation

1. Introduction

Histopathology is a crucial tool in modern clinical practice. It consists in the microscopic observation of biological tissues surgically extracted from a patient, in order to collect information regarding the presence or extent of a particular disease in the sample. In particular, it is part of the standard experimental protocol for the definitive diagnosis, grading and staging of most cancer types and plays an essential role in the design of appropriate patient-specific treatments. Histopathological examinations usually aim at searching for a certain kind of anatomical structure, like biomarkers, cancer cells or necrotic areas, whose presence or proportion within the tissue has to be quantitatively estimated. Although this procedure is traditionally conducted under a standard optical microscope, digital acquisitions of entire slices can be performed at comparable resolutions and are increasingly used by pathologists in their clinical workflow as well as for educational and research purposes (Farahani et al., 2015). Moving from optical to digital examinations has been shown to maintain similar diagnosis performances (Jukić et al., 2011; Bauer et al., 2013) and offers numerous additional advantages such as the applicability of image analysis algorithms, easier recordings and safer storage of patient data, and the pos-

sibility of displaying the scanned tissue to several examiners simultaneously (Al-Janabi et al., 2012). However, because of their high resolution, the size of digitally acquired images is very large and commonly reaches the order of a billion of pixels (Cooper et al., 2012). This increases the time required for manual quantification procedures: beyond the tediousness of annotating objects in images, a pathologist also spends a lot of time navigating through the large slide looking for evidence of the disease of interest. Moreover, the objects to localize may only be scarcely distributed, for instance at early stages of diseases or after a treatment has been applied. In such situations, the exploration phase even becomes the bottleneck of the process, since most of the time of the pathologist is spent scrolling through uninformative areas (Fig. 1).

Some characteristics of the field of histopathology bring specific challenges for an automated analysis of the acquired images. First, the accurate identification of diseased areas based on their visual appearance can be a very difficult task requiring a lot of expertise. Pathologists are typically trained several years before reaching satisfactory diagnosis abilities (Jaarsma et al., 2014), and the variability between experts remains nevertheless significant for several applications (Meyer et al., 2005; Gonul et al., 2006; Gilles et al., 2007; Eefting et al., 2009). An-

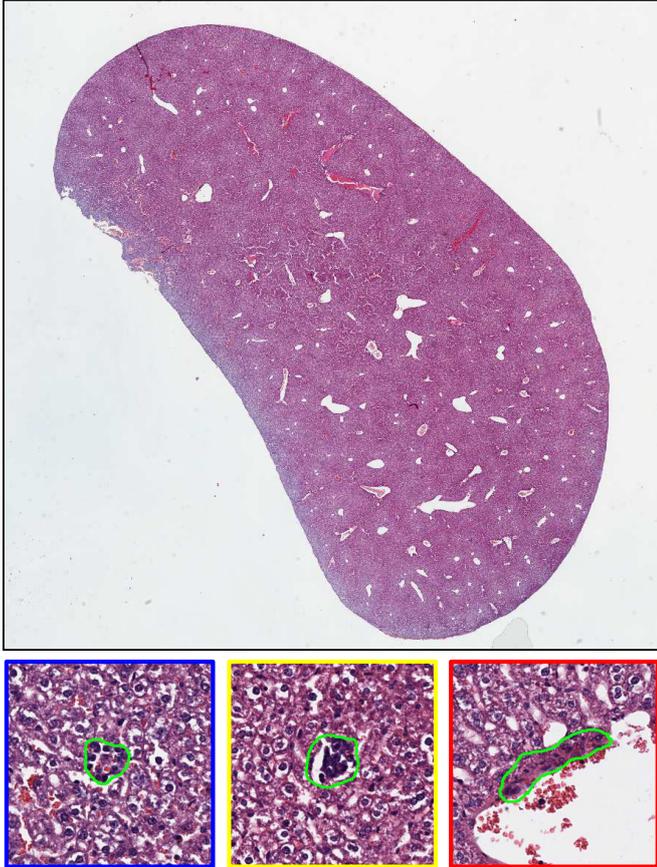


Figure 1: **Typical histopathological slide.** Three regions (blue, yellow and red squares) containing a cluster of hematopoietic cells are highlighted. Our method aims at retrieving such regions of interest within a large slide which mostly contains irrelevant background areas. Note the absence of large-scale context to guide the visual search, which would require an exhaustive screening of the slide in the case of a manual examination.

other common challenge in histopathological image analysis is the visual variability between two acquisitions. In particular, the consistency of the staining procedure is difficult to control experimentally leading to variations in terms of dye concentration (Fig. 2). Therefore, an algorithm that has been trained or designed on labeled data may not generalize well to newly acquired samples. To mitigate this source of inaccuracies, color normalization can be performed as a preprocessing step and remains an active field of research (Rabinovich et al., 2003; Macenko et al., 2009; Khan et al., 2014; Onder et al., 2014; Bautista and Yagi, 2015; Vahadane et al., 2015), together with generic techniques for online domain adaptation (Sec. 3.3). Finally, a tissue extracted surgically and observed under a microscope is less structured than other kinds of medical data such as body scans, while being of a much larger size. Objects of interest are expected to appear anywhere within the tissue, so that location or connectivity priors are rarely available.

In this work, we introduce an interactive method to assist a pathologist in exploring and quantifying large histological slides (Fig. 3). Our approach builds on a *pixelwise segmentation model* provided by a pre-trained random forest classifier

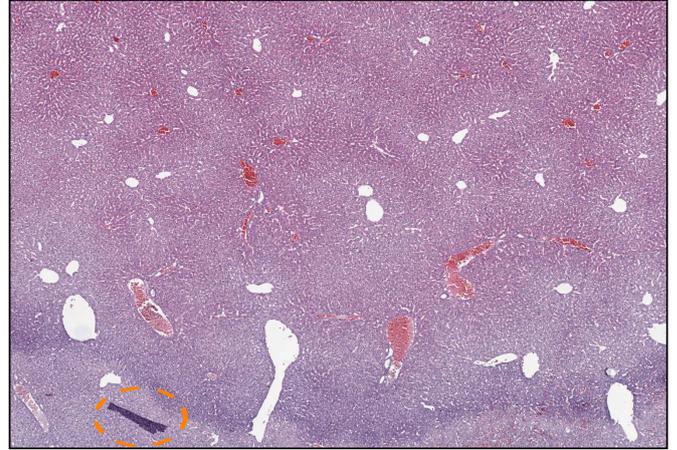


Figure 2: **Examples of visual artifacts.** A portion of a slide is displayed here. The staining itself is inhomogeneous and presents a vertical shading. Moreover, a dark artifact is present (circled in orange). Such a visual variability between and within slides complicates the application of supervised learning techniques and prompts an adaptation at prediction time.

(Sec. 4.1), and uses its output to perform an *interactive slide exploration* by suggesting, in a sequential manner, a series of candidate regions of interest (Sec. 4.2). This interactive navigation framework includes a component which allows the pathologist to provide, after each suggestion, some feedback about the actual relevance of the proposed region. From these user inputs, the underlying forest-based model is modified on-the-fly via a real-time online adaptation framework. This enables a progressive adjustment to the characteristics of the data at hand and compensates for possible mismatches with the original training set without specific assumptions about their nature, in contrast to the aforementioned explicit stain normalizations. Finally, we also demonstrate how a *whole-slide quantification* can be inferred after a partial exploration of the slide (Sec 4.3). The experimental evaluation of our approach was conducted in the context of extramedullary hematopoiesis quantification within mouse liver slides. The results demonstrate the ability of our method to quickly retrieve regions of interest and confirm the benefit of the interactive online adaptation scheme. The whole-slide quantification capabilities of our approach are also evaluated depending on the duration of the exploration stage. Finally, we demonstrate how one of our update strategies can be used with one-click inputs for faster interaction without decreasing its performance (Sec. 6).

2. Contributions

We propose an interactive framework using a forest-based pixelwise segmentation to explore large digital slides according to a predefined quantification task. In our application case, this clinical objective is the assessment of the surface covered by hematopoietic cells within mouse liver slides. Two main methodological contributions are introduced:

- The design of a region scoring function to convert pixelwise predictions into a score for each region of the slide.

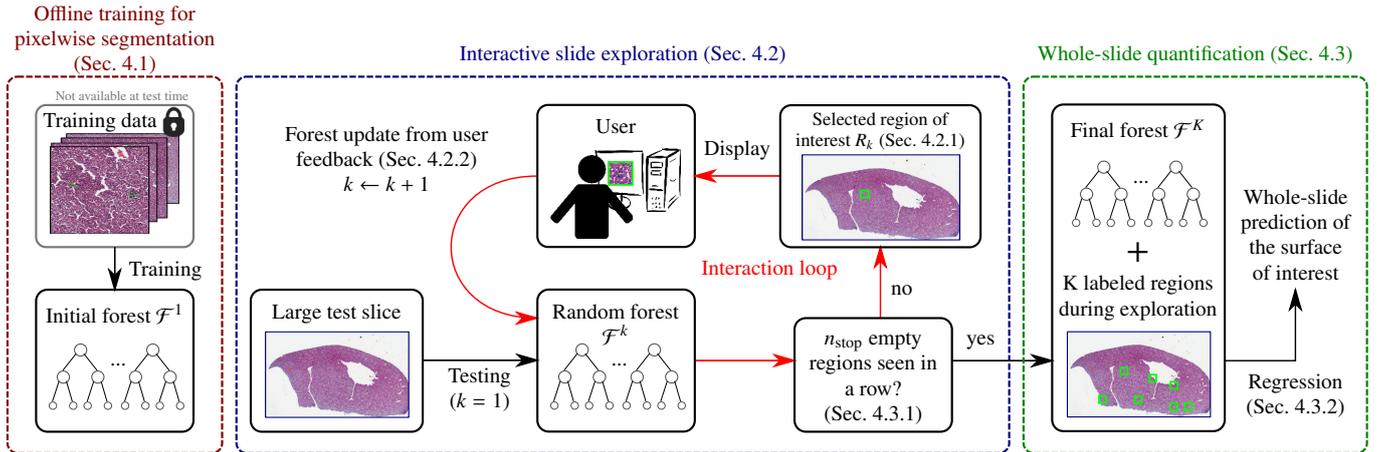


Figure 3: **Overview of our scenario.** Initially, a classification forest \mathcal{F}^1 is trained offline at pixel level to segment the structures of interest (in our case, clusters of hematopoietic cells). Given a large new slide to analyze, our method sequentially displays to the pathologist regions that are likely to contain these objects, thereby alleviating a tedious manual navigation through the slide. After the suggestion of a region R_k ($k \geq 1$), chosen as to maximize a forest-based scoring function $\phi(\cdot | \mathcal{F}^k)$, the user provides a relevance feedback about its actual content. From this input, the current forest \mathcal{F}^k is updated in real time leading to a new forest \mathcal{F}^{k+1} . By doing so, the visual specificities of the test slide are progressively incorporated into the decision model so that upcoming region suggestions can be reconsidered. The exploration is stopped after seeing a certain number n_{stop} of negative regions suggestions in a row. The updated forest and the user labels collected during the exploration can then be combined in a regression framework to predict the total surface covered by hematopoietic cells in the tissue, including the areas not observed by the pathologist.

- An online domain adaptation scheme based on interactions with the user. Three real-time strategies are compared, including a novel approach based on online gradient descent which is compatible with lighter kinds of annotations.

By means of our experimental validation, we were able:

- To demonstrate the exploration abilities of our method and the benefit of the online adaptation.
- To show how, as a by-product of an only partial exploration of the slide, a whole-slide estimate of the surface covered by hematopoietic cells can be predicted.
- To study experimentally the use of discretized user inputs during adaptation, and demonstrate how one-click inputs can be effectively used instead of accurate annotations.

The present manuscript builds on an earlier version presented at a conference (Peter et al., 2014) and includes the following extensions. Our original forest update scheme based on online gradient descent has been improved by incorporating theoretical results from the online learning literature. In addition, two alternative forest update methods are considered and studied in our experiments. The size of the dataset has been doubled and now consists of 70 fully labeled high-resolution images extracted from 16 different mouse slides. The cross-validation used in our experiments now includes the optimization of the update-related hyperparameter λ on a validation set. We also demonstrate how a whole-slide quantification can be performed based on an only partial exploration of the slide. Finally, we extended the validation regarding the discretization of user inputs. The impact of the amount of discretization is studied experimentally and we introduce and evaluate an alternative quantization procedure showing a more robust behavior.

3. Related Work

3.1. Learning-Based Image Segmentation in Histopathology

A large number of methods have been introduced for the automated analysis of histological slides and are progressively put into practice, as demonstrated by the development of general-purpose toolboxes such as Ilastik (Sommer et al., 2011), CellProfiler (Carpenter et al., 2006) and CellCognition (Held et al., 2010). We refer to surveys (Gurcan et al., 2009; Veta et al., 2014) for a broader overview of the field and focus here more specifically on the case of segmentation of histological images. In this context, learning-based techniques have been successfully applied for different tasks, including cell segmentation within follicular lymphoma images (Kong et al., 2011) and within lung and brain tumor samples (Su et al., 2015), segmentation of cancer tissue within colon images (Xu et al., 2014), and whole-slide segmentation of necrotic areas (Homeyer et al., 2013). The prediction of the Gleason grading, which is one of the most important quantitative measures for prostate cancer staging, has also gathered a particular interest in the field. Several statistical learning methods such as support vector machines (Nguyen et al., 2014), AdaBoost (Gorelick et al., 2013) and randomized forests (Khurd et al., 2010) were used towards an automatic prediction of this score. To overcome the difficulty of efficiently processing large whole-slide images, multi-resolution approaches were designed to find and segment regions of interest in a hierarchical manner (Sertel et al., 2009; Roullier et al., 2011; Huang et al., 2011; Doyle et al., 2012). These approaches simulate the behavior of a pathologist, starting from the lowest resolution and progressively refining the analysis towards presumably interesting areas. In this work, our segmentation model is a pixelwise forest classifier trained

with Haar-like features (Sec. 4.1), whose efficiency allows us to operate directly at the highest resolution instead.

3.2. Assisted Navigation within Large Digital Slides

While the aforementioned methods focus on the segmentation task, a few other approaches aim at identifying regions of interest within histological data. A method for classifying regions as relevant or irrelevant using support vector machines was introduced (Bahlmann et al., 2012), and extended to a scenario where the ground truth is generated by analyzing the actual behavior of a pathologist with viewport tracking data (Mercan et al., 2014). These classification techniques are closer to our goal but differ methodologically in two aspects. First, the methods above model the region retrieval task as a classification problem, ignoring the differences between positive regions. In particular, one may desire to display in priority regions containing larger structures of interest. Building our region scoring scheme on an underlying segmentation model naturally provides such a ranking of regions and gives the opportunity to extrapolate the quantification estimate to unobserved areas. Secondly, our method is more flexible, as it includes the ability to update the region selection rule from user annotations collected after each suggestion, in an online domain adaptation fashion.

3.3. Online Domain Adaptation

Experimental constraints during the preparation of a tissue may induce inconsistencies in terms of visual aspect between acquisitions. In particular, a newly acquired sample may differ from the data used to train the initial classifier. The problem of domain adaptation consists in the correction of such a shift between the distributions of the training and testing data. Most domain adaptation strategies retrain a new classifier once samples from the target domain have been observed. In our case, new samples are collected every time a suggested region is labeled by the user, after which the current classifier is accordingly adapted. To keep this interaction loop tractable in practice, the updates must be performed in real time, which excludes a retraining of the classifier between two suggestions. Because of this constraint, our scenario is more precisely an *online domain adaptation* task. This relatively recent paradigm has been addressed in a few works only. A generic unsupervised method based on Gaussian process regression was introduced to adapt the decision boundary of any black-box classifier to a target image (Jain and Learned-Miller, 2011). Other approaches combined a classifier trained on the source data with an online classifier continuously updated from the target data (Zhao and Hoi, 2010; Tommasi et al., 2012). Originally designed with kernel-based classifiers, transferring this technique to forest models poses some challenges. First, since it treats the initial classifier as a whole, applying this procedure to a forest would result in entire trees being discarded if they do not suit the testing distribution anymore. Thereby, one would ignore the fact that some areas of the feature space, i.e. some tree leaves, may remain valid. Moreover, it requires a strategy to build decision trees online which is not straightforward. One of the most popular strategies for online forest training

is to grow trees progressively, starting from a root node, by turning a leaf node into an internal node as soon as a split of sufficient quality can be found, both in terms of information gain and statistical representativity (Saffari et al., 2009). Moreover, this approach includes, in the context of tracking, the idea of online domain adaptation by discarding trees when they no longer match the distribution of the arriving samples. Another recent work models trees as samples from Mondrian distributions (Lakshminarayanan et al., 2014) and improves over existing online approaches, yet at the cost of losing the compatibility of forests with high-dimensional feature spaces. In our case, the large size of histological data and the fact that online forest updates take place between two human interactions impose strict computational constraints to ensure the practical applicability of our method. In particular, modifications of the structure of the trees, such as in the two aforementioned approaches, are compromised. By acting on the leaf probabilities only, we achieve real-time updates between two interactions (Sec. 4.2.2).

3.4. Active Learning

Finally, querying user labels in order to improve a classifier can be seen as a form of active learning (Settles, 2010), which inspired a few approaches in the context of histopathology (Homeyer et al., 2011). In general, active learning algorithms query the label of the most uncertain samples given the knowledge of the current classifier, in order to minimize the labeling effort from the user. The spirit of our approach is different: from a clinical perspective, a pathologist is only interested in seeing positive examples in a short amount of time. This asymmetry between positive and negative observations leads us to focus on finding and displaying positive regions as quickly as possible, so that they can be visually inspected and validated by the user. The annotations obtained during the process are used to assess the accuracy of the initial model and correct it if necessary. Moreover, by doing so, any erroneous region suggestion naturally provides a challenging negative example to include in the online adaptation process.

4. Methods

This section exposes our methodology. Our slide examination method is based on an initial forest-based model whose goal is to segment the objects of interest within the tissue. This initial forest, denoted \mathcal{F}^1 , is trained offline on some labeled examples and encodes the available prior knowledge before observing the test data. After training, the original training data are no longer considered available. This assumption is driven by practical aspects: while sharing and transferring a classifier from a machine to another is straightforward, this is usually less feasible with patient data which are of larger size and subject to ethical considerations. The training mechanism generating \mathcal{F}^1 from labeled images is conducted in a standard way (Sec. 4.1).

In addition to its use as a navigation tool (Sec. 4.2), we also discuss whole-slide quantification abilities for our method. Indeed, after the exploration phase has been completed, it can be of interest to estimate the total amount of cells in the slide,

including in the areas that have not been seen by the pathologist. After a discussion on a relevant definition of a stopping criterion for the exploration stage in Sec. 4.3.1, we expose our regression-based strategy for this task in Sec. 4.3.2.

4.1. Forest-Based Segmentation Model

Our slide analysis method builds on a segmentation model assessing the surface covered by a certain type of anatomical structure within the acquired tissue. In this work, we follow a pixelwise classification approach by modeling the segmentation process as a series of independent decisions for each pixel \mathbf{x} in an image. Each decision is conducted by a random forest classifier \mathcal{F} which outputs a conditional probability $\mathbb{P}(\mathbf{x} \in \mathcal{P}|\mathcal{F})$ that the true label $y(\mathbf{x})$ is 1 given the classifier \mathcal{F} , with $\mathcal{P} = \{\mathbf{x} \in \mathcal{X} | y(\mathbf{x}) = 1\}$ denoting the set of positive samples. \mathcal{X} denotes the theoretical set of observable samples.

The visual aspect around each pixel has to be quantitatively modeled by a set of features on which the node splitting functions are based. In this work, we use Haar-like features which describe each pixel by its visual content at offset locations. The precomputation of an integral image for each color channel allows a fast access to any of these feature values, so that this large set of descriptors can be efficiently handled at a low memory footprint. These generic and computationally efficient features were originally used in combination with a boosting classifier, first in the context of face detection (Viola and Jones, 2004) and later extended to object recognition and segmentation (Shotton et al., 2006). Their use within the random forest framework (Criminisi et al., 2009) has been successfully applied to a great variety of tasks and imaging modalities (Pauly et al. (2011); Montillo et al. (2011); Chatelain et al. (2013); Gauriau et al. (2014); Ebner et al. (2014); Kontschieder et al. (2014); Zikic et al. (2014)).

The forest training was conducted using axis-aligned splitting functions and the Gini index (Breiman et al., 1984) as node purity measure. More precisely, n_{tries} candidate Haar-like features are randomly drawn at each node and $n_{\text{thresholds}}$ thresholds are tried for each of them. The candidate features are drawn in a fine-to-coarse fashion instead of the standard uniform sampling over a patch, so that an appropriate visual scale can be automatically inferred at each node (Peter et al., 2015). Among all tried splits, the best one is retained and the procedure recursively repeated until purity, or until none of the n_{tries} candidate splits send at least $n_{\text{samples/leaf}}$ training samples to the two child nodes. As soon as one of these stopping criteria is satisfied, a leaf L is created. A probability π_L is computed from the class histogram of training samples reaching L and stored at this leaf.

4.2. Interactive Slide Exploration

Considering a new test slide, we partition it into a predefined set \mathcal{R} of non-overlapping regions of fixed size $\delta \times \delta$. The first step of our algorithm consists in retrieving the region $R_1 \in \mathcal{R}$ of highest interest to the user. Since \mathcal{F}^1 provides a pixelwise estimate, this choice of region is made according to a region scoring function $\phi(R|\mathcal{F}^1)$, which predicts the expected interest of a region R given the knowledge carried by the forest model \mathcal{F}^1 . The first region displayed to the pathologist is

$R_1 = \operatorname{argmax}_{R \in \mathcal{R}} \phi(R|\mathcal{F}^1)$. Once R_1 has been shown, the pathologist reports the actual relevance of its content. To do so, two possibilities of user labelings are considered in this work: either a full delineation of the object of interest in R_1 , which is accurate but time-consuming, or a one-click input obtained by discretization, which is faster to provide but more ambiguous (Sec. 6). Using the input of the pathologist on R_1 , the forest \mathcal{F}^1 is accordingly modified, leading to a new forest \mathcal{F}^2 . This procedure is repeated several times by showing, at each iteration, the region $R_k = \operatorname{argmax}_{R \in \mathcal{R} \setminus \{R_1, \dots, R_{k-1}\}} \phi(R|\mathcal{F}^k)$. In Sec. 4.2.1, we describe in more details our choice of scoring function ϕ . Section 4.2.2 is dedicated to the techniques for online domain adaptation, where three real-time alternatives are described including a novel approach based on online gradient descent.

4.2.1. Region Scoring Function

After training on a set of labeled images, we obtain a random forest classifier \mathcal{F}^1 which outputs, for every pixel \mathbf{x} in an image, the probabilistic estimate $\mathbb{P}(\mathbf{x} \in \mathcal{P}|\mathcal{F}^1) \in [0, 1]$ that \mathbf{x} is a positive instance, i.e. belongs to one of the sought structures. Since the goal of our approach is to display regions of interest to a pathologist, we use these pixelwise forest predictions to build a region scoring function ϕ . We propose to define the score of a region R as

$$\phi(R|\mathcal{F}) = \sum_{\mathbf{x} \in R} \mathbb{P}(\mathbf{x} \in \mathcal{P}|\mathcal{F}) \quad (1)$$

given a pixelwise classification forest \mathcal{F} . This scoring function can be interpreted as the mathematical expectation of a random variable counting the number of positive pixels in the region R . In particular, regions containing larger objects obtain a higher score. Since the quantification task consists in the estimation of the total surface covered by the structures of interest within the slide, this amounts to showing first regions which have a greater contribution to this quantity. Due to its simplicity, our scoring function possesses important properties in the context of forest updates (computational efficiency and convexity) which will be detailed in Sec. 4.2.2. These advantages result from the fact that $\phi(R|\mathcal{F})$ can be rewritten as a scalar product between the vector of leaf models of the forest \mathcal{F} and a sparse vector characterizing the region R . The derivation of this equivalent formulation is exposed in the next paragraph. Finally, although the leaf models are probabilistic estimates of a classification task in this work, a similar scoring function could be equivalently used for a regression output. For instance, if the forest predicts a density of objects (Fiaschi et al., 2012), this scoring function would count a number of objects in the region instead of their surface.

Expressing $\phi(R|\mathcal{F})$ as a scalar product. Let us first introduce some notations. The set of leaf nodes belonging to the t^{th} tree is denoted \mathcal{L}_t . $\mathcal{L} = \cup_{1 \leq t \leq n_{\text{trees}}} \mathcal{L}_t$ is the set of all leaf nodes contained in the forest \mathcal{F} , and we denote $\text{tree}(L) \in \{1, \dots, n_{\text{trees}}\}$ the index of the tree to which a leaf $L \in \mathcal{L}$ belongs. We arbitrarily order the finite set \mathcal{L} and consider the leaf probabilities jointly as a (finite-dimensional) vector $\boldsymbol{\pi} = (\pi_L)_{L \in \mathcal{L}}$. We denote $\boldsymbol{\Sigma} = (\sigma_t)_{1 \leq t \leq n_{\text{trees}}}$ the list of routing functions $\sigma_t : \mathcal{X} \rightarrow \mathcal{L}_t$, which assign to each sample $\mathbf{x} \in \mathcal{X}$ the leaf $\sigma_t(\mathbf{x})$ that it reaches when passed through the t^{th} tree (Fig. 4). Intuitively, $\boldsymbol{\Sigma}$ encodes

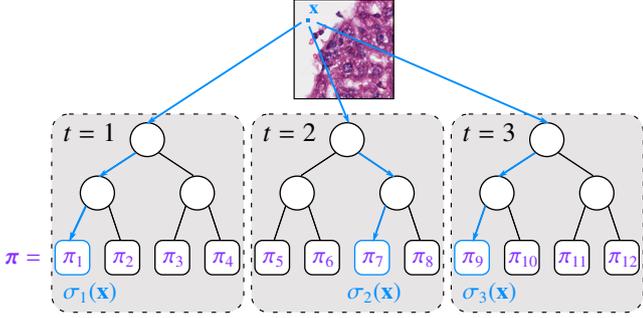


Figure 4: **Notations.** This figure illustrates our notations in the simplified case of a small forest (3 trees of depth 2). The leaves are arbitrarily ordered and the leaf models are jointly considered as a vector π . For each pixel \mathbf{x} , $\sigma_t(\mathbf{x})$ denotes the leaf reached by \mathbf{x} when passed through the t^{th} tree. The list of functions $\Sigma = (\sigma_t)_{1 \leq t \leq n_{\text{trees}}}$ encompasses information about the arrangement of the trees and the node splitting functions.

the structure of the forest determined by the arrangement of the nodes and the splitting functions, i.e. the way the forest partitions the space of observations \mathcal{X} , while the vector π defines the label predictions stored in the terminal nodes. Σ and π fully determine the forest decision rule, defined as

$$\mathbb{P}(\mathbf{x} \in \mathcal{P}|\mathcal{F}) = \frac{1}{n_{\text{trees}}} \sum_{t=1}^{n_{\text{trees}}} \pi_{\sigma_t(\mathbf{x})}. \quad (2)$$

By incorporating Eq. 2 into the definition of the scoring function (Eq. 1) and rearranging the sum signs (see Appendix A for details), we obtain the identity

$$\phi(R|\mathcal{F}) = \phi(R|\Sigma, \pi) = \langle \rho(R|\Sigma), \pi \rangle, \quad (3)$$

where $\rho(R|\Sigma) = (\rho_L(R|\Sigma))_{L \in \mathcal{L}}$ is a vector of dimension n_{leaves} characterizing the region R and defined as

$$\rho_L(R|\Sigma) = \frac{1}{n_{\text{trees}}} \frac{\#\{\mathbf{x} \in R \mid \sigma_{\text{tree}(L)}(\mathbf{x}) = L\}}{\text{number of pixels in } R \text{ falling in the leaf } L}. \quad (4)$$

Hence, the scoring function of a region R appears as a scalar product between the vector of leaf models π and a vector $\rho(R|\Sigma)$, which only depends on how the samples from the region R are sent to the leaves (Fig. 5). Moreover, since every pixel \mathbf{x} of a region R falls in exactly n_{trees} leaves, each vector $\rho(R|\Sigma)$ is sparse (or of small size) with at most $n_{\text{trees}} |R|$ non-zero elements.

4.2.2. Interactive Forest Adaptation

In Sec. 4.2.1, we described how to score regions of a large histological slide so that they can be ranked and displayed in decreasing order of interest to a pathologist. This ranking is based on the output of a pixelwise classification forest learned on labeled data. If the data at hand differs from the training images, for instance because of variations in terms of dye concentration or because of the presence of artifacts, this initial forest model can be prone to errors (Fig. 6). However, the fact that regions of interest are shown sequentially to the human expert offers the opportunity to let the user report the actual validity of the suggestions and, thereby, to recalibrate the forest model to

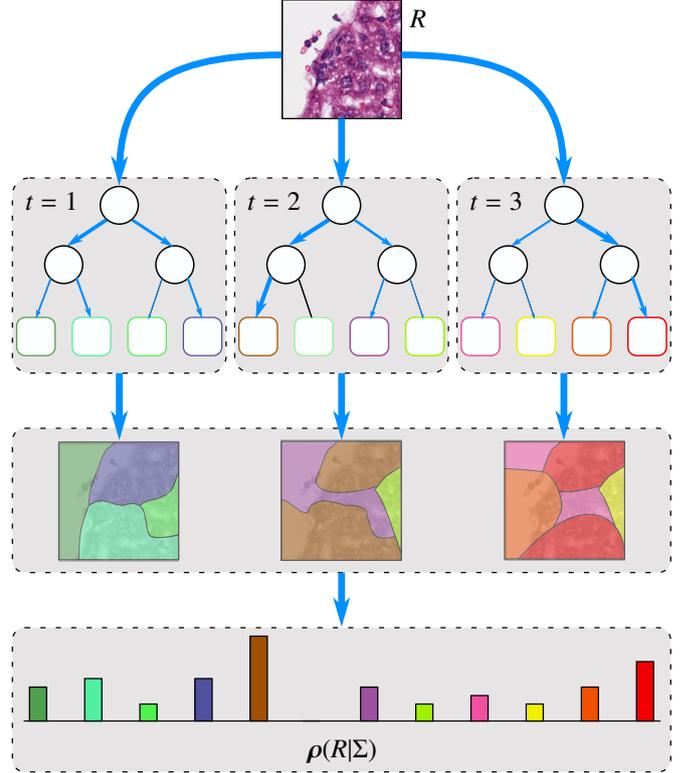


Figure 5: **Characteristic vector of a region R .** Applying a forest on all the pixels of a region R leads to n_{trees} different partitions of R , defined by the leaves reached by the sent pixels. By counting and concatenating the leaf occurrences, one obtains a characteristic vector $\rho(R|\Sigma)$ of the region R which only depends on the structure Σ of the forest (Eq. 4). Consequently, the score of any region R can be written as $\phi(R|\mathcal{F}) = \langle \rho(R|\Sigma), \pi \rangle$.

take into account the characteristics of the slide to analyze. This scenario corresponds to an online domain adaptation problem, for which we consider three different strategies. The first two require accurate delineations of the objects of interest by the user, whereas the third approach only requires a weaker form of labeling stating the actual surface covered by such objects within a suggested region. By discretizing this quantity, faster user interactions can be performed (see Sec. 6).

As exposed at the beginning of Sec. 4, the adaptation procedure generates, starting from a forest \mathcal{F}^1 , a series of forests $\mathcal{F}^2, \mathcal{F}^3, \dots$ where each forest \mathcal{F}^{k+1} is created after k regions have been observed by the pathologist and the k corresponding inputs collected. At each iteration k , the region R_k is chosen as the one maximizing the scoring function $\phi(\cdot|\mathcal{F}^k)$ over the set of remaining regions. The three alternative strategies described below are based on an assumption of fixed structure for all the forests \mathcal{F}^k , so that only the leaf probabilities are modified. This assumption offers the following computational advantage. For all $k \geq 1$, the structure Σ^k of the forest \mathcal{F}^k is equal to the structure Σ^1 of the initial forest \mathcal{F}^1 . In particular, the vectors $\rho(R|\Sigma^k), R \in \mathcal{R}$ are now kept unchanged during the whole exploration process, so that they can be precomputed once for all at the first iteration and compactly stored in memory due to their sparsity. For simplicity, we omit their dependency in Σ^1 and denote these vectors $\rho(R), R \in \mathcal{R}$. The score

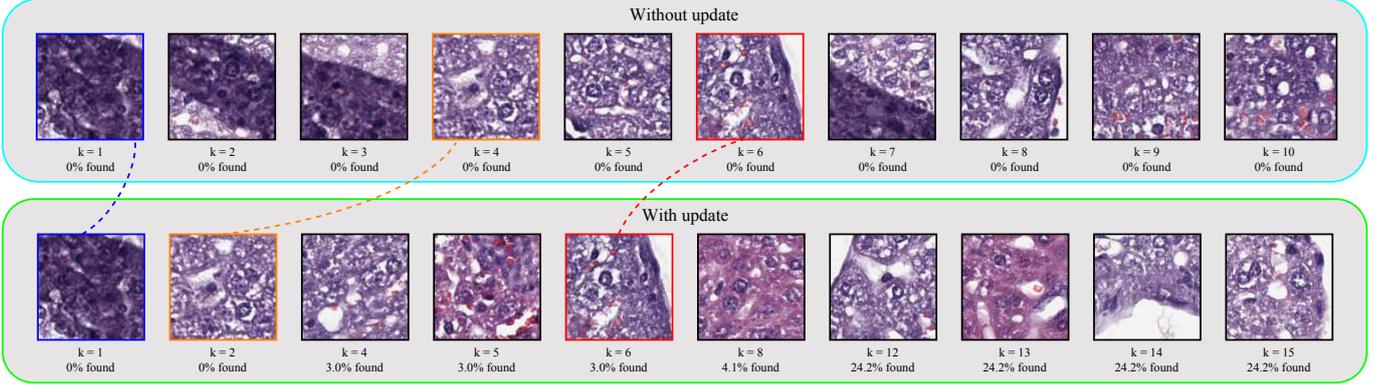


Figure 6: **Benefit of the interactive adaptation illustrated by the negative region suggestions.** In challenging cases (here the slide containing the portion shown in Fig. 2), a direct application of the pre-trained forest model leads to difficulties in the exploration. The first 10 regions that are suggested are in fact negative (first row). These confusions are due to the variation of staining and the presence of an artifact shown in Fig. 2, which confirms the necessity of adapting the initial classifier. When interactive updates are performed, each negative suggestion is signaled by the user and the underlying model is accordingly calibrated. The benefit of this adaptation scheme can be illustrated by looking at the first 10 negative suggestions (second row, using the ALM update). First, we can see that positive suggestions occur earlier in the exploration: the 3rd, 7th, 9th, 10th and 11th suggestions were positive and allowed the pathologist to localize already 24.2% of the hematopoietic patterns present in the slide. This demonstrates that adapting the classifier clearly improves the quality of the exploration. Moreover, we can see that the adaptive approach ‘learns from its mistakes’ through the greater diversity of its negative suggestions: regions with different shades of stain are proposed and multiple suggestions within the dark artifacts are avoided. Among their 10 first negatives, only 3 suggestions are shared by the two approaches (emphasized with colors), which further illustrates their difference.

$\phi(R|\mathcal{F}^k) = \langle \rho(R|\Sigma^k), \pi^k \rangle$ of a region R described in Eq. 3 can be rewritten $\phi(R|\mathcal{F}^k) = \langle \rho(R), \pi^k \rangle$. Hence, to obtain the updated scoring functions $\phi(R|\mathcal{F}^k)$ for a new forest \mathcal{F}^k , one only needs to recompute the sparse scalar products $\langle \rho(R), \pi^k \rangle$ with the new leaf models π^k . The efficiency of this operation yields real-time updates between two region suggestions, and the histological slide has to be passed only once through a forest (the initial \mathcal{F}^1) as a preliminary step before starting the exploration.

We expose now our three alternative real-time update strategies. They are all equivalent from a computational point of view, with a worst-case complexity of $O(n_{\text{trees}} |R|)$, and depend on one hyperparameter $\lambda > 0$ weighting the importance of the prior knowledge in comparison to the newly observed samples.

Update of Leaf Statistics (ULS) If the input provided by the pathologist is a full object delineation in the displayed region R_k , the pixels $\mathbf{x} \in R_k$ can be seen as new training samples whose label is known. Therefore, the leaf statistics can be updated (Criminisi et al., 2012). We denote $N_L^{1,+}$ (resp. $N_L^{1,-}$) the number of positive (resp. negative) samples which arrived in the leaf L during the training of the initial forest \mathcal{F}^1 , leading to the leaf models $\pi_L^1 = \frac{N_L^{1,+}}{N_L^{1,+} + N_L^{1,-}}$. We also denote $N_L^{k+1,+}$ and $N_L^{k+1,-}$ the total number of positive and negative samples collected in the regions R_1, \dots, R_k . Given these quantities, the ULS strategy updates the probability of each leaf L after labeling the region R_k as

$$\pi_L^{k+1} = \frac{N_L^{k+1,+} + \lambda N_L^{1,+}}{N_L^{k+1,+} + N_L^{k+1,-} + \lambda (N_L^{1,+} + N_L^{1,-})}. \quad (5)$$

Average of the Leaf Models (ALM) In the same conditions than the update described above, we propose an alternative

leaf update consisting in computing a separate probability π_L^{new} based on the pixels observed in R_1, \dots, R_k only (i.e. originating from the test slide) and averaging it with the initial probability π_L^1 . By doing so, the choice of λ is made independent of the initial number of samples in the leaf. This update can be written in vectorial form as

$$\pi^{k+1} = \frac{1}{1 + \lambda} (\pi^{\text{new}} + \lambda \pi^1) \quad (6)$$

where, for each leaf L , $\pi_L^{\text{new}} = \frac{N_L^{k+1,+}}{N_L^{k+1,+} + N_L^{k+1,-}}$ if some new samples have been observed in the leaf L (i.e. if $N_L^{k+1,+} + N_L^{k+1,-} > 0$). Otherwise, we define $\pi_L^{\text{new}} = \pi_L^1$.

Online Gradient Descent (OGD) The two previous updates require a pixelwise labeling provided by the user. Instead, this last update method uses the (weaker) information $Q(R_k)$ stating the amount of positive pixels located in the region R_k . This quantity is, in fact, what the score $\phi(R_k|\mathcal{F}^k) = \langle \rho(R_k), \pi^k \rangle$ used to assess the relevance of the region R_k estimates (Sec. 4.2.1). We propose to measure the discrepancy between the true value $Q(R_k)$ revealed by the user and the prediction from the set of leaf models π with the squared loss $l_k(\pi) = (\langle \rho(R_k), \pi \rangle - Q(R_k))^2$. Hence, at iteration k , the incurred loss is $l_k(\pi^k)$. The convexity of the loss function l_k , which directly results from the linear rewriting of our scoring function (Eq. 3), allows us to see the update problem as an online convex optimization scenario (Shalev-Shwartz, 2012). We solve this problem via an online gradient descent strategy (Zinkevich, 2003), which leads to the update rule

$$\begin{aligned} \pi^{k+1} &= \Pi_{[0,1]^{\mathcal{L}}} [\pi^k - \eta \vec{\nabla} l_k(\pi^k)] \\ &= \Pi_{[0,1]^{\mathcal{L}}} [\pi^k - 2\eta (\langle \rho(R_k), \pi^k \rangle - Q(R_k)) \rho(R_k)], \end{aligned} \quad (7)$$

where η is a learning rate. $\Pi_{[0,1]^{|L|}} : \mathbb{R}^{|L|} \rightarrow [0, 1]^{|L|}$ is the projection operator on $[0, 1]^{|L|}$ which projects each individual component π_l of a vector $\boldsymbol{\pi} \in \mathbb{R}^{|L|}$ onto the set $[0, 1]$, ensuring that the leaf probabilities stay in $[0, 1]$ after each update. By transferring generic considerations on online gradient descent to our forest-based scenario (see Appendix B), we choose a learning rate η of the form

$$\eta = \frac{1}{2\lambda\delta^4} \sqrt{\frac{n_{\text{trees}} |L|}{|\mathcal{R}|}}, \quad (9)$$

where λ is a positive hyperparameter, $|L|$ (resp. n_{trees}) is the number of leaves (resp. trees) in the forest, $|\mathcal{R}|$ is the number of regions in the slide and $\delta \times \delta$ is the predefined region size.

The difference between ALM and ULS can be seen by considering their respective asymptotic behavior when the number of new samples increases. In Eq. 5, we have $\boldsymbol{\pi}^k \sim \boldsymbol{\pi}^{\text{new}}$, whereas, in Eq. 6, $\boldsymbol{\pi}^k$ always includes a fixed contribution from $\boldsymbol{\pi}^1$ regardless of the number of new samples which have been collected. These two variants correspond to the simplest way to update an existing tree (Criminisi et al., 2012), where we introduce a parameter weighting old and new training data. More sophisticated but computationally costly strategies would involve further splitting or the replacement of old trees by new ones (Saffari et al., 2009). In our case, since the original training data used to train \mathcal{F}^1 are no longer available at testing time, we keep the structure of the old trees which represent the only available prior knowledge about the quantification task. This emphasizes the scope of our adaptation procedure, which should be seen as adjusting a known supervised segmentation task (e.g. within a same clinical study) to the variations of visual appearances that may occur experimentally. However, a different task cannot be accommodated a priori and would first require the training of a new segmentation model. Finally, due to the nature of the required input, OGD can be used for a lighter kind of user interaction (see Sec. 6), which is not supported by existing forest online learning algorithms.

4.3. From Partial Exploration to Whole-Slide Quantification

4.3.1. Stopping the Exploration Stage

In practice, the exploration process is meant to be interrupted before seeing the whole slide. Since the amount of hematopoietic cell clusters is variable from a slide to another, some slides intrinsically require more time from the pathologist than others. Therefore, fixing in advance the number of iterations for the exploration would be inappropriate. Instead, we propose a stopping criterion based on the density of positive suggestions, and interrupt the exploration as soon as n_{stop} negative regions were suggested in a row, i.e. when most positive regions were presumably seen. Once the stopping criterion is reached, a whole-slide prediction can be made (Sec. 4.3.2).

The chosen value for n_{stop} is directly depending on the amount of time that the pathologist is ready to spend for the analysis. Strictly reasoning in terms of accuracy, it is always

preferable to let the pathologist see a maximum number of regions. Defining a recommended value for n_{stop} is hence subjective and result from a tradeoff between accuracy and human effort. Our experiments regarding the whole-slide quantification were conducted for several values of n_{stop} , encoding different amounts of effort that the pathologist is ready to invest.

4.3.2. Whole-Slide Quantification via Regression

Once the stopping criterion has been reached, we predict an estimate \hat{q} of the surface covered by hematopoietic cells within the whole slide with linear regression. Denoting K the total number of regions that have been seen during the exploration stage, a partial knowledge on \hat{q} is available via the quantity $q^{\text{labeled}} = \sum_{k=1}^K Q(R_k)$ obtained as the user annotated the regions R_1, \dots, R_K during the exploration. In addition, the updated forest model \mathcal{F}^{K+1} obtained at the end of the exploration phase provides a prediction $\phi(R|\mathcal{F}^{K+1})$ of the quantity of positive pixels in each region $R \in \mathcal{R}$ of the slide. In particular, this gives a total prediction $\Phi^{\text{total}} = \sum_{R \in \mathcal{R}} \phi(R|\mathcal{F}^{K+1})$ and a prediction on the labeled regions $\Phi^{\text{labeled}} = \sum_{k=1}^K \phi(R_k|\mathcal{F}^{K+1})$. We formalize our regression problem by considering that the relative change between the total quantity \hat{q} and the partial quantity q^{labeled} is proportional to the relative change between total prediction and the partial prediction, i.e.

$$\frac{\hat{q} - q^{\text{labeled}}}{q^{\text{labeled}}} \propto \frac{\Phi^{\text{total}} - \Phi^{\text{labeled}}}{\Phi^{\text{labeled}}}. \quad (10)$$

This corresponds to a prediction rule of the form

$$\hat{q} = q^{\text{labeled}} + a \frac{\Phi^{\text{total}} - \Phi^{\text{labeled}}}{\Phi^{\text{labeled}}} q^{\text{labeled}}. \quad (11)$$

The regression parameter $a \in \mathbb{R}$ is learned on a validation set.

5. Experiments

5.1. Dataset and Medical Motivation

The presence of hematopoietic cells outside the bone marrow, also called extramedullary hematopoiesis, is a marker of an extensive stimulation of the immune system (Tao et al., 2008). There is accumulating evidence that the amount of infiltrating immune cells such as cytotoxic CD8-positive T-lymphocytes into the tumor can be considered as a tumor biomarker for measuring clinical outcome (Balermipas et al., 2016). We evaluated our approach in this clinical context on a dataset addressing the aspect of lymphocytic infiltration into mouse liver tissues, for which the amount of these cells within histological samples must be estimated.

Slides from 16 mice were digitally acquired at the resolution $0.5 \mu\text{m}$ per pixel and downsized by 2 to speed up the training and testing steps. 70 large representative subimages were extracted from these slides and fully segmented, covering approximately 20% of the total tissue (Fig. 7). Resorting to a set of subimages follows the clinical practice and was necessary to obtain accurate labels for a sufficient number of different slides. This is particularly important in our study which

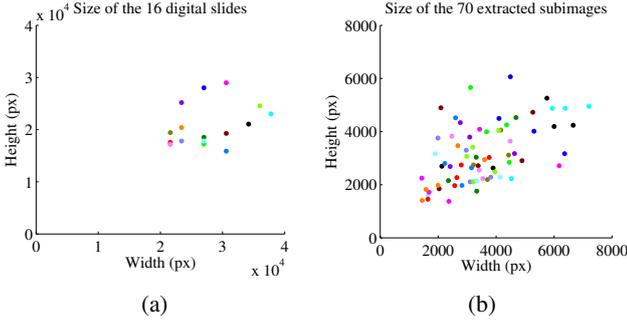


Figure 7: **Dataset dimensions.** Our dataset consists of 16 digital slides, an example of which is provided in Fig. 1. Figure 7a shows their dimensions, where each colored point corresponds to one slide. Instead of working on the slide, we extracted and labeled entirely a total of 70 subregions such as the one shown in Fig. 2. The dimensions of these subregions are reported in Fig. 7b. The color code of Fig. 7a is respected so that the slide from which each subimage is extracted can be identified by its color. In total, the extracted subimages cover around 20% of the acquired tissues.

focuses on issues arising from the visual variability between acquisitions. Yet, it comes at the cost of possibly introducing a few natural biases such as underrepresentation of border areas or of straightforwardly negative objects (e.g. large white parts).

5.2. Experimental Settings

We randomly split our dataset \mathcal{D} into 4 sets \mathcal{D}_i of 4 slides each and we performed a 4-fold nested cross-validation. The goal of this procedure is to optimize the update-related hyperparameter λ independently of the test data to avoid overfitting. Let Λ be a set of candidate values for λ and $l^v(\lambda, \mathcal{F}^1, I)$ a loss measuring the error of our method on the slide I when using the hyperparameter λ and an initial forest \mathcal{F}^1 . The nested cross-validation consists of 4 runs, each run corresponding to a set $\mathcal{D}_{i_{\text{out}}}$ (with $i_{\text{out}} \in \{1, 2, 3, 4\}$) left out for testing. For each run, a second cross-validation (called inner cross-validation) is performed over the 3 remaining sets $(\mathcal{D}_i)_{i \neq i_{\text{out}}}$, where 2 sets are used to train the forest and the remaining one is used as a validation set. At the end of the inner cross-validation, i.e. when 3 forests have been trained and each of the 3 sets has been used as a validation set, we define the optimal hyperparameter $\lambda_{i_{\text{out}}}$ of this run as the one minimizing the total loss over the 3 bags, i.e.

$$\lambda_{i_{\text{out}}} = \underset{\lambda \in \Lambda}{\operatorname{argmin}} \sum_{i \neq i_{\text{out}}} \sum_{I \in \mathcal{D}_i} l^v(\lambda, \mathcal{F}_{\mathcal{D} \setminus (\mathcal{D}_i \cup \mathcal{D}_{i_{\text{out}}})}, I). \quad (12)$$

$\mathcal{F}_{\mathcal{D} \setminus (\mathcal{D}_i \cup \mathcal{D}_{i_{\text{out}}})}$ denotes the forest obtained by training on the two remaining sets after excluding \mathcal{D}_i and $\mathcal{D}_{i_{\text{out}}}$. Using the hyperparameter value $\lambda_{i_{\text{out}}}$, we then report independently the prediction of each of the 3 forests $(\mathcal{F}_{\mathcal{D} \setminus (\mathcal{D}_i \cup \mathcal{D}_{i_{\text{out}}})})_{i \neq i_{\text{out}}}$ on the left-out set $\mathcal{D}_{i_{\text{out}}}$. This procedure allows us to learn automatically the hyperparameter independently of the testing set, and outputs 3 different predictions for each test slide which gives an idea of their dependency on the original training data. This results in a total of 48 predictions. Note that, to conduct the entire nested cross-validation, only 6 forests $(\mathcal{F}_{\mathcal{D}_i \cup \mathcal{D}_j})_{1 \leq i < j \leq 4}$ have to be trained. The user interaction was automatically simulated from the ground truth delineations. Regions were chosen of

size $\delta \times \delta$ with $\delta = 60 \mu\text{m}$. Every time a region is displayed, the user can easily extend the field of view around it if necessary. We simulated this behavior automatically by showing the neighboring positive region(s) in the case where an object of interest is not fully included in the displayed region.

The forests were initially trained on labeled pixels which were densely collected every $8 \mu\text{m}$ in the two directions. Parallelized on 10 threads, this training step took between 3 and 6 hours depending on the cross-validation run (with a corresponding number of training samples comprised between 5×10^5 and 10^6). Given an incoming slide, testing was performed on all pixels, which is tractable since it has to be done only once at the beginning of the process (see Sec. 4.2.2). This preliminary step took around 1 minute, after which the interaction loop could take place in real-time conditions. More precisely, the update of forest leaf models and recomputation of box scores between two iterations took between 10 and 100 ms without any parallelization. The visual features were computed on the Lab color space. The following forest parameters were used: $n_{\text{samples/leaf}} = 10$, $n_{\text{trees}} = 30$, $n_{\text{tries}} = 500$, $n_{\text{thresholds}} = 10$, and the bagging rate was 0.5.

Due to the efficiency of Haar-like features, our segmentation algorithm is able to work directly at the highest level of magnification, processing approximately 2.0×10^7 pixels per minute. This order of magnitude is, for instance, the same as in a recent boosting-based hierarchical segmentation approach (Doyle et al., 2012) which analyzes around 1.4×10^7 pixels in less than 3 minutes (with parallelization on 2 threads instead of 10). However, this latter work used more complex features for their application, hence justifying a hierarchical strategy.

5.3. Evaluation of the Exploration Stage

We studied the ability of our approach to retrieve regions of interest as quickly as possible within large slides. The following alternatives were compared:

- the simplest forest-based exploration approach without update from the pathologist (No Update (Forest)), i.e. only relying on the pre-trained forest \mathcal{F}^1 ,
- the three update strategies (ULS, ALM and OGD) exposed in Sec. 4.2.2,
- a baseline showing, at each iteration, a region randomly (uniformly) drawn among the remaining regions (Random exploration),
- an oracle whose scoring function is extracted from the ground truth, hence serving as a gold standard showing the highest achievable performance (Oracle).

In addition, to position the forest-based performance among other classification methods, we trained an AdaBoost classifier and used it as segmentation model instead of the forest (No Update (AdaBoost)). The chosen weak classifiers were decision stumps based on Haar-like features such as the splitting functions stored in tree nodes. 100 boosting iterations were conducted, and 500 stumps tried at each iteration. These design choices led to a training time similar to the forest one.

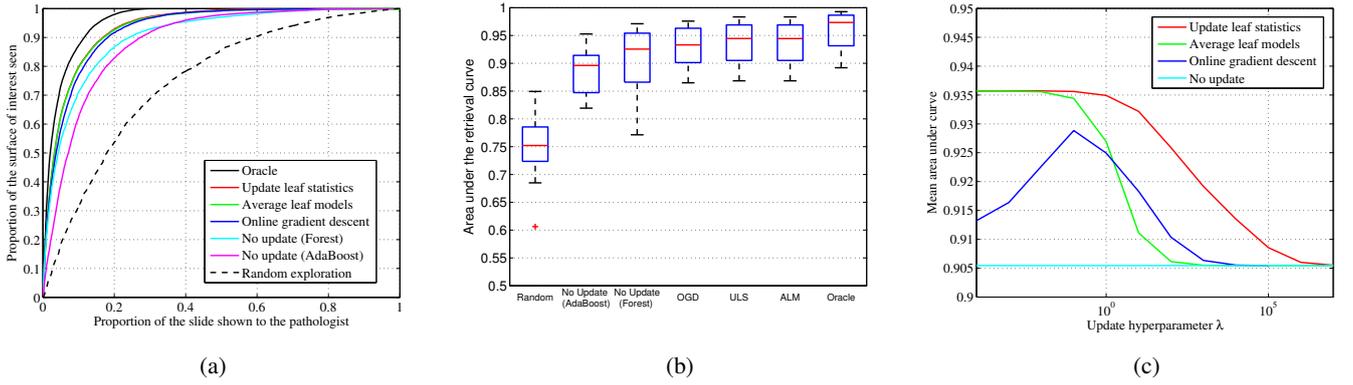


Figure 8: **Experimental validation of the slide exploration stage.** (a) For each method, we plot the mean retrieval curve which shows the proportion of positive pixels seen by the pathologist after having seen a certain proportion of the slide. We use the area under these curves to measure quantitatively the slide exploration abilities. (b) Statistical distribution of the area under the curve for each method. Each box plot is computed over the 48 measurements obtained during the nested cross-validation. (c) Influence of the hyperparameter λ on the performance. For each of the three update strategies, we studied the behavior of the mean area under the curve when λ varies, i.e. without optimization on a validation set.

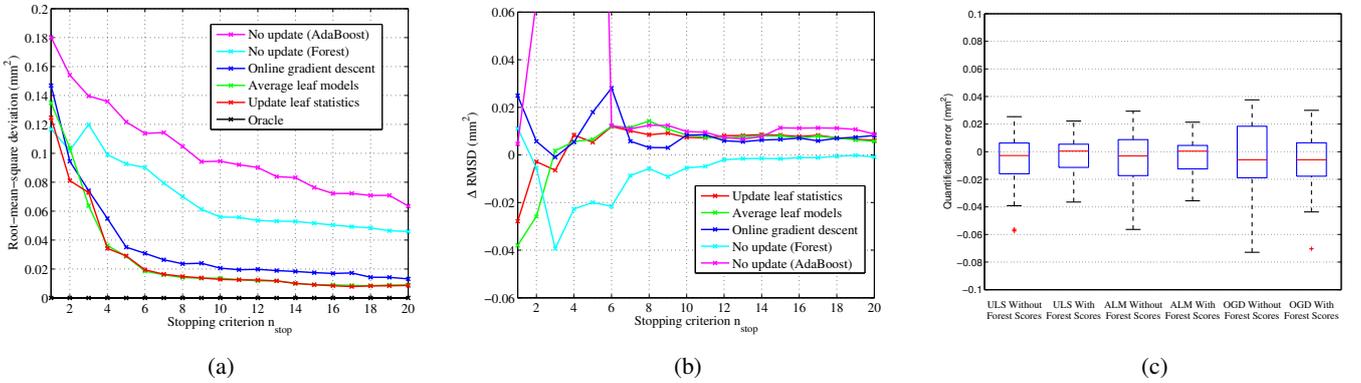


Figure 9: **Experimental validation of the whole-slide quantification stage.** (a) We report the root-mean-square deviation for several values of the stopping criterion n_{stop} which represents the amount of interaction provided by the pathologist. (b) Difference ΔRMSD between the root-mean-square deviation obtained when using a regression of the form $\hat{q} = (1 + a)q^{\text{labeled}}$ based on the user inputs alone, and the one obtained with the regression of Eq. 11 whose results are reported in Fig. 9a. Positive values of ΔRMSD correspond to a gain of accuracy when including the forest scores Φ^{total} and Φ^{labeled} in the regression task. As soon as n_{stop} is large enough to obtain stable results, the information carried by an updated forest improves the performance. (c) Distribution of the signed quantification error over slides for $n_{\text{stop}} = 10$, with and without including the forest scores in the regression.

To assess quantitatively the performance of each method, we consider the curve showing the proportion of positive pixels that have been displayed after having shown a certain percentage of the slide to the pathologist (Fig. 8a). A good exploration method is expected to lead to a curve converging quickly towards 1. We summarize quantitatively the performance on a slide I by computing the area $A(\lambda, \mathcal{F}^0, I)$ under this curve. The nested cross-validation procedure described in Sec. 5.2 was accordingly performed using the loss function $l^{\text{cv}}(\lambda, \mathcal{F}^1, I) = 1 - A(\lambda, \mathcal{F}^1, I)$ and optimizing λ over a logarithmic grid. The statistical distribution of the area under curves obtained at prediction time for each method are shown in Fig. 8b. We performed statistical pairwise comparisons between methods by conducting paired Wilcoxon’s signed-rank tests over these values. To maintain the independence between samples, we repeated each test 100 times retaining at random one of the 3 runs for each slide and considered the median p-value over these 100 runs. Denoting Method 1 < Method 2

the fact that Method 2 is significantly better than Method 1 and Method 1 \approx Method 2 the absence of demonstrated statistical difference between the two methods, the series of tests provided the following ranking:

Random < No Update (AdaBoost) < No Update (Forest) < OGD < ULS \approx ALM < Oracle.

All p-values showing statistical difference were lower than 10^{-3} , and the p-value obtained when comparing ULS and ALM was 0.5. This ranking confirms what was intuitively expected. The three methods proposing a model update from the user inputs improve over a non-interactive exploration, and the two methods using accurate pixelwise labelings outperform the online gradient descent technique which is based on a weaker but lighter type of information. We also see, from the performance of a random exploration, that using a pre-trained forest drastically helps finding relevant objects more quickly. Note that, in theory, one might have expected a straight ‘y = x’ line for the random exploration. In fact, when a suggested region be-

longs to a larger object, the user extends the field of view to see the object in its totality. Hence, a positive suggestion may be immediately followed by other positive ones due to the user intervention. This bias explains why, in spite of a random exploration, one obtains a slightly ‘better than random’ curve.

The impact of the update-related parameter λ was assessed experimentally (Fig. 8c). As expected, when $\lambda \rightarrow \infty$, the three methods converge towards the method without update. The choice $\lambda = 0$ leads to a nearly maximal performance for the two update strategies based on pixelwise labelings (ULS and ALM). Since, moreover, these two methods are equivalent for $\lambda = 0$, they behave very similarly after optimization on a validation set, as observed on Fig. 8a and Fig. 8b. Note that choosing $\lambda = 0$ does not mean that the initial forest \mathcal{F}^1 is ignored. The prior knowledge contained in \mathcal{F}^1 is used through both the tree structures and their leaf models. Moreover, a leaf remains unchanged as long as it does not appear in a selected region, which may happen if it accurately predicts background areas.

5.4. Evaluation of the Whole-Slide Quantification Stage

To assess, at prediction time, the accuracy of a list of n_{pred} estimates $(\hat{q}_i)_{1 \leq i \leq n_{\text{pred}}}$ given the corresponding true quantities $(q_i)_{1 \leq i \leq n_{\text{pred}}}$, we use the root-mean-square deviation

$$\text{RMSD} = \sqrt{\frac{1}{n_{\text{pred}}} \sum_{i=1}^{n_{\text{pred}}} (q_i - \hat{q}_i)^2}. \quad (13)$$

In our case, $n_{\text{pred}} = 48$. This corresponds to the 3 predictions obtained for each of the 16 slides during the cross-validation.

The experimental evaluation of the whole-slide quantification abilities was conducted as follows. We kept the cross-validation setup described in Sec. 5.2 and learned the hyperparameter a on a validation set, as was done for the update parameter λ , using here a squared loss $\ell^{\text{cv}}(a, \mathcal{F}^1, I) = (q_I - \hat{q}_I)^2$ between true and predicted whole-slide estimates. This procedure was performed independently for several values of the stopping criterion n_{stop} . We report the resulting curves in Fig. 9a and an example of the correspondence between estimates and true quantities in Fig. 10. The ranking of methods obtained while studying the exploration abilities is preserved, due to the fact that the quality of the exploration phase is directly linked to the amount of regions which are eventually labeled. Asymptotically, if all regions containing positive samples are labeled, choosing $a = 0$ provides a perfect prediction.

During our experiments, we observed that the sum of segmentation probabilities over the whole slide, i.e. predicting $\hat{q} = \Phi^{\text{total}}$ (with the notation of Eq. 11), does not form a reliable whole-slide quantification and overestimates the quantity of positive pixels due to two effects. First, the random nature of trees leads in general to small nonzero probabilities on negative pixels. When summed over all pixels, these small errors aggregate. Moreover, since we mainly retrieve positive examples during the slide exploration, the distribution of incoming samples for the update is strongly biased towards positive instances. These difficulties motivate the use of a regression. In Fig. 9b and Fig. 9c, we show that our regression approach (Eq. 11) outperforms a regression based on the user inputs alone of the form

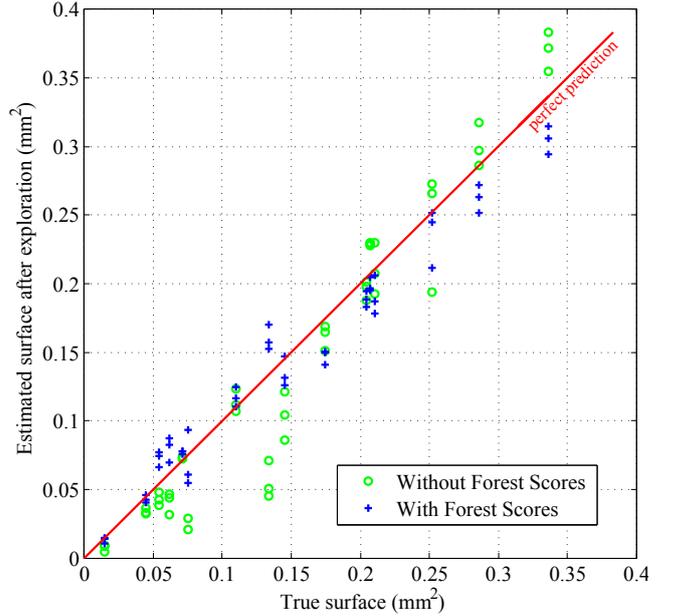


Figure 10: **Correlation between estimated and true hematopoietic surface within whole slides.** We show an overview on the whole dataset of the whole-slide estimates after exploration. Each slide tested during the cross-validation appears 3 times corresponding to different initial forests (Sec. 5.2). Perfect predictions would lie on the red line. This example was obtained using the ULS adaptation for $n_{\text{stop}} = 6$, leading to $\text{RMSD} = 1.9 \times 10^{-2} \text{ mm}^2$. The median proportion of the slide which was seen during the exploration was 5.2% in this case. The predictions obtained without using the final forest scores (see Fig. 9b and Fig. 9c) are also reported.

$\hat{q} = (1 + a)q^{\text{labeled}}$. This demonstrates that, in spite of its global overestimation, the forest estimate can be effectively exploited by a regression procedure.

6. Input Discretization for Lighter Interactions

In the experiments presented in Sec. 5, the forest adaptation techniques assumed that the pathologist provides a full pixelwise labeling of the objects of interest in the displayed regions. In this section, we demonstrate how our OGD scheme can be used with one-click inputs instead without decreasing its performance, thereby allowing faster user interaction.

6.1. One-Click User Inputs

Unlike the two other techniques based on individual labels for each pixel in a region, the OGD forest update employs as user input for a region R_k the amount of positive pixels $Q(R_k)$ contained in R_k . This is a different kind of input, which can be inferred from a delineation or communicated directly instead. Here, we propose to discretize the input values into bins and ask the user to select the bin to which the proportion of positive pixels belongs. This interaction is performed with only one click, or possibly without a mouse (e.g. via voice recognition).

Formally, the user annotations are discretized as follows. Instead of providing the exact quantity $Q(R_k)$, the user simply indicates an interval within which the proportion $\tilde{Q}(R_k) = \frac{Q(R_k)}{|R_k|}$

lies. The list of available ranges is predefined as $\{0\}$, $]0; \frac{1}{m}]$, \dots , $] \frac{m-2}{m}; \frac{m-1}{m}]$, $] \frac{m-1}{m}; 1[$, $\{1\}$, where m is a positive integer which encodes the fineness of the quantization. Accordingly, by taking the middle-value of each bin, the proportions $\tilde{Q}(R_k)$ provided by the user take their values in the finite set $D_m = \{0, \frac{1}{2m}, \frac{3}{2m}, \dots, \frac{2m-1}{2m}, 1\}$. This gives in total $m + 2$ input possibilities for the discrete input $\tilde{Q}(R_k)$, including the 2 trivial ones corresponding to an empty ($\tilde{Q}(R_k) = 0$) or a full ($\tilde{Q}(R_k) = 1$) region. By doing so, only one click per region is required from the pathologist, resulting in a lesser amount of interactions.

Given a discrete region label $\tilde{Q}(R_k) \in D_m$ provided by the user, we have to compute the actual quantity $Q(R_k)$ eventually used in the adaptation process (Eq. 8) and recorded for an eventual whole-slide quantification (Eq. 11). The simplest idea consists in directly taking $Q(R_k) = |R_k| \tilde{Q}(R_k)$, but has the drawback of losing information due to the discretization. To attenuate this aspect, we propose to perform updates only if the forest estimate $\phi(R_k|\mathcal{F}^k)$ (whose objective is to predict the quantity $Q(R_k)$) deviates too strongly from the user label. More precisely, we define $Q(R_k) = |R_k| \tilde{Q}(R_k)$ if $\left| \frac{\phi(R_k|\mathcal{F}^k)}{|R_k|} - \tilde{Q}(R_k) \right| \geq \frac{1}{2m}$, and $Q(R_k) = \phi(R_k|\mathcal{F}^k)$ otherwise. In other words, we fully trust the forest estimate as long as it leads to the same bin as the one indicated by the user. This distinction is only made if the label is ambiguous, i.e. different than 0 and 1. Otherwise, it corresponds in fact to an exact labeling and is treated as such.

6.2. Evaluation

The experiments involving online gradient descent presented in Sec. 5.3 and Sec. 5.4 were repeated using these discretized inputs instead of the exact ones, for every level of quantization $m \in \{1, \dots, 5\}$. In terms of retrieval performance, discretizing the inputs does not show any clear difference in comparison to the use of exact user inputs, and this from $m = 1$ on (Fig. 11). For $m = 1$, the paired Wilcoxon’s signed-rank test leads to a p-value of 0.65. Additionally, if we assume the differences to be normally distributed, the confidence interval for the mean difference is $[-4.8 \times 10^{-3}, 2.3 \times 10^{-3}]$. Hence, by making available to the user 3 buttons (corresponding to the choice $m = 1$) stating respectively whether a region is empty, full of hematopoietic cells or partially covered, the exploration phase is of equivalent quality as the one provided by the online gradient descent method with accurate user labelings.

Since the task of whole-slide quantification from the exploration phase (Sec. 5.4) relies strongly on the user inputs $\tilde{Q}(R_k)$ (see Eq. 11), obtaining satisfactory results for this task with discretized inputs requires a more accurate quantization. This minimum level was experimentally found to be $m = 3$, which remains nevertheless tractable in practice (Fig. 12).

7. Conclusion

We introduced an interactive framework able to help a pathologist to navigate efficiently through large digital slides. Our approach is based on a pixelwise random forest classifier pre-trained to segment objects of interest within the tis-

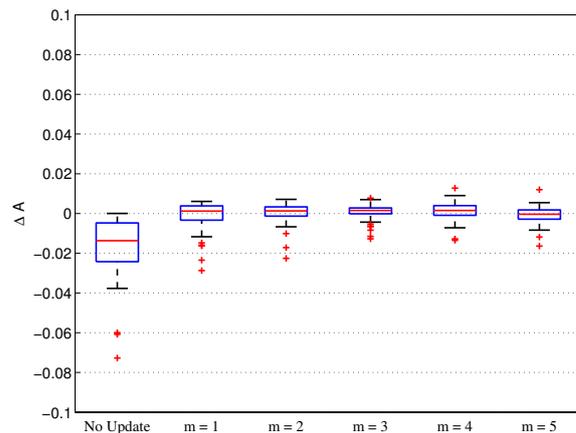


Figure 11: **Impact of input quantization on the exploration phase.** We consider the difference ΔA of area under the curve observed when using discretized user inputs instead of exact ones in the OGD adaptation. The parameter m encodes the fineness of the quantization. No statistical difference is observed from $m = 1$ on. To put the variations in perspective, we also report the difference obtained when no updates are performed.

sue whose predictions are used to score, rank and display regions according to their expected interest. By allowing the user to provide labels on each suggested region, the leaf nodes of the forest model are adjusted in real time during the exploration procedure so that visual specificities of the data at hand can be gradually incorporated into the region selection process. For this purpose, in addition to two standard leaf update techniques, we introduced a novel adaptation scheme based on online gradient descent which supports one-click inputs from the pathologist instead of more tedious accurate object delimitations. Experimental validation was conducted on the task of extramedullary hematopoiesis quantification within mouse liver slides. Beyond its slide exploration abilities, we demonstrated how our method can successfully exploit both the forest segmentation output and the labels collected during the exploration stage to provide accurate estimates of the surface covered by hematopoietic cells in the whole slide.

Acknowledgments

This work was partially supported by the Collaborative Research Center 824: “Imaging for Selection, Monitoring and Individualization of Cancer Therapies”. We are also very grateful to the pathologist Dr. Katja Steiger for supervising the acquisition and labeling of the data used in this study.

References

- Al-Janabi, S., Huisman, A., Van Diest, P.J., 2012. Digital pathology: current status and future perspectives. *Histopathology* 61, 1–9.
- Bahlmann, C., Patel, A., Johnson, J., Ni, J., Chekkoury, A., Khurd, P., Kamen, A., Grady, L., Krupinski, E., Graham, A., et al., 2012. Automated detection of diagnostically relevant regions in H&E stained digital pathology slides,

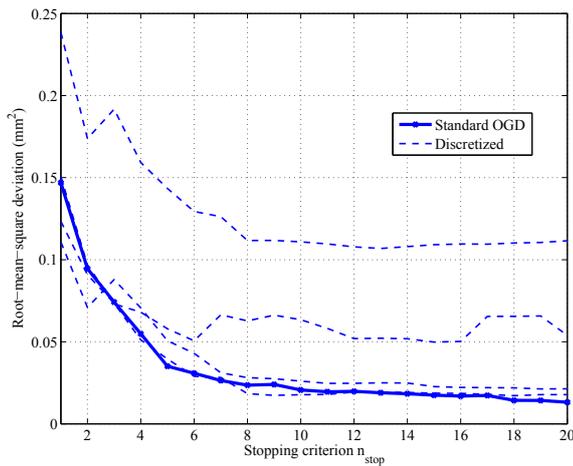


Figure 12: **Whole-slide quantification from discrete inputs.** From $m \geq 3$, the discretized approach reaches similar quantification performances as an OGD update with exact inputs.

in: SPIE Medical Imaging, International Society for Optics and Photonics. pp. 831504–831504–8.

Balermas, P., Rödel, F., Rödel, C., Krause, M., Linge, A., Lohaus, F., Baumann, M., Tinhofer, I., Budach, V., Gkika, E., Stuschke, M., Avlar, M., Grosu, A.L., Abdollahi, A., Debus, J., Bayer, C., Stangl, S., Belka, C., Pigorsch, S., Multhoff, G., Combs, S.E., Mönnich, D., Zips, D., Fokas, E., 2016. CD8+ tumour-infiltrating lymphocytes in relation to HPV status and clinical outcome in patients with head and neck cancer after postoperative chemoradiotherapy: A multicentre study of the german cancer consortium radiation oncology group (DKTK-ROG). *International Journal of Cancer* 138, 171–181.

Bauer, T.W., Schoenfeld, L., Slaw, R.J., Yerian, L., Sun, Z., Henricks, W.H., 2013. Validation of whole slide imaging for primary diagnosis in surgical pathology. *Archives of Pathology & Laboratory Medicine* 137, 518–524.

Bautista, P.A., Yagi, Y., 2015. Staining correction in digital pathology by utilizing a dye amount table. *Journal of Digital imaging*, 1–12.

Breiman, L., Friedman, J., Stone, C., Olshen, R., 1984. Classification and regression trees.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al., 2006. Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. *Genome biology* 7, R100.

Chatelain, P., Pauly, O., Peter, L., Ahmadi, S.A., Plate, A., Bötzel, K., Navab, N., 2013. Learning from multiple experts with random forests: Application to the segmentation of the midbrain in 3D ultrasound, in: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, pp. 230–237.

Cooper, L.A., Carter, A.B., Farris, A.B., Wang, F., Kong, J., Gutman, D.A., Widener, P., Pan, T.C., Cholleti, S.R., Sharma, A., et al., 2012. Digital pathology: Data-intensive frontier in medical imaging. *Proceedings of the IEEE* 100, 991–1003.

Criminisi, A., Shotton, J., Bucciarelli, S., 2009. Decision forests with long-range spatial context for organ localization in CT volumes, in: *MICCAI Workshop on Probabilistic Models for Medical Image Analysis*.

Criminisi, A., Shotton, J., Konukoglu, E., 2012. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trends® in Computer Graphics and Vision* 7, 81–227.

Doyle, S., Feldman, M., Tomaszewski, J., Madabhushi, A., 2012. A boosted bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE Transactions on Biomedical Engineering* 59, 1205–1218.

Ebner, T., Stern, D., Donner, R., Bischof, H., Urschler, M., 2014. Towards automatic bone age estimation from MRI: Localization of 3D anatomical

landmarks, in: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, pp. 429–437.

Eefting, D., Schrage, Y.M., Geirnaerd, M.J., Le Cessie, S., Taminiau, A.H., Bovée, J.V., Hogendoorn, P.C., et al., 2009. Assessment of interobserver variability and histologic parameters to improve reliability in classification and grading of central cartilaginous tumors. *The American Journal of Surgical Pathology* 33, 50–57.

Farahani, N., Parwani, A., Pantanowitz, L., 2015. Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International* 7, 23–33.

Fiaschi, L., Köthe, U., Nair, R., Hamprecht, F.A., 2012. Learning to count with regression forest and structured labels, in: *International Conference on Pattern Recognition (ICPR)*, IEEE. pp. 2685–2688.

Gauriau, R., Cuingnet, R., Lesage, D., Bloch, I., 2014. Multi-organ localization combining global-to-local regression and confidence maps, in: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, pp. 337–344.

Gilles, F.H., Tavaré, C.J., Becker, L.E., Burger, P.C., Yates, A.J., Pollack, I.F., Finlay, J.L., 2007. Pathologist interobserver variability of histologic features in childhood brain tumors: results from the CCG-945 study. *Pediatric and Developmental Pathology* 11, 108–117.

Gonul, I.I., Poyraz, A., Unsal, C., Acar, C., Alkibay, T., 2006. Comparison of 1998 WHO/ISUP and 1973 WHO classifications for interobserver variability in grading of papillary urothelial neoplasms of the bladder. *Pathological evaluation of 258 cases. Urologia Internationalis* 78, 338–344.

Gorelick, L., Veksler, O., Gaed, M., Gomez, J., Moussa, M., Bauman, G., Fenster, A., Ward, A., 2013. Prostate histopathology: Learning tissue component histograms for cancer detection and classification. *IEEE Transactions on Medical Imaging* 32, 1804–1818.

Gurcan, M., Boucheron, L., Can, A., Madabhushi, A., Rajpoot, N., Yener, B., 2009. Histopathological image analysis: A review. *IEEE Reviews in Biomedical Engineering* 2, 147–171.

Held, M., Schmitz, M.H., Fischer, B., Walter, T., Neumann, B., Olma, M.H., Peter, M., Ellenberg, J., Gerlich, D.W., 2010. Cellcognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nature Methods* 7, 747–754.

Homeyer, A., Schenk, A., Arlt, J., Dahmen, U., Dirsch, O., Hahn, H.K., 2013. Practical quantification of necrosis in histological whole-slide images. *Computerized Medical Imaging and Graphics* 37, 313–322.

Homeyer, A., Schenk, A., Dahmen, U., Dirsch, O., Huang, H., Hahn, H.K., 2011. A comparison of sampling strategies for histological image analysis. *Journal of Pathology Informatics* 2.

Huang, C.H., Veillard, A., Roux, L., Loménie, N., Racoceanu, D., 2011. Time-efficient sparse analysis of histopathological whole slide images. *Computerized Medical Imaging and Graphics* 35, 579–591.

Jaarsma, T., Jarodzka, H., Nap, M., Merriënboer, J.J., Boshuizen, H., 2014. Expertise under the microscope: processing histopathological slides. *Medical Education* 48, 292–300.

Jain, V., Learned-Miller, E., 2011. Online domain adaptation of a pre-trained cascade of classifiers, in: *IEEE Computer Vision and Pattern Recognition (CVPR)*, pp. 577–584.

Jukić, D.M., Drogowski, L.M., Martina, J., Parwani, A.V., 2011. Clinical examination and validation of primary diagnosis in anatomic pathology using whole slide digital images. *Archives of Pathology & Laboratory Medicine* 135, 372–378.

Khan, A.M., Rajpoot, N., Treanor, D., Magee, D., 2014. A non-linear mapping approach to stain normalisation in digital histopathology images using image-specific colour deconvolution. *IEEE Transactions on Biomedical Engineering*.

Khurd, P., Bahlmann, C., Maday, P., Kamen, A., Gibbs-Strauss, S., Genega, E., Frangioni, J., 2010. Computer-aided gleason grading of prostate cancer histopathological images using texton forests, in: *IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 636–639.

Kong, H., Gurcan, M., Belkacem-Boussaid, K., 2011. Partitioning histopathological images: An integrated framework for supervised color-texture segmentation and cell splitting. *IEEE Transactions on Medical Imaging* 30, 1661–1677.

Kontschieder, P., Dorn, J., Morrison, C., Corish, R., Zikic, D., Sellen, A., DSouza, M., Kamm, C.P., Burggraaf, J., Tewarie, P., et al., 2014. Quantifying progression of multiple sclerosis via classification of depth videos, in: *Medical Image Computing and Computer Assisted Intervention (MIC-*

CAI), Springer. pp. 429–437.

Lakshminarayanan, B., Roy, D.M., Teh, Y.W., 2014. Mondrian forests: Efficient online random forests, in: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K. (Eds.), *Advances in Neural Information Processing Systems (NIPS)*, pp. 3140–3148.

Macenko, M., Niethammer, M., Marron, J., Borland, D., Woosley, J., Guan, X., Schmitt, C., Thomas, N., 2009. A method for normalizing histology slides for quantitative analysis, in: *IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1107–1110.

Mercan, E., Aksoy, S., Shapiro, L.G., Weaver, D.L., Brunye, T., Elmore, J.G., 2014. Localization of diagnostically relevant regions of interest in whole slide images, in: *IEEE International Conference on Pattern Recognition (ICPR)*, pp. 1179–1184.

Meyer, J.S., Alvarez, C., Milikowski, C., Olson, N., Russo, I., Russo, J., Glass, A., Zehnbaauer, B.A., Lister, K., Parwaresch, R., 2005. Breast carcinoma malignancy grading by bloom–richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index. *Modern Pathology* 18, 1067–1078.

Montillo, A., Shotton, J., Winn, J., Iglesias, J., Metaxas, D., Criminisi, A., 2011. Entangled decision forests and their application for semantic segmentation of CT images, in: *Information Processing in Medical Imaging (IPMI)*. Springer, pp. 184–196.

Nemirovski, A., Juditsky, A., Lan, G., Shapiro, A., 2009. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization* 19, 1574–1609.

Nguyen, K., Sarkar, A., Jain, A., 2014. Prostate cancer grading: use of graph cut and spatial arrangement of nuclei. *IEEE Transactions on Medical Imaging* 33, 2254.

Onder, D., Zengin, S., Sarioglu, S., 2014. A review on color normalization and color deconvolution methods in histopathology. *Applied Immunohistochemistry & Molecular Morphology* 22, 713–719.

Pauly, O., Glocker, B., Criminisi, A., Mateus, D., Möller, A., Nekolla, S., Navab, N., 2011. Fast multiple organ detection and localization in whole-body MR dixon sequences, in: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, pp. 239–247.

Peter, L., Mateus, D., Chatelain, P., Schworm, N., Stangl, S., Multhoff, G., Navab, N., 2014. Leveraging random forests for interactive exploration of large histological images, in: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, pp. 1–8.

Peter, L., Pauly, O., Chatelain, P., Mateus, D., Navab, N., 2015. Scale-adaptive forest training via an efficient feature sampling scheme, in: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer, pp. 637–644.

Rabinovich, A., Agarwal, S., Laris, C., Price, J.H., Belongie, S.J., 2003. Un-supervised color decomposition of histologically stained tissue samples, in: *Advances in Neural Information Processing Systems (NIPS)*, pp. 667–674.

Roullier, V., L  zoray, O., Ta, V.T., Elmoataz, A., 2011. Multi-resolution graph-based analysis of histopathological whole slide images: Application to mitotic cell extraction and visualization. *Computerized Medical Imaging and Graphics* 35, 603–615.

Saffari, A., Leistner, C., Santner, J., Godec, M., Bischof, H., 2009. On-line random forests, in: *IEEE International Conference on Computer Vision Workshops*, pp. 1393–1400.

Sertel, O., Kong, J., Shimada, H., Catalyurek, U., Saltz, J.H., Gurcan, M.N., 2009. Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development. *Pattern Recognition* 42, 1093–1103.

Settles, B., 2010. Active learning literature survey. *Computer Sciences Technical Report* 1648.

Shalev-Shwartz, S., 2012. Online learning and online convex optimization. *Foundations and Trends  in Machine Learning* 4, 107–194.

Shotton, J., Winn, J., Rother, C., Criminisi, A., 2006. Textonboost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation, in: *European Conference on Computer Vision (ECCV)*. Springer, pp. 1–15.

Sommer, C., Straehle, C., K  the, U., Hamprecht, F.A., 2011. Ilastik: Interactive learning and segmentation toolkit, in: *2011 IEEE International Symposium on Biomedical Imaging*, pp. 230–233.

Su, H., Xing, F., Kong, X., Xie, Y., Zhang, S., Yang, L., 2015. Robust cell detection and segmentation in histopathological images using sparse reconstruction and stacked denoising autoencoders, in: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, pp. 383–390.

Tao, K., Fang, M., Alroy, J., Sahagian, G., 2008. Imagable 4T1 model for the study of late stage breast cancer. *BMC Cancer* 8.

Tommasi, T., Orabona, F., Kaboli, M., Caputo, B., 2012. Leveraging over prior knowledge for online learning of visual categories, in: *British Machine Vision Conference*, pp. 87.1–87.11.

Vahadane, A., Peng, T., Albarqouni, S., Baust, M., Steiger, K., Schlitter, A., Sethi, A., Esposito, I., Navab, N., 2015. Structure-preserved color normalization for histological images, in: *IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1012–1015.

Veta, M., Pluim, J., van Diest, P., Viergever, M., 2014. Breast cancer histopathology image analysis: A review. *IEEE Transactions on Biomedical Engineering* .

Viola, P., Jones, M.J., 2004. Robust real-time face detection. *International Journal of Computer Vision* .

Xu, Y., Zhu, J.Y., Eric, I., Chang, C., Lai, M., Tu, Z., 2014. Weakly supervised histopathology cancer image segmentation and classification. *Medical Image Analysis* 18, 591–604.

Zhao, P., Hoi, S.C., 2010. OTL: A framework of online transfer learning, in: *International Conference on Machine Learning*, pp. 1231–1238.

Zikic, D., Glocker, B., Criminisi, A., 2014. Encoding atlases by randomized classification forests for efficient multi-atlas label propagation. *Medical Image Analysis* 18, 1262–1273.

Zinkevich, M., 2003. Online convex programming and generalized infinitesimal gradient ascent. *International Conference on Machine Learning* .

Appendix A. Rewriting the Scoring Function $\phi(R|\mathcal{F})$ as a Scalar Product (Eq. 3)

We expose here in details how Eq. 3 can be derived. This equality mainly comes from the linearity of both the scoring function ϕ and the aggregation of the tree predictions. Combining Eq. 2 and Eq. 1, one obtains

$$\phi(R|\mathcal{F}) = \frac{1}{n_{\text{trees}}} \sum_{\mathbf{x} \in R} \sum_{t=1}^{n_{\text{trees}}} \pi_{\sigma_t(\mathbf{x})} \quad (\text{A.1})$$

$$= \frac{1}{n_{\text{trees}}} \sum_{\mathbf{x} \in R} \sum_{t=1}^{n_{\text{trees}}} \sum_{L \in \mathcal{L}_t} \pi_L \mathbf{1}_{\{\sigma_t(\mathbf{x})=L\}}. \quad (\text{A.2})$$

We defined in Sec. 4.2.1 the quantity $\text{tree}(L) \in \{1, \dots, n_{\text{trees}}\}$ as the index of the tree to which a leaf $L \in \mathcal{L}$ belongs. Following this definition, we have for all trees $t \in \{1, \dots, n_{\text{trees}}\}$ and leaves $L \in \mathcal{L}_t$ the equality $t = \text{tree}(L)$. Thus

$$\sum_{t=1}^{n_{\text{trees}}} \sum_{L \in \mathcal{L}_t} \pi_L \mathbf{1}_{\{\sigma_t(\mathbf{x})=L\}} = \sum_{t=1}^{n_{\text{trees}}} \sum_{L \in \mathcal{L}_t} \pi_L \mathbf{1}_{\{\sigma_{\text{tree}(L)}(\mathbf{x})=L\}} \quad (\text{A.3})$$

$$= \sum_{L \in \mathcal{L}} \pi_L \mathbf{1}_{\{\sigma_{\text{tree}(L)}(\mathbf{x})=L\}} \quad (\text{A.4})$$

since the double sum $\sum_{t=1}^{n_{\text{trees}}} \sum_{L \in \mathcal{L}_t}$ amounts to summing over all leaves in the forest. Finally, by incorporating Eq. A.4 in Eq. A.2, we obtain

$$\phi(R|\mathcal{F}) = \frac{1}{n_{\text{trees}}} \sum_{\mathbf{x} \in R} \sum_{L \in \mathcal{L}} \pi_L \mathbf{1}_{\{\sigma_{\text{tree}(L)}(\mathbf{x})=L\}} \quad (\text{A.5})$$

$$= \sum_{L \in \mathcal{L}} \pi_L \left(\frac{1}{n_{\text{trees}}} \sum_{\mathbf{x} \in R} \mathbf{1}_{\{\sigma_{\text{tree}(L)}(\mathbf{x})=L\}} \right) \quad (\text{A.6})$$

$$= \sum_{L \in \mathcal{L}} \pi_L \rho_L(R|\Sigma) \quad (\text{A.7})$$

$$= \langle \rho(R|\Sigma), \pi \rangle \quad (\text{A.8})$$

using the definition of $\rho(R|\Sigma)$ given in Eq. 4.

Appendix B. Choice of Learning Rate (Eq. 9)

We expose here the theoretical considerations leading to the form of the learning rate exposed in Eq. 9. We follow a classical reasoning inspired from the online learning literature (Nemirovski et al., 2009; Shalev-Shwartz, 2012) and show how it relates to our scenario by expressing bounds in terms of the parameters of our method. As detailed in Sec. 4.2.2, at each iteration $k \geq 1$, the current set of leaf models π^k suffers the loss $l_k(\pi^k) = \left(\langle \rho(R_k), \pi^k \rangle - Q(R_k) \right)^2$, after which a new vector of leaf models π^{k+1} is chosen according to the online gradient descent update rule (Eq. 7). After T region suggestions ($T \geq 1$), the cumulated regret of having used the series of models π^1, \dots, π^T is defined as

$$\text{Regret}_T = \sum_{k=1}^T (l_k(\pi^k) - l_k(\pi^*)), \quad (\text{B.1})$$

where

$$\pi^* = \operatorname{argmin}_{\pi \in [0,1]^{|\mathcal{L}|}} \sum_{k=1}^T l_k(\pi) \quad (\text{B.2})$$

corresponds to the set of leaf models which would have incurred the smallest loss over the T iterations. The reasoning consists in computing an upper bound of Regret_T depending on the learning rate η . To do so, we use the fact that the functions l_k are convex, so that, for all k , we have $l_k(\pi^k) \leq l_k(\pi^*) + \langle \pi^k - \pi^*, \vec{\nabla} l_k(\pi^k) \rangle$ and thus

$$\text{Regret}_T \leq \sum_{k=1}^T \langle \pi^k - \pi^*, \vec{\nabla} l_k(\pi^k) \rangle. \quad (\text{B.3})$$

To find an upper bound of $A_k = \langle \pi^k - \pi^*, \vec{\nabla} l_k(\pi^k) \rangle$, we use the update rule of Eq. 7 as follows. For all k , denoting $\Pi = \Pi_{[0,1]^{|\mathcal{L}|}}$ and $D_k = \|\pi^k - \pi^*\|$, we have

$$D_{k+1}^2 = \|\pi^{k+1} - \pi^*\|^2 \quad (\text{B.4})$$

$$= \left\| \Pi \left[\pi^k - \eta \vec{\nabla} l_k(\pi^k) \right] - \pi^* \right\|^2 \quad (\text{B.5})$$

$$= \left\| \Pi \left[\pi^k - \eta \vec{\nabla} l_k(\pi^k) \right] - \Pi \left[\pi^* \right] \right\|^2 \quad (\text{B.6})$$

$$\leq \left\| \pi^k - \eta \vec{\nabla} l_k(\pi^k) - \pi^* \right\|^2 \quad (\text{B.7})$$

$$= D_k^2 - 2\eta A_k + \eta^2 \left\| \vec{\nabla} l_k(\pi^k) \right\|^2, \quad (\text{B.8})$$

which leads to the inequality

$$A_k \leq \frac{1}{2\eta} \left(D_k^2 - D_{k+1}^2 + \eta^2 \left\| \vec{\nabla} l_k(\pi^k) \right\|^2 \right). \quad (\text{B.9})$$

The inequality between Eq. B.6 and Eq. B.7 results from the fact that, in the Hilbert space $\mathbb{R}^{|\mathcal{L}|}$, performing a projection on the closed convex set $[0, 1]^{|\mathcal{L}|}$ does not increase the distance be-

tween two points. Using Eq. B.9 in Eq. B.3, we obtain

$$\text{Regret}_T \leq \frac{1}{2\eta} \sum_{k=1}^T \left(D_k^2 - D_{k+1}^2 + \eta^2 \left\| \vec{\nabla} l_k(\pi^k) \right\|^2 \right) \quad (\text{B.10})$$

$$= \frac{1}{2\eta} \left[D_1^2 - D_{T+1}^2 + \eta^2 \sum_{k=1}^T \left\| \vec{\nabla} l_k(\pi^k) \right\|^2 \right] \quad (\text{B.11})$$

$$\leq \frac{1}{2\eta} D_1^2 + \frac{\eta}{2} \sum_{k=1}^T \left\| \vec{\nabla} l_k(\pi^k) \right\|^2. \quad (\text{B.12})$$

To obtain a final bound on the regret, we need to find an upper bound of D_1^2 and of the norm of the gradient $\left\| \vec{\nabla} l_k(\pi^k) \right\|^2$. First, since both π^1 and π^* belong to $[0, 1]^{|\mathcal{L}|}$, we have

$$D_1^2 = \|\pi^1 - \pi^*\|^2 \leq |\mathcal{L}|. \quad (\text{B.13})$$

Secondly, for all k and π , we have

$$\vec{\nabla} l_k(\pi) = 2 \left(\langle \rho(R_k), \pi \rangle - Q(R_k) \right) \rho(R_k). \quad (\text{B.14})$$

The quantity $\langle \rho(R_k), \pi \rangle$ estimates the surface covered by positive pixels in the region R_k , while $Q(R_k)$ is the actual value of this surface revealed by the user. Since, by definition, both $\langle \rho(R_k), \pi^k \rangle$ and $Q(R_k)$ are comprised between 0 and the size δ^2 of the region, we have $\left| \langle \rho(R_k), \pi^k \rangle - Q(R_k) \right| \leq \delta^2$. By definition of ρ (see Eq. 4), we also know that each individual component ρ_L does not exceed $\frac{\delta^2}{n_{\text{trees}}}$ (since at most the number of pixels in the region δ^2 can fall in a leaf L), and, moreover, that these components sum to δ^2 . Thus

$$\|\rho(R_k)\|^2 = \sum_{L \in \mathcal{L}} \rho_L^2(R_k) \quad (\text{B.15})$$

$$\leq \frac{\delta^2}{n_{\text{trees}}} \sum_{L \in \mathcal{L}} \rho_L(R_k) \quad (\text{B.16})$$

$$= \frac{\delta^4}{n_{\text{trees}}}, \quad (\text{B.17})$$

hence the following upper bound on the gradient:

$$\left\| \vec{\nabla} l_k(\pi^k) \right\|^2 \leq 4 \frac{\delta^8}{n_{\text{trees}}}. \quad (\text{B.18})$$

Finally, including Eq. B.13 and Eq. B.18 in Eq. B.12 gives the bound

$$\text{Regret}_T \leq \frac{|\mathcal{L}|}{2\eta} + \frac{2\eta T \delta^8}{n_{\text{trees}}}. \quad (\text{B.19})$$

We choose the value of η providing the best regret bound, i.e. minimizing the right side of Eq. B.19. This is obtained for

$$\eta = \frac{1}{2\delta^4} \sqrt{\frac{n_{\text{trees}} |\mathcal{L}|}{T}}. \quad (\text{B.20})$$

While the relevant number of iterations T for the practical applicability of our scenario is unknown, it should at least be proportional to the size of the test slide, and thus to the number of regions $|\mathcal{R}|$. This leads us to define $T = \lambda^2 |\mathcal{R}|$ as proportional to this quantity, resulting in Eq. 9, and learn the hyperparameter λ on a validation set.

Leveraging Random Forests for Interactive Exploration of Large Histological Images*

Loïc Peter¹, Diana Mateus^{1,2}, Pierre Chatelain^{1,3}, Noemi Schworm⁴, Stefan Stangl⁴, Gabriele Multhoff^{4,5}, Nassir Navab^{1,6}

¹ Computer Aided Medical Procedures, Technische Universität München, Germany

² Institute of Computational Biology, Helmholtz Zentrum München, Germany

³ Université de Rennes 1, IRISA, France

⁴ Department of Radiation Oncology, Technische Universität München, Germany

⁵ CCG - Innate Immunity in Tumor Biology, Helmholtz Zentrum München, Germany

⁶ Computer Aided Medical Procedures, Johns Hopkins University, USA

Abstract. The large size of histological images combined with their very challenging appearance are two main difficulties which considerably complicate their analysis. In this paper, we introduce an interactive strategy leveraging the output of a supervised random forest classifier to guide a user through such large visual data. Starting from a forest-based pixelwise estimate, subregions of the images at hand are automatically ranked and sequentially displayed according to their expected interest. After each region suggestion, the user selects among several options a rough estimate of the true amount of foreground pixels in this region. From these one-click inputs, the region scoring function is updated in real time using an online gradient descent procedure, which corrects on-the-fly the shortcomings of the initial model and adapts future suggestions accordingly. Experimental validation is conducted for extramedullary hematopoiesis localization and demonstrates the practical feasibility of the procedure as well as the benefit of the online adaptation strategy.

1 Introduction

Analyzing histological images is usually an extremely challenging task. Due to the complex appearance of objects of interest, the accuracy of fully-automatic techniques is generally insufficient for clinical use, e.g. in the case of mitosis detection [1] for which a variety of automatic approaches has been recently compared quantitatively [2]. In fact, the accurate identification of patterns within such images is often only achievable by well-trained human experts and remains prone to inter-experts disagreements in some cases [3]. Moreover, histological data are very tedious to process for a human because of their large dimension, which commonly reaches tens of thousands of pixels along each direction. When aiming at finding rare objects within images of this size, a manual search for these instances requires a painstaking exploration of the whole content, and a huge amount of time is spent scrolling through uninteresting background areas.

To overcome this, we propose to leverage fully-automatic pixelwise classification techniques to recover candidate areas of interest (i.e. where positive instances are found), leaving the task of interpreting the visual content within these regions to an expert user. To this end, assuming that an available automatic method provides a label confidence at each pixel, an intuitive strategy would be to rank subregions of the images in decreasing order of confidence and display them accordingly one after the other. Having thereby patterns of interest shown early in the process, an exhaustive search would not be necessary anymore. However, such a scenario relies entirely on the accuracy of the automatic detector and can fail if the latter suffers from difficulties on the images at hand. In the case of detectors based on supervised learning procedures, difficulties may arise due to the common mismatches between the training and test data, caused by e.g. differences in noise or illumination. In this context, it would be beneficial to allow the user

* The final publication is available at http://link.springer.com/chapter/10.1007%2F978-3-319-10404-1_1

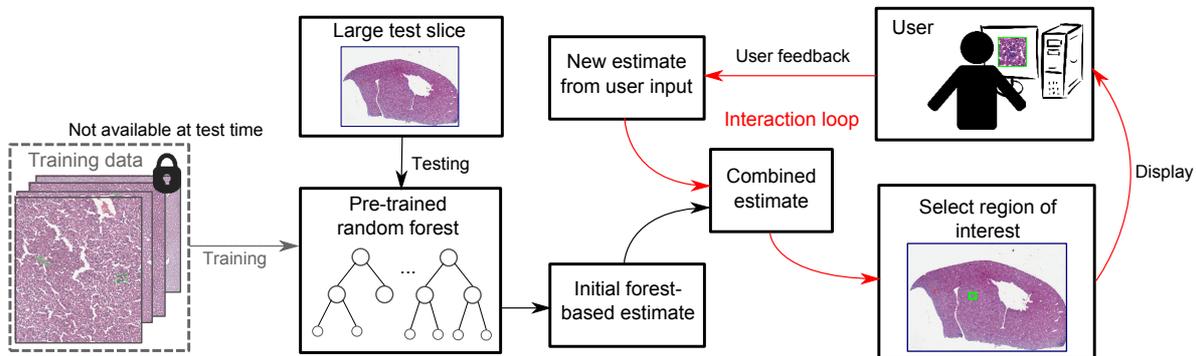


Fig. 1. Summary of our approach

to interact after each suggestion (in particular after wrong ones) so that the procedure can be adapted to the characteristics of the test data and eventually avoid mistakes for the upcoming region suggestions.

Online transfer learning (OTL) frameworks [4,5] address this problem by combining (i) a kernel-based classifier pre-trained on some source data, and (ii) an online classifier continuously trained on arriving instances from the target data, given by the user input in our case. On the closely related problem of domain adaptation, an unsupervised algorithm based on Gaussian process regression was introduced to adapt the decision boundary of a face detector to a test image [6]. In the case of medical data, random forest classifiers [7,8] demonstrated high accuracy and tractability and are especially used for their ability to handle high amounts of training data and descriptors. Although an online version of random forests exists in the literature [9], replacing the kernel classifiers of the OTL approach by forests is not straightforward: only few samples are available to train the target classifier, and a forest-based OTL scheme would handle the source classifier as a whole, without exploiting the diversity and possibly heterogeneous relevance of its local models on the target data.

In this work, we introduce a regression-based method able to leverage the knowledge carried by a pre-trained random forest classifier to guide the user through the data by suggesting candidate regions of interest. Our approach exploits the multiple partitionings of the feature space defined by the trained forest to perform local updates of the model from one-click inputs provided by the user. Thereby, characteristics of the test data can be captured and future suggestions are adapted accordingly. These updates are performed in real time between two iterations and do not require the original training data, which makes our method particularly tractable for clinical use. Experimental evaluation is conducted on a high-resolution dataset for hematopoiesis identification within mouse liver slices, as well as on synthetic deformations of it, and demonstrates the general feasibility of our approach and the benefit of the adaption scheme.

2 Methods

We first summarize the general scenario of our approach. Let us consider a given set of images composed of pixels. Each pixel \mathbf{p} has a true label $y(\mathbf{p}) \in \{0, 1\}$ and we aim at detecting the positive instances within our images, i.e. the pixels \mathbf{p} such that $y(\mathbf{p}) = 1$. We make the assumption that a random forest classifier has been trained beforehand on some available labeled data and hence provides a probability $P^{\text{RF}}(y(\mathbf{p}) = 1) \in [0, 1]$ for each pixel \mathbf{p} . Here, we implicitly assume that this forest has been originally designed for a binary classification task, but any multi-class forest could also be used by grouping as positive the labels considered as interesting for the given application. For tractability, the original training data is not made available during the whole process and is thus only encoded through this forest. In our scenario, a user runs this classifier which is based on prior data on a new set of images that are correlated but potentially slightly different, e.g. because of different imaging conditions. From the resulting pixelwise

output, an area of interest, expected to contain positive labels, is then displayed to the user, who provides in return an estimate of the proportion of positive pixels actually observed. In doing so, the bias of the initial model can be progressively assessed and compensated to ultimately increase the relevance of the upcoming region suggestions.

Under these conditions, we can formalize our scenario mathematically. The test images are partitioned into a set of rectangular regions $\mathcal{R} = \{R_1, \dots, R_n\}$ whose sizes are tractable for a human user. At iteration k , i.e. after the user saw $k - 1$ regions, we proceed as follows:

1. The region $\hat{R} \in \mathcal{R}$, maximizing a scoring function $\phi_k : \mathcal{R} \rightarrow \mathbb{R}$ stating the expected relevance of each region, is displayed to the user.
2. The user interacts to provide information about the content observed in \hat{R} .
3. Using this feedback, ϕ_k is updated, resulting in a new scoring function ϕ_{k+1} .
4. \hat{R} is removed from the pool of candidate regions \mathcal{R} .

The initial scoring function ϕ_1 is based on the initial random forest classifier only. We summarize the random forest model in Sec. 2.1. In Sec. 2.2, we expose our choice of scoring functions ϕ_k and how they can be parametrized. Finally, the update strategy of ϕ_k from the user inputs is described in Sec. 2.3.

2.1 Random Forest Model

Our approach uses a trained random forest classifier as starting point, which provides for each pixel \mathbf{p} a probability of being a positive instance. A random forest is a collection of T decorrelated binary decision trees. Each tree (indexed by $t \in \{1, \dots, T\}$) is a hierarchical collection of decision rules based on visual features, leading to a partition of the domain into $l(t)$ leaves $\mathcal{L}_{t,1}, \dots, \mathcal{L}_{t,l(t)}$. Each leaf $\mathcal{L}_{t,i}$ contains a probability model $\pi_{t,i} \in [0, 1]$. For each tree t , every newly observed pixel \mathbf{p} reaches exactly one leaf, and we will denote $\pi_t(\mathbf{p})$ the corresponding probabilistic model. Averaging over trees, the final pixelwise probability provided by the random forest for each pixel \mathbf{p} is

$$P^{\text{RF}}(y(\mathbf{p}) = 1) = \frac{1}{T} \sum_{t=1}^T \pi_t(\mathbf{p}). \quad (1)$$

We follow the standard procedure [7,8] to train such a classifier, i.e. to infer the decision rules and the leaf models. Starting from a root node, the decision rules are recursively chosen within a set of randomly drawn splitting functions to maximize an information gain criterion. The proportion of positive training samples contained in a leaf defines its probabilistic model.

2.2 Region Scoring Function

In this subsection, we describe our region scoring model. Given a pixelwise estimate $P_k(y(\mathbf{p}) = 1)$ at iteration k , we define the relevance of a region $R \in \mathcal{R}$ as the expectation of the proportion of positive pixels in R . Modeling the pixel labels as independent Bernoulli distributions of success probability $P_k(y(\mathbf{p}) = 1)$, this expectation and hence the scoring function ϕ_k can be written as

$$\phi_k(R) = \frac{1}{|R|} \sum_{\mathbf{p} \in R} P_k(y(\mathbf{p}) = 1), \quad (2)$$

where $|R|$ denotes the number of pixels in the region R . In this work, we would like to define the pixelwise probabilistic model P_k as a combination of (i) the prior knowledge acquired on training data that is encoded by the initial random forest classifier, and (ii) the $k - 1$ inputs provided by the user based on the test data. Therefore, we propose to obtain $P_k(y(\mathbf{p}) = 1)$ from the forest estimate of Eq. 1 by adding to each prior leaf model $\pi_{t,i}$ a signed offset $\epsilon_{t,i}^k \in \mathbb{R}$ whose value will be progressively adapted from the

user inputs. Let us denote ϵ_k the vector $(\epsilon_{t,i}^k)_{t,i}$ and rewrite the scoring function $\phi_k(R)$ as $\phi(R|\epsilon_k)$ so that this parametrization clearly appears. Introducing this additive model in Eq. 2 and after some algebraic manipulations, the scoring function can be rewritten as

$$\phi(R|\epsilon_k) = \phi^{\text{RF}}(R) + \phi^{\text{new}}(R|\epsilon_k), \quad (3)$$

where $\phi^{\text{RF}}(R)$ is obtained by using the forest probabilistic output (Eq. 1) as pixelwise model in Eq. 2, and

$$\phi^{\text{new}}(R|\epsilon_k) = \frac{1}{|R|} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{l(t)} \epsilon_{t,i}^k h_{t,i}(R) = \langle \mathbf{r}, \epsilon_k \rangle. \quad (4)$$

In Eq. 4, $h_{t,i}(R)$ denotes the number of pixels \mathbf{p} in R which reached the leaf $\mathcal{L}_{t,i}$, and $\mathbf{r} = (\frac{1}{|R|} \frac{1}{T} h_{t,i}(R))_{t,i}$. Before the user starts providing inputs, the $\epsilon_{t,i}^1$ are set to 0 such that $\phi_1 = \phi^{\text{RF}}$. As the user gives information after each suggestion, the acquired knowledge is encoded through the leaf-dependent $\epsilon_{t,i}^k$ that allow the scoring function ϕ_k to vary around the initial model ϕ^{RF} . The next subsection describes how these $\epsilon_{t,i}^k$ are inferred from the collected user feedbacks.

2.3 Online Model Update

After each suggestion of a region R_j , a feedback $\mathcal{P}^{\text{user}}(R_j) \in [0, 1]$ stating the true proportion of positive pixels in R_j is requested from the user. We assume for the sake of formalism that the exact proportion is given, but we will demonstrate in our experiments that providing a discretized approximation of this input does not decrease the performance (Tab. 1). This reduces drastically the necessary amount of interaction, since giving a discrete estimate of this proportion requires only one click from the user, whereas computing the exact value would involve a pixelwise region labeling. Note that this feedback does not provide any information at pixel level and hence excludes a direct update of the leaf statistics [9]. Assuming that k regions R_1, \dots, R_k have been suggested and their respective user inputs $\mathcal{P}^{\text{user}}(R_1), \dots, \mathcal{P}^{\text{user}}(R_k)$ collected ($k \geq 1$), we would like to find a new scoring function ϕ_{k+1} , or equivalently a coefficient vector ϵ_{k+1} , which matches with these k user inputs. Considering a squared loss function, we solve the following minimization problem

$$\epsilon_{k+1} = \underset{\epsilon}{\operatorname{argmin}} \sum_{j=1}^k (\phi(R_j|\epsilon) - \mathcal{P}^{\text{user}}(R_j))^2 + \lambda \|\epsilon\|^2, \quad (5)$$

which can be rewritten from Eq. 3 and Eq. 4 as a linear ridge regression problem

$$\epsilon_{k+1} = \underset{\epsilon}{\operatorname{argmin}} \sum_{j=1}^k (\langle \mathbf{r}_j, \epsilon \rangle - \delta_j)^2 + \lambda \|\epsilon\|^2, \quad (6)$$

where $\delta_j = \mathcal{P}^{\text{user}}(R_j) - \phi^{\text{RF}}(R_j)$. The regularization term aims at keeping the deviation from the initial random forest classifier small and is driven by a hyper-parameter λ . The minimization problem defined by Eq. 6 is equivalent to a linear ridge regression problem with k training instances of feature vectors $\mathbf{r}_1, \dots, \mathbf{r}_k$ and labels $\delta_1, \dots, \delta_k$, i.e. a regression problem at the region level. As a consequence, Eq. 6 could be naturally extended to a kernel ridge regression problem if one desires to remove the linearity assumption. The solution ϵ_{k+1} of Eq. 6 can be found in closed form [10]. However, computing this exact value at each iteration involves the computationally expensive inversion of a large matrix, which could force a user to wait between two suggestions of regions. Instead, we adopt an online gradient descent scheme [11] leading to the following incremental update rule:

$$\epsilon_{k+1} = \epsilon_k - \frac{1}{\lambda} (\langle \mathbf{r}_k, \epsilon_k \rangle - \delta_k) \mathbf{r}_k. \quad (7)$$

Thereby, the k^{th} user input can be incorporated in real time in our model.

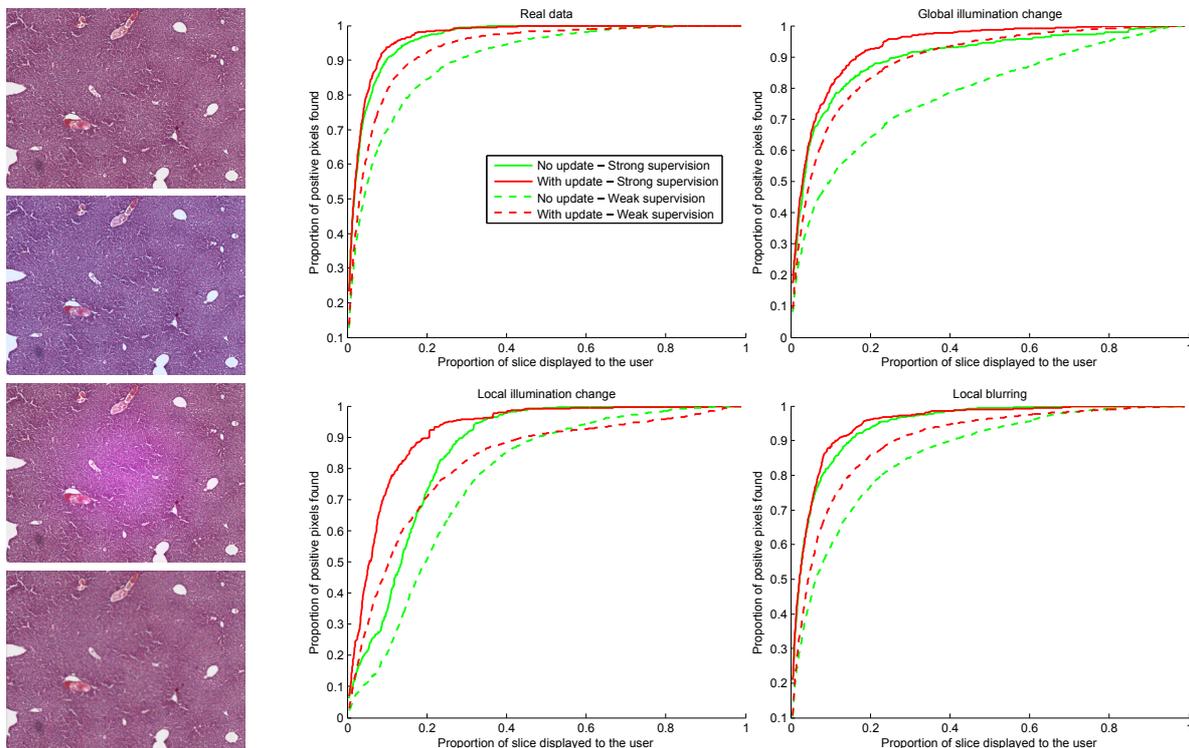


Fig. 2. Results The curves display the average proportion of recovered positive pixels as a function of the proportion of the slice shown to the user. These experiments are driven under different synthetic conditions, examples of which are shown on the left. From top to bottom: real image, global illumination change, local illumination change, local blurring. The discretized version of the user input was used in these experiments.

3 Experiments

Dataset: Extramedullary hematopoiesis, i.e. hematopoiesis located outside of the bone marrow, is a rare event in human adults. Yet, in some animals like mice, its presence in the liver can be a reaction to an immense inflammation or to a tumor stimulating extensively the immune system of the organism [12]. Localizing these relatively small lesions as well as estimating their number can be useful to compare the influence of different therapeutic strategies. 34 high-resolution images entirely labeled at pixel level were extracted from 8 mouse liver slices of resolution $0.498 \mu\text{m}/\text{pixel}$. This restriction to subimages was necessary to obtain accurate labels from a pathologist, since a slice is typically too big (34200×21052 pixels) to be labeled entirely. Since images taken from a same slice share common visual properties, we kept them together for training or testing. Images were downsized by a factor 3 to speed up the training and testing procedures.

Experimental Setup: We performed our experiments at two levels of supervision: (i) a standard leave-one-out strategy (*strong supervision*), where the initial random forest is trained on 7 slices (i.e. approximately 30 images) and tested on the 8th one, and (ii) the inverse setting where one trains only on one slice and tests on the 7 others (*weak supervision*). While the second scenario is plausible in practical situations due to the difficulty of obtaining pixelwise labels, it also suggests another application of our approach as a way to facilitate collecting hard-to-find positive samples to enrich an existing weak training set. Training samples were collected on a regular grid of step 30, resulting in approximately 500 k instances in the leave-one-out

Table 1. Area under the curves in our different settings.

Weak supervision												
	Real data			Global change			Local change			Local blurring		
	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max
No update	82.5	90.9	95.9	63.8	77.0	95.7	56.2	76.4	81.8	78.7	83.7	93.8
Exact input	87.6	92.2	96.1	74.4	87.8	95.7	62.9	82.8	91.6	78.0	86.1	94.1
Discretized input	88.5	92.3	96.1	74.6	87.7	95.9	64.0	83.9	91.6	77.9	86.3	94.1

Strong supervision												
	Real data			Global change			Local change			Local blurring		
	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max
No update	94.1	96.6	97.7	71.1	93.1	97.1	78.5	85.4	92.7	91.2	95.4	96.0
Exact input	94.7	96.7	98.1	89.1	93.6	97.2	83.2	91.9	95.0	92.4	95.8	97.1
Discretized input	94.8	96.7	98.3	88.9	93.7	97.4	82.8	91.9	95.0	92.4	95.6	97.0

setup. Each pixel was described by its visual content at offset locations. We used the CIELab color space that we enriched with the output of a bank of filters [13]. 40 trees were grown without depth limit but with at least 10 training samples per leaf. The images were partitioned into regions of size 123×123 pixels, which fits to the size of usual hematopoiesis patterns. If a hematopoiesis pattern is not fully included in a displayed region, the field of view can be extended with a simple click revealing the next adjacent region in the indicated direction. We investigated ways to handle pools of regions of heterogeneous sizes, but this raised issues in terms of scoring function (finding a reliable rule that does not always favor the smallest or the largest regions is not straightforward) and evaluation. The retained solution requires almost no effort from the user and eventually seemed to be the most practical one.

Evaluation: Our approach aims at performing an interactive online adaptation to new imaging conditions, and more generally to mismatches between training and test data. The aforementioned weakly supervised scenario favors naturally such a bias, since training on only one slice is unlikely to generalize well. To go further into this direction, we propose, as additional experiments, to keep the initial training sets and apply three types of synthetic modifications on the test images to simulate classical difficulties occurring in histological imaging: (i) global illumination change, (ii) local illumination change, and (iii) local blurring. These changes were randomly generated on each image. To evaluate our method, we studied the average proportion of positive pixels recovered over time (Fig. 2) and measured the area under these curves (Tab. 1). The results demonstrate that our update framework allows a faster retrieval of interesting areas. As mentioned in Sec. 2.3, we also show that discretizing the user input into 7 predefined bins (i.e. enforcing $P^{\text{user}}(R_j) \in \{0, 0.1, 0.3, \dots, 0.9, 1\}$) does not reduce the performance and thus lightens the user interaction. Finally, in the leave-one-out setting on real data, we assessed the influence of the parameter λ (Eq. 6). From Eq. 7, low (resp. high) values of λ are expected to lead to numerical instabilities (resp. negligible updates). These asymptotic behaviors were experimentally observed for $\lambda \leq 0.05$ and $\lambda \geq 100$, respectively. Between these two extremes, a wide range of values of λ led to improvements, and we chose $\lambda = 1$ for simplicity.

4 Conclusion

We proposed an interactive method able to suggest sequentially regions of interest within large data from the output of a pre-trained random forest classifier at pixel level. After each suggestion, one-click user inputs are collected to capture the specificities of the images at hand and update the region scoring function accordingly and in real time *via* an online gradient descent scheme. We evaluated our approach in the context of hematopoiesis localization, with and without synthetic deformations simulating common sources of variability in histological imaging. As future work, we would like to extend our method to mitosis detection tasks [2] and generalize it to 2D video sequences and 3D volumes.

Acknowledgments This work was partially supported by the DFG-funded Collaborative Research Centre 824: “Imaging for Selection, Monitoring and Individualization of Cancer Therapies”.

References

1. Veta, M., Pluim, J., van Diest, P., Viergever, M.: Breast cancer histopathology image analysis: A review. *IEEE Transactions on Biomedical Engineering* (2014)
2. <http://amida13.isi.uu.nl/>
3. Crowley, R.S., Naus, G.J., Stewart, J., Friedman, C.P.: Development of visual diagnostic expertise in pathology - an information-processing study. *Journal of the American Medical Informatics Association* **10**(1) (2003) 39–51
4. Zhao, P., Hoi, S.C.: OTL: A framework of online transfer learning. In: *Proceedings of the 27th International Conference on Machine Learning*. (2010) 1231–1238
5. Tommasi, T., Orabona, F., Kaboli, M., Caputo, B.: Leveraging over prior knowledge for online learning of visual categories. In: *BMVC*. (2012)
6. Jain, V., Learned-Miller, E.: Online domain adaptation of a pre-trained cascade of classifiers. In: *CVPR*. (2011) 577–584
7. Breiman, L.: *Random forests*. *Machine Learning* (2001)
8. Criminisi, A., Shotton, J.: *Decision Forests for Computer Vision and Medical Image Analysis*. Springer (2013)
9. Saffari, A., Leistner, C., Santner, J., Godec, M., Bischof, H.: On-line random forests. In: *ICCV Workshops*. (Sept 2009) 1393–1400
10. Hastie, T., Tibshirani, R., Friedman, J.: *The elements of statistical learning: data mining, inference and prediction*. 2 edn. Springer (2009)
11. Shalev-Shwartz, S.: Online learning and online convex optimization. *Found. Trends Mach. Learn.* **4**(2) (February 2012) 107–194
12. Tao, K., Fang, M., Alroy, J., Sahagian, G.: Imagable 4t1 model for the study of late stage breast cancer. *BMC Cancer* **8**(1) (2008)
13. Winn, J., Criminisi, A., Minka, T.: Object categorization by learned universal visual dictionary. In: *ICCV*. (2005)